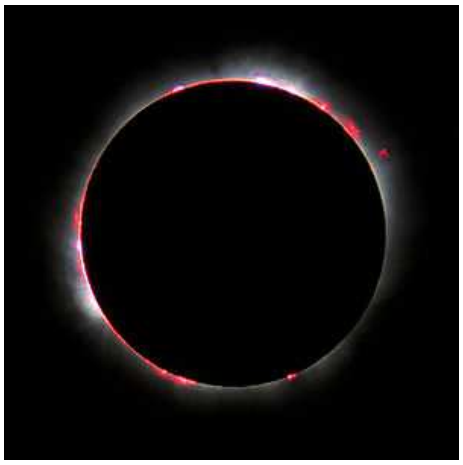
Ground-Based Solar Physics in the Era of Space Astronomy

A White Paper Submitted to the 2012 Heliophysics Decadal Survey

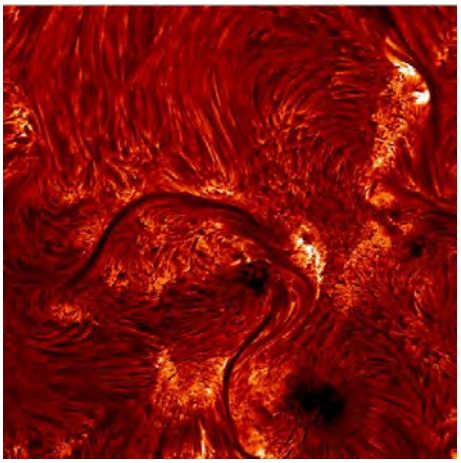
T. Ayres[1], D. Longcope[2]

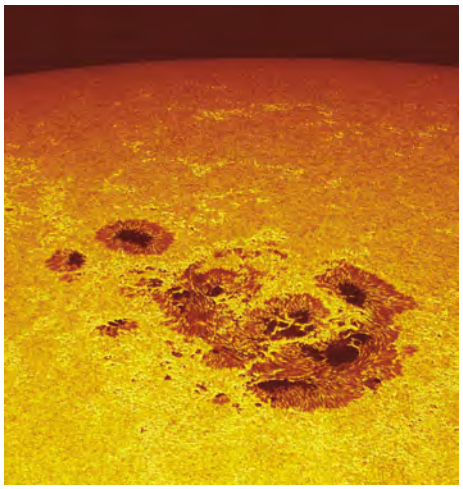
(on behalf of the 2009 AURA Solar Decadal Committee)

Chromosphere-Corona at eclipse

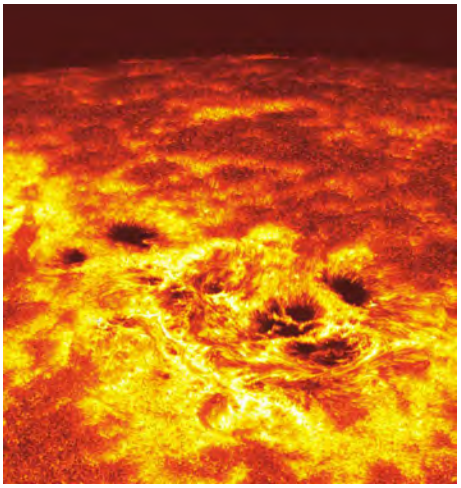


Hα filtergram





Photospheric spots & bright points



Same area in chromospheric Ca^+

[1]Center for Astrophysics and Space Astronomy, 389 UCB, University of Colorado, Boulder, CO 80309; Thomas.Ayres@Colorado.edu (corresponding author)

[2]Montana State University

SUMMARY. *A report, previously commissioned by AURA to support advocacy efforts in advance of the Astro2010 Decadal Survey, reached a series of conclusions concerning the future of ground-based solar physics that are relevant to the counterpart Heliophysics Survey. The main findings: (1) The Advanced Technology Solar Telescope (ATST) will continue U.S. leadership in large aperture, high-resolution ground-based solar observations, and will be a unique and powerful complement to space-borne solar instruments; (2) Full-Sun measurements by existing synoptic facilities, and new initiatives such as the Coronal Solar Magnetism Observatory (COSMO) and the Frequency Agile Solar Radiotelescope (FASR), will balance the narrow field of view captured by ATST, and are essential for the study of transient phenomena; (3) Sustaining, and further developing, synoptic observations is vital as well to helioseismology, solar cycle studies, and Space Weather prediction; (4) Support of advanced instrumentation and seeing compensation techniques for the ATST, and other solar telescopes, is necessary to keep ground-based solar physics at the cutting edge; and (5) Effective planning for ground-based facilities requires consideration of the synergies achieved by coordination with space-based observatories. Additional findings — more broadly relevant to solar physics as a whole — involved workforce and training issues, and the support of theory and modeling. A final recommendation considered the possible transfer of the National Solar Observatory and its facilities from NSF's Math & Physical Sciences directorate to Geosciences, urging that full consideration be given to ensuring that key intellectual and material links of solar to broader astrophysics are not broken. We also take this opportunity to advocate the solar-stellar connection, especially its role in understanding aspects of our star that are difficult to capture with contemporary measurements, such as the post-T-Tauri stage, relevant to the evolution of primitive planetary atmospheres, and the far distant future of the Sun, and what it might hold for the ultimate fate of the solar system.*

Background

In 2009, a report[1] concerning the future of ground-based solar physics was commissioned by the Association of Universities for Research in Astronomy (AURA), which manages the National Solar Observatory on behalf of the National Science Foundation. The report considered the standing of ground-based solar in an era when funding for heliosciences by NASA greatly eclipses that by the NSF for solar facilities on the ground and competitive grants. The report was written before construction of NSO's premier new ground-based facility — the 4m Advanced Technology Solar Telescope — had been approved by the National Science Board, and MREFC funding for the initial phases had been secured. The intent of the report was to inform AURA's efforts to advocate the value of ground-based solar physics, especially at the NSO and particularly the ATST, to the NRC's Astronomy & Astrophysics Decadal Survey. The report never was formally submitted as a white paper to Astro2010, however, and much of its impetus evaporated when the ATST construction was approved, shortly before the deliberations of the Astro2010 panels began, and the ATST then was removed from reconsideration (and possible re-ranking).

At the same time, the findings and recommendations of the report are pertinent to the new 2012 Heliophysics Decadal Survey. It is our intent here to provide a synopsis of those

[1]See: http://www.aura-astronomy.org/news/Future of Ground Based Solar Physics.pdf

results, on behalf of the original AURA committee.

We also take the opportunity to discuss the role of the "solar-stellar connection" in the heliosciences context. It is a multidisciplinary effort to, on the one hand, provide a conduit of solar knowledge to the broader astrophysical community; and on the other, passing stellar lore back to solar physics to address aspects of the Sun — especially historical and futuristic — that are not easily accessible with contemporary measurements.

Solar Physics: Support, Demographics, and Facility Usage

Before delving into the specific recommendations of the AURA report, it is worth mentioning some of the background information that was woven through that narrative (and in other solar-related Astro2010 white papers), as a backdrop for those recommendations.

Solar physics plays a unique role in astronomy. It is the only field in classical astronomy that can claim significant societal relevance, extending from the obvious direct influence of the solar luminosity on planetary habitability, to the more subtle impacts of Space Weather on human civilization and commerce. Solar physics also is unique in contemporary stellar astronomy in having a next-door view of its (albeit singular) subject, even to the extent of directly capturing particulate matter in the extreme outer limits of the solar atmosphere (i.e., the coronal wind). The uniqueness of solar physics extends to its (divided) support at the Federal level, to the demographics of its adherents, and to the facilities upon which it relies.

Unlike classical astronomy, which is funded almost exclusively by NASA Astrophysics and the AST division of NSF's Math & Physical Sciences directorate, most solar grant support and that for the High Altitude Observatory comes out of the AGS division in a separate directorate, Geosciences. At the same time, facility support — the National Solar Observatory and its two current observing sites — comes almost exclusively from the AST division, but few solar-related competitive grants are awarded from that source. Similarly, NASA supports solar physics almost entirely through its Heliophysics division, the focus of the present Decadal survey. Solar facilities governed by NSF's AST division are assessed separately in the decadal survey covering NASA's Astronomy portfolio. Furthermore, NSF AST conducted a "Senior Review" of facilities several years ago, which was applied to NSO together with the classical dark-side observatories, but independently of the part of solar physics funded through the AGS division.

Most self-identified solar physicists are located at federally-funded research centers such as NSO, HAO, the Naval Research Lab, NASA Goddard and Marshall Space Flight Centers, and in for-profit organizations such as the Lockheed-Martin Palo Alto Research Labs and the Southwest Research Institute. In contrast to other areas of astronomy, there are relatively few solar physicists on the faculties of major research universities. On the one hand, the concentration of solar astronomers in a few large organizations achieves an important intellectual critical mass that would be lacking in a typical Astronomy department hosting one or a few solar faculty. On the other hand, the diffuseness of solar presence in traditional research universities limits access to students, and makes the training of the next generation of solar astronomers problematic and a cause for concern. Strong REU programs at the national solar centers can allay the concerns somewhat, but does not change the fact that

solar training at the top research universities has been neglected. The U.S. situation differs from that in Europe, for example, where solar research groups typically have strong ties with university departments, and consequently many of the top solar researchers in the U.S. today have had their graduate training elsewhere.

Another unique aspect of solar physics is its ability to make occasional use of ground-based facilities designed primarily for night-time astronomy. Examples include the Very Large Array, which operates in the centimetric radio band sensitive to flare processes, and the James Clerk Maxwell Telescope (draped by a teflon sheet to protect the dish from direct solar heating) to image the Sun at sub-mm wavelengths. It is rarer for the exchange to go in reverse, although the historical night-time program at NSO's McMath-Pierce should be acknowledged. The Atacama Large Millimeter Array (ALMA), currently under construction in Chile, promises to be a powerful tool for solar astronomy when it is commissioned later this decade; as will the recently completed Expanded VLA. The awkwardness here is that while these ground-based night-time facilities are potentially quite valuable to solar astronomers, the reviews of the existing facilities (and proposals for new facilities and instruments) historically have been conducted in a forum where solar interests are only lightly represented, if at all (as in the recent Astro2010 Decadal Survey). To be fair, the flagship facility for ground-based solar — the 4m Advanced Technology Solar Telescope (ATST) — was the outcome of the 2000 Astronomy decadal review, so it cannot be said that solar has fared badly in the past. Nevertheless, it is important in the future that a mechanism exists to allow solar interests to be represented in discussions of night-time facilities and instruments that might have unique applications for day-time astronomy.

A major theme of this white paper is the role of ground-based observations in the era of space astronomy, wherein literally billions of dollars have been invested in a series of increasingly sophisticated orbiting solar observatories that have revolutionized our view of the Sun over the past four decades. These sums dwarf the $300M to be spent on NSO's flagship ATST, the major expenditure on solar-related ground-based experiments in the foreseeable future. When we can obtain uninterrupted sequences of $0.''1$ resolution narrow band images from *Hinode,* for example, or high quality vector magnetograms and helioseismic measurements 24/7 from HMI on the Solar Dynamics Observatory, is there even any need for ground-based observations in the future? The AURA committee concluded, "Of course there is the need!" Here are a few of the more compelling reasons:

(1) High-precision spectropolarimetry of the smallest-scale — dominantly magnetic — features in the solar atmosphere requires a large-aperture telescope to collect sufficient photon flux to ensure accurate measurements of the Stokes parameters at many positions across the line profile in a time short enough that the scene has not changed significantly due to plasma motions. For many reasons, principally cost, the aperture of the largest ground-based telescope always will exceed that of the largest orbiting telescope, solar or otherwise.

(2) Ground-based facilities can be easily maintained, repaired and upgraded; whereas most space observatories (with the conspicuous exception of *Hubble*) are subject to unrecoverable failures that can limit their useful lifetimes. Because solar space experiments tend to be of a one-off nature, there usually are no backup facilities to replace a lost unique capability. Further, the long lead times for design, approval, development and fabrication of space ex-

periments, and the restriction to flight-qualified hardware, limits the amount of advanced technology that can be incorporated. And once the experiment has reached orbit, upgrading is not an option. Ground-based experiments, on the other hand, can be supplied with state-of-the-art technology to respond agilely to new opportunities. For example, the venerable GONG helioseismic network could be retrofitted to observe multiple spectral features, to provide a sharper view of the solar interior beyond that achieved by a single-channel instrument (like HMI).

(3) While space observatories provide unique access to parts of the electromagnetic spectrum invisible to ground-based experiments — gamma and X-rays, and the ultraviolet — there are other key parts of the spectrum for which orbiting platforms have little, or no, advantage. For example, the thermal IR — rich in molecular diagnostics — requires a large aperture telescope to achieve diffraction limited performance, and, as mentioned earlier, we are unlikely to see such an instrument in orbit anytime soon. The sub-mm to metric radio bands are valuable diagnostics of thermal structure in the chromosphere and of nonthermal particle acceleration in flares, but again the large antenna arrays required to collect the radiation, and interferometrically image the Sun, are impractical for space deployment. In fact, there would be little advantage of the space environment in the first place, except perhaps for the very short sub-mm range, since the atmosphere is relatively transparent throughout the radio band and seeing is not an issue. A final example is the observation of solar neutrinos, whose detectors require shielding from cosmic rays by a kilometer or so of solid rock: an experiment clearly ill-suited for Earth orbit.

(4) Space weather encompasses a wide range of time scales, especially decadal, and longer, variations associated with magnetic cycles. The study of such long-term phenomena requires a consistent base of synoptic observations. The collection of synoptic data is rarely accomplished from space, however, due to undependable access, high cost, and NASA's choice to focus on novel science. In contrast, ground-based facilities have produced such invaluable data sets as the 400-year record of sunspot numbers, 8 decades of solar images from Mount Wilson, and the long-term magnetogram series from NSO's Kitt Peak facility. NSO is in the process of consolidating its major synoptic components — GONG and SOLIS — into a single coherent program with the intent to maintain a long-term magnetic — and full-disk imaging — context for future solar research.

We hope that these few examples clearly demonstrate that ground-based observations of the Sun are not merely an adjunct to the space-based counterparts, but rather a full partner in the broad enterprise of solar physics.

Specific Recommendations for Ground-Based Solar Physics

(1) Advanced Technology Solar Telescope

At the time of writing the AURA report, NSF construction funding for ATST was not yet approved, and there was concern within the community that its premier ground-based project would be reevaluated in the Astro2010 Survey, and possibly re-ranked to the detriment of its funding prospects. Thankfully, the re-consideration did not occur, but in preparation for that possibility, the AURA committee placed ATST at the top of its recommendations. In an era when 1.5m solar telescopes are beginning to replace historical 50–70cm aperture in-

struments, the leap to 4m, while technically challenging, promises to maintain, if not restore, U.S. leadership in ground-based solar physics. Many non-solar astronomers misinterpret the need for ATST, imagining that it is the raw diffraction-limited resolution that is important. Rather, it is the light gathering power of the large aperture that will be essential to support the spectropolarimetry of the smallest scale magnetic features that are at the heart of many contemporary solar puzzles. When one divides the incoming light in a spatial pixel into a half dozen or more polarization states, and over a dozen or so frequencies in the line profile, at a cadence fast enough to avoid velocity smearing by surface flows, one quickly runs out of photons; hence the need for large aperture. Nevertheless, the 4m, all-reflecting design of ATST will be a welcome innovation for thermal-IR solar astronomers, who currently are struggling to achieve even $1''$ resolution with the largest existing solar telescopes.

Hand in hand with the construction of ATST, the AURA committee recommended strong (financial) support by NSF for consolidation of NSO into a single headquarters site, and for the full-up operations of ATST together with the existing synoptic program.

(2) Full-Sun Measurements

The new mantra is to treat the Sun as a physical system that contains a hierarchical series of fundamental scale lengths that should be accessed simultaneously to build a coherent picture of the underlying physical processes. ATST will be the premier solar facility for peering into the crucial fine scale details of surface structure, but its observations on the upper end will be limited only to the sizes of active regions. Synoptic instruments like SOLIS and GONG on the ground, and SDO in space, will play key roles in building the broader spatial context for interpreting ATST's measurements, but there is additional complementarity that could be contributed by novel new mid-scale initiatives such as the Coronal Solar Magnetism Observatory (COSMO) and the Frequency Agile Solar Radiotelescope (FASR). One aspect of FASR that should be emphasized is that it, unlike the EVLA or ALMA, is exclusively a solar instrument, so it can provide the radio context without the scheduling vagaries of the dark-side facilities.

(3) Synoptic Solar Astronomy

Synoptic measurements are at the core of elucidating many types of astronomical phenomena — eclipsing binaries, pulsars, flare stars, stellar cycles, AGN reverberation mapping, and so forth — and such time-domain studies are no less important to solar physics. In fact, solar synoptic measurements are a key to Space Weather assessment and prediction, as mentioned earlier, and thus have an operational role beyond the mainly curiosity-driven nature of other parts of astronomy, and indeed much of solar physics itself. Synoptic observing is a mainstay of helioseismology, and its offshoots like far-side imaging; as well as long-term solar cycle studies. Just as the full-Sun view provides a key context for the more narrowly focused stare of large aperture solar telescopes, so too does the long-term synoptic gaze provide the crucial context for shorter-term phenomena. Thus, in the era of 'giant' solar telescopes, we must not lose sight of the value of the synoptic backdrop.

(4) Instrumentation and Technology

To keep ground-based solar physics at the cutting edge, and to complement space-borne assets, it is not simply enough to build a large aperture solar telescope like ATST, but

it is equally important to provide that facility with first-class instrumentation to exploit the large collecting area, and with other technologies to mitigate the main drawback of a ground facility: atmospheric seeing. Furthermore, beyond the complement of first-light experiments, there must be provision for advanced instrument development; to build on the initial experiences, exploit the latest technologies (especially in cameras), and perhaps shift emphasis based on new discoveries. At the outset, ATST will be equipped with Adaptive Optics to deliver crisp imaging, although only over relatively small fields. The development of Multi-Conjugate Adaptive Optics promises to provide stable corrections over much larger fields, but it will be challenging to implement. If MCAO can be demonstrated, it will pave the way for wide-field simultaneous spectral imaging, using for example massive fiber bundle arrays or advanced reflecting image slicers. These types of spatial multiplexing techniques will move solar astronomy beyond the age of the scanned slit, into a brave new world.

(5) Synergy of Coordinated Ground and Space Measurements

Ever since the first solar satellites were launched into orbit, there has been a strong desire to coordinate their observations with those from the ground, again because the context matters. Some types of coordination happen naturally, as with synoptic facilities that simply carry out their canned programs and provide their data products in well characterized, non-proprietary ways. Other types of coordination require considerable effort: to align the space and ground instruments to the same solar target at the same moment, and coordinate the joint analysis of the resulting perhaps not so compatible data sets. Given that the synergy between ground and space observations is firmly established, and will be even more essential in the future, the AURA committee urged special consideration for the planners of new ground facilities to explicitly accommodate joint scheduling with space observatories.

(6) Transfer of Facility Support to Geosciences

A final recommendation considered the then hypothetical, but now imminent, transfer of the National Solar Observatory and its facilities from NSF's Math & Physical Sciences directorate (AST division) to the Geosciences directorate (AGS division), urging that full consideration be given to ensuring that key intellectual and material links of solar to broader astrophysics are not broken. With all of the nation's solar assets supported by AGS, there will be little direct influence on the night-time facilities wholly residing in AST. Under this arrangement, it will be necessary to exercise more diligence to identify and exploit the possible synergies that might be achieved using such facilities during the daytime (as mentioned earlier for EVLA and ALMA). For example, without solar advocacy in AST, there would be the danger that potential solar uses of astronomical facilities might be precluded by inadvertent design decisions, or cost tradeoffs. This is a two-way street, in fact, because one also loses advocacy from the night-time community concerning possible uses of, say, the unique capabilities of ATST for stellar or exosolar planet research ("night-time solar physics" according to one pundit). The concentration of solar into Geosciences also raises flags with the practitioners of the solar-stellar connection, who already are virtually homeless in terms of advocacy in NSF, let alone NASA. More on that, next.

The Solar Stellar Connection

We also take this opportunity to mention the solar-stellar connection, a cross-disciplinary field where the same synergies between ground- and space-based observations exist. The solar-stellar connection is especially valuable for exploring aspects of our star that are difficult to characterize with contemporary measurements, such as the operation of magnetic cycles over evolutionary timescales; or the coronal activity of the Sun at a very young age, of relevance to the survival of primitive planetary atmospheres; or the far distant future of the Sun, when it bloats into a red giant, frying — if not also swallowing — the inner planets of the solar system in the process. In fact, even the near term evolution of our star's luminosity will guarantee the nonhabitability of our home world in a mere billion years or so, regardless of humanity's efforts to tame our contemporary bout of global warming.

From humble beginnings in the early 1970's, with exploratory high-resolution measurements of the Ca II 395nm lines in nearby bright stars, and tentative access to the key ultraviolet wavelengths with *Copernicus,* the study of stars as suns in their own right has blossomed into a mature field, fueled in large part by remarkable instrumental advances. These include *HST, FUSE, Chandra, XMM–Newton,* and now *Kepler,* to mention a few recent examples in space. Similar progress has occurred on the ground in high-resolution spectroscopy, photometry, and spectropolarimetry at optical telescopes; radio imaging at the VLA; and soon the bright promise of mm/sub-mm astronomy with ALMA, again to mention just a few examples. Solar physics has witnessed equally impressive strides with recent missions such as *SoHO, TRACE, Hinode,* and now *SDO,* in space; together with steady advances in ground-based solar imaging and spectropolarimetry, especially with the development of modern adaptive optics systems.

With all these excellent discovery tools at hand, one would imagine that the solar-stellar connection community would be making progress by leaps and bounds. This has been true to some extent, but a major hindrance has been lack of a funding home for this multidisciplinary field. Unlikely purely solar astronomy (seeking grants through AGS programs, or NASA Heliophysics), or its purely stellar counterpart (proposing to the AST AAG program, for example, or NASA ADP), solar-stellar researchers unfortunately seem to be regarded somewhat as orphans by both sides. The major sources of support, in fact, have come almost accidentally through NASA Guest Investigator programs (e.g., *IUE, HST, FUSE, Chandra,* and so forth) rather than via a well-defined, long-term funding opportunity (like NASA's Living with a Star targeted research). Those of us in the solar-stellar community of course desire a higher profile for our science, and additional opportunities to support our work (as well as to hire postdocs and train grad students). The situation of solar-stellar is not unlike that of other multidisciplinary efforts whose advocacy falls between the cracks of traditional fields. It would be self-serving, if not also shameless, of us to propose a specific targeted effort to support solar-stellar research within NASA and/or NSF, but we do wish to raise the awareness of the heliosciences community that an important opportunity to exploit solar-stellar insight for both disciplines could be languishing for lack of direct support.

The use of a fast Z-pinch facility as a platform for laboratory solar and heliophysics

B. Bach, A. Covington, T. Darling, R. Presura
Department of Physics, University Of Nevada, Reno, Nevada 89557-0220

Introduction

There are a number of Z-pinch facilities in operation around the world. These facilities create and investigate the physics of high-energy-density (HED) plasmas. The plasmas produced in these laboratories are in an energy density regime, which is relevant to a number of astrophysical environments such as supernovae, astrophysical jets, black holes and gamma-ray bursts [1]. While HED plasmas are in a high-energy-density regime, they should have implications for solar phenomena.

This concept paper will explore the use of existing Z-pinch facilities to produce laboratory plasmas, relevant to the environment of the sun and or sun-earth interactions.

Summary of the Science

The Z-pinch device:

A fast Z-pinch device is a machine designed to confine and compress (pinch) a plasma column, creating a hotter, denser plasma. The pinch force or compression is a consequence of a self-generated Lorentz or **J** x **B** force that acts on current-carrying plasmas. In a Z-pinch, an electrical current is driven into a plasma (typically along the "z-axis"). The self-generated **J** x **B** force causes the current-carrying plasma to experience a magnetically driven pinch or implosion. The magnetic energy associated with the electrical current is transformed into kinetic energy of the imploding plasma. As the imploding plasma accelerates to the z-axis, the kinetic energy of the plasma is transformed into heat when the plasma compresses and stagnates along the z-axis [2]. Stagnation occurs when the outward pressure of the compressed plasma is equal to the inward pressure of the pinch force. Lighting and solar flares are both examples of self-constricting plasmas that occur in nature.

A fast Z-pinch device is used to implode a variety of "targets" or "loads". A target or load could be something as simple as a single wire or puff of gas. Loads could also be complex configurations of hundreds of wires, metal or plastic foils, foams, hollow or solid objects; even fibers of frozen hydrogen have been used [1]. In application, the target and Z-pinch device are configured to achieve a desired effect. For example, an implosion load might be designed to perform as a bright x-ray or UV light source [3] or to launch plasma jets for use in a shock experiment. Table I. presents the parameters of a typical Z-pinch device and target [4].

Table I. Typical parameters of a Z-pinch device	
Target size(cm)	1-2
Target mass (mg/cm)	1-2
Maximum current(MA)	1-10
Maximum voltage (MV)	1
Implosion time (ns)	100

Solar plasmas and pulsed-power plasmas:

The pinch force causes any current-carrying plasma filament to contract on itself. As many astrophysical objects have magnetic fields, the pinch force is usually present. Our "local" astrophysical object, the sun, regularly hosts dramatic displays of the pinch force, transforming magnetic energy into the kinetic energy and heating of the magnetized plasma. In the sun, magnetized plasmas show up as specific solar features: solar flares, sunspots, and faculae, each with its own spectral features. Solar magnetic phenomena are not well understood, in part because such events are transient and the data sparse. Studies of solar magnetic phenomena have been almost exclusively observational in nature: using ground based or space based telescopes and instruments. The magnetohydrodynamic (MHD) equations of motion, which describe how plasma is accelerated by magnetic and hydrodynamic forces, are independent of scale [5]. In theory the pinch force or other magnetic phenomena, taking place in the solar atmosphere should be reproducible in the laboratory. A Z-pinch facility could create magnetized, solar-like plasmas in a laboratory setting, where plasma conditions can be varied and explored. The diagnostics associated with a typical Z-pinch facility include time-resolved X-ray and UV spectrometers, streak cameras, bolometers as well as D-dot and B-dot sensors for monitoring voltage and magnetic field. Such diagnostics would measure and record the physical state of the plasma.

A pulsed power facility such as a Z-pinch machine can achieved a wide range of plasma conditions, dependent on the configuration of the "target" or "load". Table II presents the parameters characteristic of pulsed-power plasmas [6].

Table II. Typical Pulsed-Power Plasma Conditions	
Electron density(cm^{-3})	10^{13}-10^{23}
Temperature(eV)	10^{0}-10^{3}
Maximum magnetic field(T)	10^{3}
Maximum electric field (GV/m)	1
Density (g/cm^{3})	10^{-6}-10^{1}

As Table II demonstrates, pulsed-power plasmas consist of highly ionized atoms in the presence of strong external fields; this suggests that it may be possible to produce solar relevant or solar-like plasmas in the laboratory environment. To flesh-out this suggestion it is necessary to compare the characteristics of pulsed-power plasmas to those of solar plasmas. Tables III-VI, tabulate typical conditions found in various solar regions [7, 8]. The reader should be aware that all solar regions are very in homogenous, so the stated values only represent rough order of magnitude estimations.

Table III. Solar Convection Zone	
Electron density(cm^{-3})	?
Temperature(eV)	10
Maximum magnetic field(T)	10
Density (g/cm^3)	10^{-1}

Table IV. Solar Photosphere	
Electron density(cm^{-3})	10^{13}
Temperature(eV)	10^{-1}
Maximum magnetic field(T)	10^{-1}
Density (g/cm^3)	10^{-7}

Table V. Solar Chromosphere	
Electron density(cm^{-3})	10^{9}-10^{11}
Temperature(eV)	10^{1}
Maximum magnetic field(T)	10^{-1}
Density (g/cm^3)	10^{-14}-10^{-7}

Table VI. Solar Corona	
Electron density(cm^{-3})	10^{8}-10^{9}
Temperature(eV)	10^{2}
Maximum magnetic field(T)	10^{-1}
Density (g/cm^3)	10^{-17}-10^{-14}

While the characteristics of pulsed-power plasmas and solar plasmas do not correspond in a one-to-one fashion, the comparison does suggest how to design or modify a Z-pinch experiment to observe solar relevant plasmas. Typical Z-pinch loads are designed such that the implosion time (time to the "z-axis") matches the duration of the applied current; this is done to achieve maximum compression on the "z-axis". One option would be to use time-resolved spectroscopy and gated detectors to observe the Z-pinch plasma early in its evolution. Observe the plasma sometime after the load or target has vaporized but before the plasma stagnates on the "z-axis", where it is too dense and hot to be solar-like. Another option would be to design a load or target with the proper acceleration characteristics. Design a "heavy" or "large" load whose implosion time is greater than the duration of the electrical current driven into the plasma. The load would be vaporized but the **J** x **B** force would terminate before the plasma is maximally compressed on the "z-axis".

Relevance to Solar and Space Physics (An Overview)

Z-pinch experiments could make fundamental contributions to understanding the basic physics of the solar process responsible for transforming magnetic energy in to kinetic energy and heat. A laboratory programs to produce solar-like or solar-relevant plasmas in a pulsed-power device would be a small, but coherent part of an integrated strategy for solar research. In certain circumstances, such experiments could run concurrent or piggyback to HED plasma experiments. This environment would put the solar community in close contact with the laboratory astrophysics and HED communities, who have developed laboratory techniques and diagnostics for these experiments. This type of interaction would encourage potential cross-fertilization between the disciplines and across the scientific community. An individual HED plasma campaign can cost approximately $\$10^5$ per year [9]. This number serves as a good estimate as to the cost of a solar-like campaign, as the laboratory equipment and support would be similar to an HED plasma experiment.

Relevance to Solar and Space Physics (In Detail)

Laboratory plasma physics, spectroscopy:

Observational techniques deliver a wealth of data from objects such as sunspots, solar flares and coronal streamers. Much of this information is in the form of photons. The accurate interpretation of observed data is strongly dependent on laboratory benchmarks. Laboratory pulsed-power devices that generate solar-relevant plasmas would advance the study of atomic properties and processes that occur in the presence of strong external fields. Such laboratory data would greatly assist and broaden the interpretation of space-based observations [10].

Space-based instrumentation routinely image at wavelengths ranging from the Infrared to x-ray. The data returned has provided the scientific community with a wealth of spectra. Some data remains to be analyzed because the spectral features are unidentified or unassigned [11]. It is suspected that unassigned features may originate from highly ionized atoms or from plasma backlight by hotter plasma. Unassigned spectral may also originate from high B-field regions in the active sun, the feature being Zeeman or Paschen-Back shifted away from a known low-field location [12].

Visible, UV and x-ray images of the sun reveal dramatic interactions between plasma and the solar magnetic field. Recent extreme-ultraviolet (EUV) images of the earth's magnetosphere, reveal an equally complex interaction between the earth's magnetic field and the solar wind. Variations in the sun's irradiance are relatively large in the UV portion of the spectrum, larger than the relative variation in the total luminosity. This variation in the UV irradiance is believed to have a significant effect on the earth's atmosphere and climate [13].

A laboratory program in pulsed-power plasma spectroscopy (visible, UV and x-ray) could contribute to the fundamental understanding of observed solar and heliospheric phenomena. Such data would provide benchmarks for theoretical and computational efforts or future high-resolution observations. The program would further, our understanding of highly ionized atoms in strong external fields, this is of primary interest to solar, plasma and astrophysical scientists. EUV spectroscopy of solar-like plasmas could help us understand the link between solar activity and variations in the sun's UV irradiance.

Laboratory plasma experiments, empirical:

Z-pinch loads are not just bright sources of x-ray or UV radiation; they can also be designed to launch plasma jets into strong external fields [14]. This type of Z-pinch target would allow laboratory scientists to study the way such plasmas interact with magnetic field lines. This would provide empirical insight into the formation of solar jets and streamers, as well as the interaction between the solar wind and the earth's magnetosphere.

The sun's magnetic field is responsible for the complex plasma dynamics we observe in the sun. Stream interfaces, flux tubes and current sheets are not well understood. Empirical studies of how pulsed-power plasmas behave in the strong external field of the Z-pinch, would contribute to our fundamental understanding of the basic physics involved in the phenomena.

The development of plasma jet experiments could be used as benchmarks for computational and theoretical modeling. Z-pinch experiments could address magnetic reconnection, turbulence, shock and plasma instabilities. These experiments would inform magnetohydrodynamic models and shed light on the process in which magnetic energy is converted into kinetic energy and heat.

Review

The technologies and facilities to pursue the described programs in laboratory solar physics are currently in place. The cost to fund individual research groups and design appropriate targets would be minimal. The laboratory aspects of the programs would be attractive to undergraduate and graduate students who might otherwise choose to pursue a degree in a different discipline. The laboratory results would be of interest to an audience outside the solar physics community.

[1] National Research Council, *Frontiers in Higher Energy Density Physics*, (National Academies Press, Washington, D.C., 2003).
[2] R. P. Drake, *High-Energy-Density Physics*, (Springer-Verlag Berlin Heidelberg 2006), pg380.
[3] Song Inho, Monhanty R.Smruti, Wantanabe Masato, Kawamura Toru, Okino Akitoshi, Horioka Kazuhiko and Hotta Eiki, J. Plasma Fusion Res. Vol 81, No. 9, 647-648, (2005).
[4] D.D. Ryutov, M.S. Derson, M.S. Matzen, Reviews of Modern Physics Vol. 72, No.1, January 2000.
[5] P.M. Bellan, D. Kumar, E.V. Stenson, S. K. P. Tripathi, G.S. Yun, A. L. Moser, Laboratory simulations of astrophysical jets and solar coronal loops, http://ve4xm.caltech.edu/webpub/2010-Bellan-Como-submit-Jan-25-2010.pdf
[6] G.A. Rochau, J.E. Bailey, Y. Maron, Applied spectroscopy in pulsed power plasma, Physics of Plasmas 17, 055501 (2010).
[7] Dermott J. Mullan, *Physics of the Sun,* (CRC Press, Taylor and Francis Group, 2010).
[8] V.H. Hansteen, *Heliophysics*, C. Schrijver, G.L.Siscoe editors, (Cambridge University Press, 2009).
[9] Private communication, A. Covington, R. Mancini.
[10] National Research Council, *Plasma Science: From Fundamental Research to Technological Applications*, National Academy Press, Washington, D.C., 1995, pp. 25, 119.
[11] K. Wilhelm et al., Astronomy and Astrophysics 439, 701-711 (2005).
[12] *Spectral line shapes modeling in laboratory and astrophysical plasmas*, R. Stamm et al., New Astronomy Reviews **53**, 272-276 (2009).
[13] Le Mouël, J.-L., E. Blanter, M. Shnirman, and V. Courtillot (2010), Solar forcing of the semi-annual variation of length-of-day, *Geophys. Res. Lett.*, 37, L15307, doi:10.1029/2010GL043185.
[14] R. Presura, C. Plechaty, S. Stein, D. Martinez, S. Neff, Y. Sentoku, A. Esaulov, V. V. Ivanov, Y. Stepanenko, *Plasma Streaming Across Magnetic Fields*, 2010 NASA Laboratory Astrophysics Workshop, Oct -25th-28th, 2010

The Next Step in Heliospheric Modeling – Increasing the Interplanetary Observing Network

A white paper submitted to the NRC Decadal Strategy for Solar and Space Physics

Christopher Balch (christopher.balch@noaa.gov) – Space Weather Prediction Center

Tremendous advances in physical modeling of the interplanetary medium have occurred over the past decade. A major benefit to society is currently nearing realization as the WSA/ENLIL solar wind model (e.g. Odstrcil & Pizzo, 2009 and references therein) is in active transition to operations at the NOAA Space Weather Prediction Center and is scheduled to become operational in Fiscal Year 2011. However, as was recently pointed out in a talk presented by Gombosi (2010), model uncertainties continue to be problematic due to the sparsity of interplanetary observations with which to initialize or adjust the model calculations. It seems likely that no matter how much more one may invest in modeling the heliosphere, there is a fundamental limit of capability because of this lack of observations.

This situation contrasts dramatically with the modeling of the thermosphere and the ionosphere where techniques have been developed to assimilate ~100 observations of ionospheric quantities (Hajj et al, 2004, Codrescu et al, 2004, Scherliess et al, 2006). These results show that methods for assimilating data to improve the accuracy physical models are well developed.

Therefore, it seems clear that the next substantial step forward for improving physics-based models of the solar wind and interplanetary CME's is to deploy a large number of relatively inexpensive, in-situ solar wind monitors, a suggestion proposed by Gombosi (2010). If such a fleet were deployed in space between the Sun and the Earth, and the observations collected in near-real time, assimilative models of the interplanetary medium could be developed which would be significantly more accurate than the current capability. In particular, a major step forward could be realized through the in-situ measurement of the interplanetary magnetic field; such measurements could be assimilated into the models and provide a new capability for predicting the internal magnetic field of the ICME driver. Such predictions are currently beyond the capability of the existing models. In addition, corrections for the morphology, speed, and mass of interplanetary CMEs could be made which would improve the prediction of the spatial evolution of the CME. This in turn will lead to a better estimate of CME arrival time and duration of CME passage at Earth, a major space weather objective.

In addition to improving model forecasting capabilities for solar wind and ICME's, it is expected that this capability will clearly illuminate deficiencies in current understanding which has been used in developing state-of-the art physics-based models. As is typical with all of science, new observations typically force investigators to re-examine and revise basic concepts and often lead to new understanding of the subject.

An additional benefit of the concept is that the technology to build and deploy the required sensors is well known and should be relatively inexpensive (i.e. in comparison to developing an entirely new sensor). Gombosi (2010) estimates a single 'CubeSat' complemented with a magnetometer, a plasma sensor (for density, velocity, and temperature), and a simple energetic particle sensor, could be deployed for about $1 million. Gombosi suggests that 1000 such sensors, for example, would be cost-comparable to four expensive science missions. In order to determine the optimal number of sensors, however, the concept would require studies evaluating the trade-offs in accuracy versus number of sensors. In addition, the architecture for real-time data telemetry, ingest and assimilation would need to be developed as part of the observing network concept of operations.

In summary, the deployment of an enhanced interplanetary observing network represents an opportunity to use known technologies and modeling techniques to make an order-of-magnitude (or more) improvement in the accuracy of physics-based space weather models for the solar wind and interplanetary CME's, and should thereby result in a significant societal benefit by improving the predictability of adverse space weather at Earth.

References

Codrescu, M.V., Fuller-Rowell, T.J., Minter, C.F, 2004. An ensemble-type Kalman filter for neutral thermospheric composition during geomagnetic storms, Space Weather, Volume 2, Issue 11 (DOI 10.1029/2004SW000088)

Gombosi, Tamas, 2010. Is SWMF Ready for R20?, presented at Space Weather Workshop, April 27-30, 2010, Boulder, Colorado.

Hajj, G.A., Wilson, B.D., Wang, C., Pi, X., Rosen, I.G. 2004. Data assimilation of ground GPS total electron content into a physics-based ionospheric model by use of the Kalman filter, Radio Science, 39, 1 (DOI 10.1029/2002RS002859).

Odstrcil, D., Pizzo, V.J., 2009. Numerical Heliospheric Simulations as Assisting Tool for Interpretation of Observations by STEREO Heliospheric Imagers, Sol. Phys., 259, 1-2, pp 297-309.

Scherliess, L., Schunk, R.W., Sojka, J.J., Thompson, D.C., Zhu, L. 2006. Utah State University Global Assimilation of Ionospheric Measurements Gauss-Markov Kalman filter model of the ionosphere: Model description and validation, J. Geophys Res., 111, A11 (DOI 10.1029/2006JA011712)

High Spectral Resolution, High Cadence, Imaging X-ray Microcalorimeters for Solar Physics

S.R. Bandler[a], J.S. Adams[a], C.N. Bailey[a], J.A. Bookbinder[b], N. Brickhouse[b], S.D. Christe[a], S. Deiker[c], B.R. Dennis[a], E.E. DeLuca[b], J.A. Chervenak[a], W.B. Doriese[d], M.E. Eckart[a], F.M. Finkbeiner[a], G.C. Hilton[d],K.D. Irwin[d], R.L. Kelley[a], D.P. Kelly[a], C.A. Kilbourne[a], J.M. Laming[e], D. Martinez-Galarce[c], F.S. Porter[a], J.E. Sadleir[a], S.J. Smith[a], R.K. Smith[b], R. Stern[c], J. Ullom[d]

[a]NASA/Goddard Space Flight Center, [b]Smithsonian Astrophysical Observatory,
[c]Lockheed Martin Solar and Astrophysics Laboratory,
[d]National Institute of Science and Technology, [e]Naval Research Laboratory

ABSTRACT

A microcalorimeter-based instrument will enable detailed studies of fundamental plasma physics of the solar corona on the finest spectral, spatial, and temporal scales between 0.2 and 10 keV. It will enable us to simultaneously acquire high spectral and spatial resolution data on short timescales; this in contrast to previous instruments that have several minute cadence, sample only a narrow part of the spectrum, and require rastering to build up images. The microcalorimeter capabilities allow us to follow the evolution of the radiative emission during the impulsive phase of flares and study with unprecedented spectral resolution how magnetic energy is released in magnetic reconnection events such as solar flares and coronal mass ejections. The characteristics required of x-ray microcalorimeters for solar physics are significantly different from conventional ones developed for astrophysics because they need to accommodate much higher count-rates (300-1000 cps) and because they require much smaller pixel sizes. We advocate the support of continued research of this type of technology development, and the serious consideration of this technology for future solar missions in the next decade, from sounding rockets, to Explorers, to medium and possibly large scale missions. It is also a part of a trilogy of white papers advocating a new mission concept called the Solar Spectroscopic Explorer (SSE) described by Bookbinder et al.,[1] which will be able to answer the science questions described by Laming et al.[2] In this paper we describe the motivation for this type of technology, the types of microcalorimeter pixels developed to date, future pixel designs, strawman instrument designs, and the resources necessary for operating such an instrument in space.

1. INTRODUCTION

The source of coronal heating of the solar outer atmosphere has puzzled solar physicists for nearly half a century. High quality imaging data have revealed that the solar atmosphere is significantly more dynamic and turbulent than had been previously appreciated. The plasma evolution and the non-equilibrium processes occur on timescales comparable to the integration time for Bragg and slit crystal spectrometers to record a spectrum or to the time needed to raster-scan an image. What would be preferable is an instrument that is simultaneously capable of imaging spectroscopy with high throughput and high spatial and spectral resolution. An x-ray microcalorimeter instrument has the potential to meet such requirements, and it will become ready to be used on a solar mission in the forthcoming decade.

X-ray microcalorimeters have been in development for astrophysics for over 20 years,[3,4] and NASA has invested heavily to bring their technical readiness to levels suitable for satellite missions. For some time it has been suggested that this technology could be suitable for making ground-breaking measurements of the x-ray emission from the solar corona.[5–7] However, questions remained whether large enough arrays could be fabricated, whether individual pixels could accommodate high enough count-rates, and whether the individual pixels could be made small enough. Recently, a NASA ROSES program (Solar and Heliospheric Physics) has

Further author information:
E-mail: Simon.R.Bandler@nasa.gov, Telephone: 1 301 286 1363

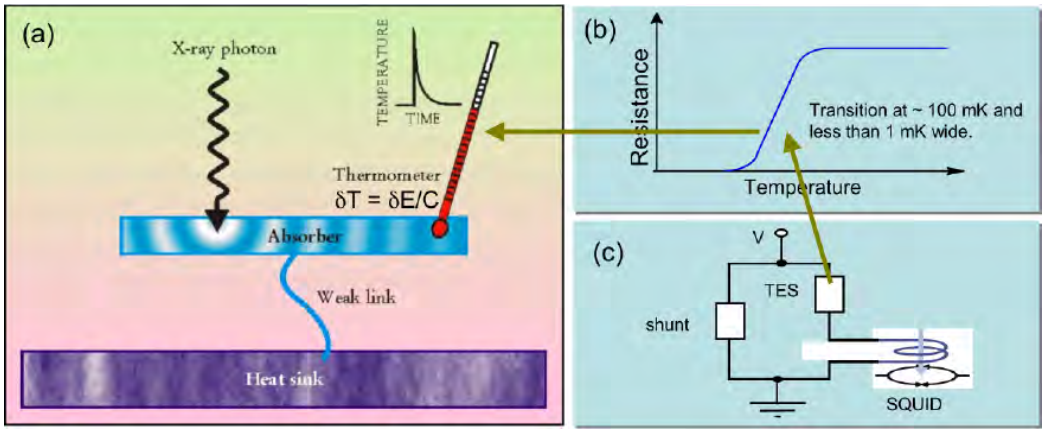


Figure 1. (a) Schematic representation of how a microcalorimeter works. (b) Graph depicting the resistance of a superconductor as it goes through its superconducting transition temperature. (c) Microcalorimeter read-out circuit.

supported the development of microcalorimeter arrays for solar physics with pixels that are small and fast,[8] leveraging a similar technological approach as that used for the International X-ray Observatory (IXO).[9,10] The progress has been extremely promising[11] and in Section 2, we describe designs and results from this technology development. We discuss the current and prospective capabilities of this technology in terms of the energy resolution, focal plane size and layout, and the x-ray count-rates that can be accommodated. In Section 3, we outline the technical readiness level of the various components of a satellite microcalorimeter instrument. In Section 4, we briefly discuss the possibilities for developing a microcalorimeter sounding rocket and explorer payloads. A separate white paper[1] discusses a strawman mission concept and various options one could consider for optimizing missions that utilize a microcalorimeter, and how the different instrument complement options map to the different scientific capabilities. The white paper by Laming et al.[2] discusses many of the different types of measurements that this technology is ideally suited to carry out. Some other white papers being submitted present compelling science objectives and key measurement capabiltiies that could be ideally addressed by this technology. These include Lin et al.,[12] Golub et al.,[13] Doschek (in his new instrument concept paper)[14] and Doschek et al.[15]

2. MICROCALORIMETER CAPABILITIES

2.1 Microcalorimeter design

The operating principle of microcalorimeter x-ray detection is conceptually simple (Figure 1(a)). An x-ray absorber is connected to a low-temperature bath through a weak thermal link. When a photon is incident on the absorber, the photon's energy causes the absorber to warm; it subsequently cools as heat flows out of the absorber and into the heat sink. The height of the temperature pulse is $\Delta T = E/C$ where E is the photon energy and C is the heat capacity of the device. The type of thermometer that is the most promising for building large microcalorimeter arrays is called a transition-edge sensor (TES).[4] In this device, a metal is biased at a temperature in-between its superconducting and normal-metal state (Figure 1(b)), where a small change in temperature causes a relatively large change in resistance. This superconductor is voltage biased so that changes in resistance cause changes in the current flowing through the TES, which is sensed using an extremely sensitive ammeter, a SQUID (superconducting quantum interference device), as shown in Figure 1(c).

Our TES microcalorimeter design is shown in Figure 2. Current flows from left to right across the TES, entering and departing the TES via the (purple) superconducting leads as shown in Figure 2(b). As the leads enter and exit the sensor region, they are widened to the width of the TES. The large yellow square represents the area of the X-ray absorber, which is cantilevered above the substrate and the TES, and only contacts the TES and substrate in the dark "T" region. The design also incorporates interdigitated stripes of thicker Au perpendicular to the current flow to optimize the design of the superconducting transition region. For solar applications, the TES is directly fabricated on the solid substrate; there is no membrane isolation (that exists



Figure 2. (a) A three-dimentional view of a microcalorimeter array, with the front left two pixels shown without an absorber attached. (b) An overhead view. The (purple) TES leads enter and exit the TES horizontally. The TES is a square that is outlined by the edge of the leads and by the green horizontal lines that join the edges of the leads. The large yellow square shows the region of the gold overhanging absorbers which are fabricated on top of the TES sensor but are transparent in this figure for the purpose of clarity.

in most designs for astrophysical detectors), simplifying the design. The relatively high thermal conductance of the pixel is determined by the Kapitza (boundary) thermal resistance between the TES film and the underlying substrate.

We have successfully fabricated these devices with absorbers of two thicknesses, 9.1 μm and 4.5 μm and a range of sizes between 34 and 75 μm. The TESs underneath ranged from 12 to 50 μm on a side. Examples of arrays of these pixels are shown in Figure 3.

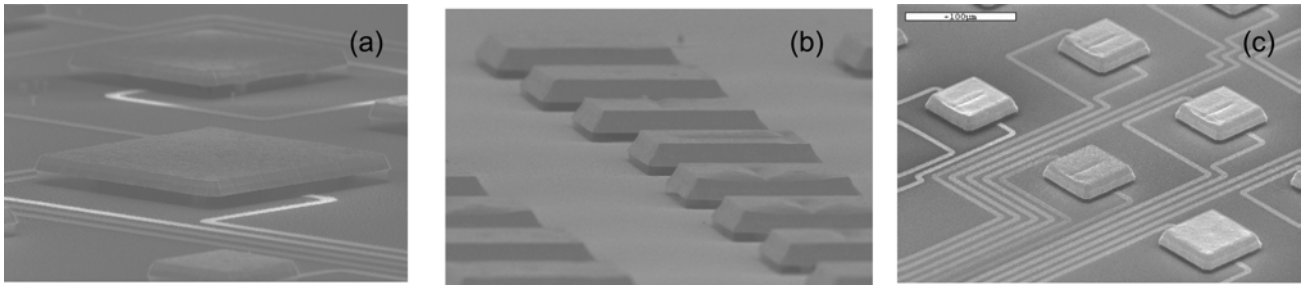


Figure 3. (a) Scanning electron microscope (SEM) image of a single pixel showing a 4.5 μm gold absorber cantilevered above the substrate surface. (b) and (c) SEM images of an array of uniform sized pixels with 9.1 μm gold absorbers fabricated on top of TES sensors.

2.2 Energy resolution

Since the energy resolution of microcalorimeters depends on their operating temperature, they are cooled to cryogenic temperatures, usually $\sim$ 50 mK. A bias current provides Joule heating that keeps them at their transition temperatures (T_c), typically 130 mK for our solar microcalorimeters. In Figure 4 we show K_α spectra from one of our pixels when irradiated by three different x-ray sources to demonstrate the achievable energy resolution at these energies. The microcalorimeter energy resolution ΔE_{FWHM} is 1.26 eV at 1.5 keV, 1.58 eV at 6 keV, and 1.94 eV at 8 keV. The responsivity is slightly non-linear, but this is reproducible and can be calibrated.[11] This small non-linear dependence of signal height on energy is the main reason why the energy resolution of the microcalorimeter is energy dependent. While the complete gain curve up to higher energies has not yet been calibrated, the energy resolution is expected to degrade gracefully as a function of energy and is expected to remain below 10 eV up to 20 keV. The energy resolution is so good that there is hardly any difference between the observed spectra and intrinsic line shape for the Kα spectra from Mn and Cu. The stability of the pulse height for a 6 keV x-ray as a function of time is excellent, with no measurable drift in the gain apparent during the collection of the spectra, better than 1 eV over 10 hours.[11]

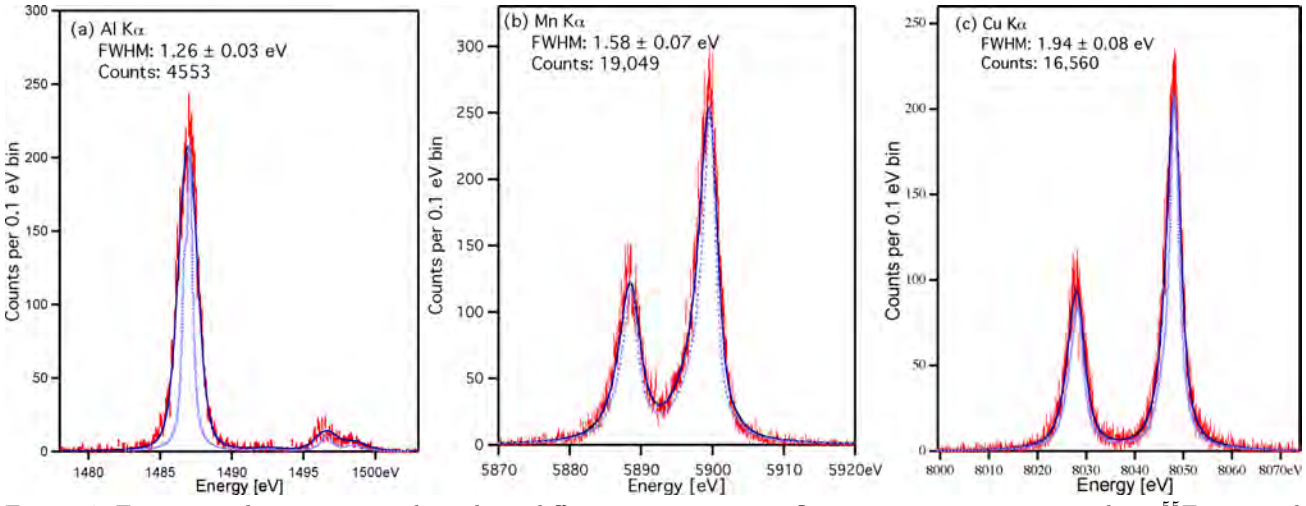


Figure 4. Energy resolution measured at three different energies using fluorescent x-ray sources and an ^{55}Fe internal conversion source. The pixel dimensions for these measurements are 57 μm x 57 μm x 4.5 μm, with an underlying 35 μm TES. In each spectrum, the light blue dotted line shows the intrinsic line shape of the Kα transition. The red lines represent the measured number of counts in each 0.1 eV bin and the error bar for each bin. The darker blue line is the intrinsic line shape convolved with the best fit instrumental broadening (assumed to be gaussian).

2.3 Count-Rate

The performance of the microcalorimeter as a function of count-rate is extremely important for solar physics where the photon fluxes needed for useful measurements are relatively high. Due to the practical limitations of the strength of x-ray fluxes in our experimental setup, the energy resolution as a function of count-rate has not yet been tested. However, we can predict the performance from the measured properties of the microcalorimeters through modeling. We typically use an optimal (matched, digital) filter to calculate the energy of each x-ray pulse and the longer the digitized record length (in time) of a particular x-ray pulse, the more accurately the photon energy can be determined.[18] Pile-up can also cause resolution degradation if the dead-time is too short. Figure 5 shows a plot of the expected energy resolution and throughput as a function of count-rate, for three choices of the total dead-time for measuring each x-ray photon. It is assumed that the detector has 1.6 eV energy resolution at low count rates, and conservatively that all the incoming photons are at 6 keV (in practice, the spectrum of x-rays coming from the Sun will be much softer and the degradation will be lower). As an example, for a dead time of 0.8 ms, an energy resolution of less than 2.1 eV is readily achievable for input count-rates up to 300 cps, and the throughput is 80% at this count rate, and 50% for an input count-rate of 1000 cps. Degradation from 2.1 eV to 2.5 eV would only occur at 1000 cps if all of x-rays were at 6 keV, the degree of degradation depending on the distribution of x-ray energies. For count-rates greater than 1000 cps the energy resolution will degrade, but will still provide very useful data; further studies are necessary to determine the relation between the energy resolution and count-rate in this regime. The other key contribution to the achievable count-rate capability is the magnitude of the cross-talk between pixels. We have developed a design, utilizing a buried copper layer that acts as a thermal ground, in which the thermal cross-talk between pixels has been reduced to the very low level of less than 0.01% of the neighboring pulse-height.[19] Operationally, the use

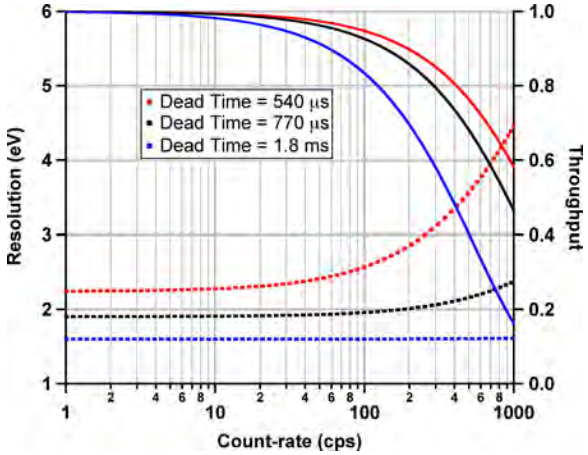


Figure 5. An example of the expected throughput and energy resolution as a function of count-rate for different total dead times surrounding a pulse. The dead times consist of a time since the previous pulse. The dashed lines are for energy resolution, using the left hand axis, and the calculated resolution degradation with count-rate assumes all the x-rays are at 6 keV. The continuous lines are for the throughput using the right hand axis.

of filters with different transmissivities on a rotatable filter wheel in front of the microcalorimeter can be used to address fluctuations in source strengths.

2.4 Position-sensitive microcalorimeters - "Hydras"

The coverage of microcalorimeter arrays can be enhanced through the use of position-sensitive microcalorimeters that increase the effective number of pixel elements in an array. There are a number of different ways in which these can be designed[7,20,22] for solar physics; they can be used to increase the field-of-view of a focal plane for a fixed number of sensors and read-out channels, or can be used to increase the spatial resolution through sub-pixel imaging. The approach that we believe has the greatest potential in terms of optimizing the energy resolution is to use a "Hydra" geometry, in which a single TES is coupled to more than one discrete absorber, as depicted schematically for four absorbers in Fig. 6(a).[21] Each absorber element has a different thermal conductance to the sensor which results in position information being encoded in the pulse shape. This type of device has been successfully fabricated and tested in larger sizes for astrophysics, and it performs as predicted.[21,22] The count-rates that can be accommodated by each Hydra are lower by about a factor of n in terms of count-rate per absorber, where n is the number of absorbers attached to the TES. This is the inherent tradeoff in designs that use Hydras. In the next section, we discuss a strawman large-format array that uses Hydras similar to those that have already been developed for astrophysics, that have nine absorbers attached to each TES sensor. Nine-headed Hydras have already been fabricated and shown to work, and a version consistent with the strawman design is being developed.

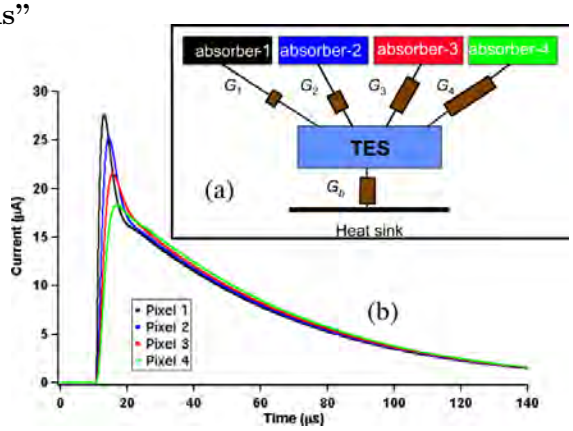


Figure 6. (a) A simple thermal model of the Hydra concept including four absorbers with different thermal couplings to a TES, and the TES is weakly thermally coupled to a heat sink. (b) Average pulse shapes for a Hydra designed for solar physics. After the initial position-dependent equilibration signal, the pulses decay with the same exponential rate.

2.5 Large-Format Arrays

Figure 7 shows a photograph of a state-of-the-art 1 kilo-pixel microcalorimeter array developed for x-ray astrophysics.[23] Larger TES bolometer arrays have been made and are being used for infrared radiation detection, such as the second Submillimeter Common User Bolometer Array (SCUBA-2), which has 12,800 pixels.[24] For SSE, Bookbinder et al.[1] describe a strawman array design that includes a core array and an outer array. The core array consists of 36x36 pixels on a 75 μm pitch behind a 15.4 meter focal length grazing incidence telescope. This corresponds to a 1" pixel pitch and a field-of-view of 0.6'. Each pixel is expected to have an energy resolution of less than 2 eV. Fine stripline wiring is used to wire out all the pixels, with the wiring feeding out underneath the cantilevered absorbers, the gaps between neighboring TESs being large enough to accommodate the largest number of these wires.[11] The 1296 pixels can be read out using 41 multiplexer columns (discussed in the next section). The outer region is a 96x96 pixel array of pixels on a 150 μm pitch (2") that provides a 3.2' field-of-view. In this region, the array incorporates a Hydras design with 9 absorbers attached to each TES. Thus, there is a 32x32 array of TESs in this outer array region (minus those not needed in the inner array region), that can be read out using a further 31 multiplexer columns. The energy resolution of the pixels in the outer array will be better than 5 eV. The outer array is fabricated onto a separate substrate that has a hole exposing the inner array (similar to IXO). While not a requirement, the goal is for a field-of-view of greater than 5'. This array design is motivated by the need for a very high angular resolution core for making the primary measurements, with a larger field-of-view from the outer array to provide sufficient context for these

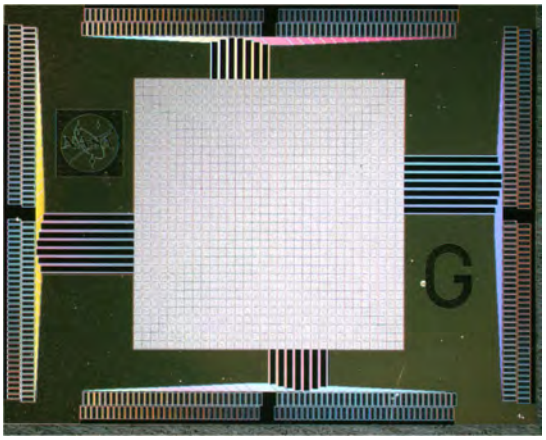


Figure 7. Photograph of an IXO prototype 32x32 array of TES microcalorimeters.

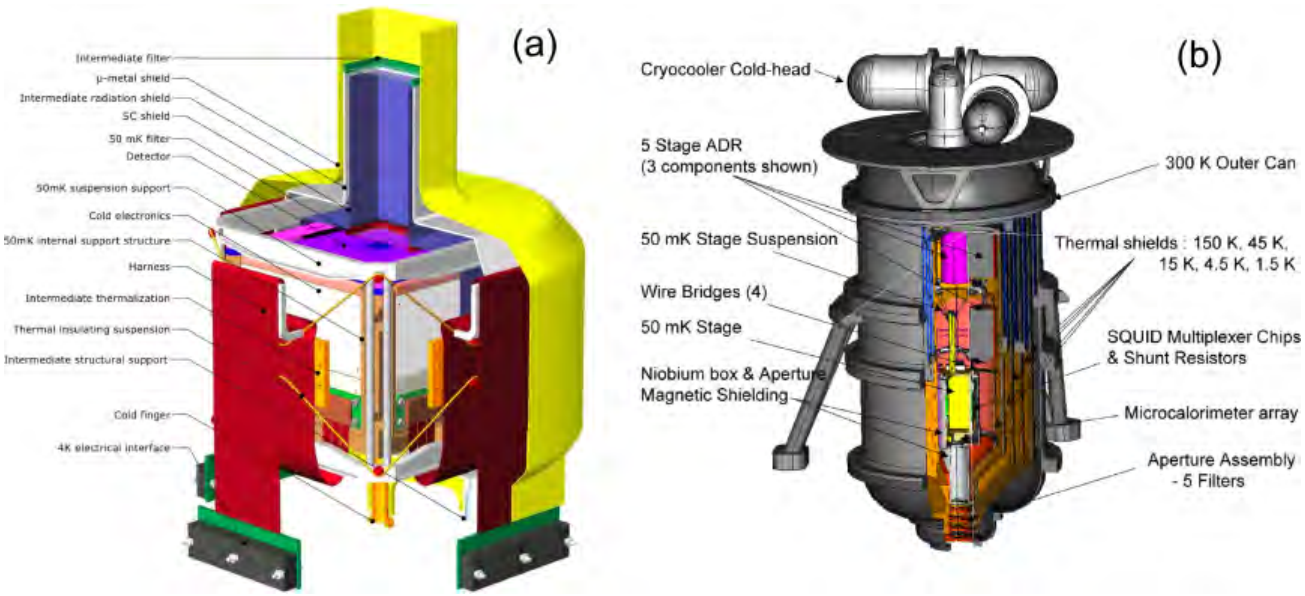


Figure 8. (a) CAD drawing of the IXO focal-plane assembly that houses the microcalorimeter. (b) The design of the IXO cryostat.

measurements. The outer array is also key for providing a sufficient field-of-view for studying the energetics of coronal mass ejection, as described by Bookbinder et al.[1] The total number of TESs is similar to IXO, and so the mass, power and cost can be estimated for this SSE design. The total mass of the microcalorimeter instrument is estimated to be 300 kg, the power required is 650 W, and the cost is approximately $100M.

2.6 Read-out

Our TESs are read out using SQUIDs, and a typical read-out scheme for a single TES detector employs a single SQUID in close proximity to the TES. If a full amplifier channel were provided for each detector in the array, the heat load from both the large number of SQUIDs and the large number of wires running to the TES cold bath would be impractical. To address these concerns, the National Institute of Standards and Technology has developed a SQUID multiplexing architecture to read out large arrays of TESs.[25–28] This multiplexer has undergone a great deal of successful development work for IXO. Solar physics observations place very challenging requirements on the multiplexer, since the rise and fall times of the x-ray pulses are so fast. With the appropriate filter in the TES bias circuit, the rise and fall times will both be approximately 30 μs. The type of multiplexer that allows us to read 32 TESs with each multiplexer for such fast pulses is called code-division multiplexing (CDM).[29] CDM has already been demonstrated to operate with our highest resolution solar pixels in a 1x4 experiment at GSFC which showed that very similar energy resolution can be achieved using CDM multiplexing as when reading out single pixels. With the increase of bandwidth expected from the development of this technology, and using a properly optimized filter to damp the pulse rise-times correctly, it should be possible to multiplex 32 rows in this way without any energy resolution degradation.

2.7 Cryogenics

The cryogenics to provide the 50 mK temperature necessary to operate a microcalorimeter is very similar to that developed for Suzaku, being developed for Astro-H, and to the one that is being designed for IXO (see Figure 8(b)). The fact that the instrument needs to be pointed towards the sun matters little to the cryogenic design. Once a temperature for the outer shield of the cryostat has been established, the cryogenics inside is essentially the same. A thermal design, suitable sunshades and pre-filters are needed to achieve a temperature close to 300 K for the outside of the dewar. The pre-filters will be nearly identical to those used on existing solar instruments such as TRACE, AIA and XRT. Further filters are needed inside the cryostat. Just like for IXO, the cooling system includes a combination of a last stage adiabatic demagnetization refrigerator, Joule Thomson coolers and 2-stage Stirling Coolers. A typical operational life for the coolers is 5 years. Figure 8(a) shows a CAD drawing of the design of the IXO focal plane assembly where the microcalorimeters are situated. Since the number of TESs being read out by the strawman solar mission is almost identical to that proposed for IXO,

the design for a solar mission focal plane assembly would be similar, but a little easier since an anti-coincidence detector will not be necessary.

3. HERITAGE AND TECHNICAL READINESS

At NASA/GSFC we have been developing microcalorimeter arrays for x-ray astrophysics for over 25 years and have successfully flown them on both suborbital and orbital observatories.[30,31] They have been routinely used at Electron Beam Ion Traps (EBITs) for over a decade to provide previously unobtainable measurements of the atomic physics of highly ionized, non-equilibrium plasmas.[32] In addition, TES microcalorimeters are part of a broad class of sub-Kelvin detectors with a long heritage in laboratories around the world. Such low-temperature detectors have been successfully deployed, using a variety of cooling techniques, in Antarctica,[33] balloon flights,[34] airborne observatories,[35] and other astrophysical satellites.[36] The premature loss of the XRS instrument on the Suzaku observatory occurred after the basic operation of the instrument was successfully verified. The cause of the failure was related to missing a requirement on the accommodation of the instrument within a closed spacecraft, and not inherently related to the principles of operation of the detector system or low temperature technology, all of which performed as designed and tested prior to launch. While there is no question that flying a microcalorimeter may be more complex than other more commonplace technologies, the failure of XRS should not be viewed as higher risk in the future. By moving entirely to cryogen-free designs, the XRS failure mode is eliminated. In all reviews of the multi-billion dollar IXO mission in the US and Europe, it has not been suggested that there is enhanced risk to having a core instrument that requires 50 mK cryogenics.

A microcalorimeter instrument consists of 4 main components - the detector, read-out, focal-plane assembly (FPA), and the cryogenics. Each of these components can be (and has been) broken down further, for studies of technical readiness for IXO. Here we briefly summarize these technical readiness levels in the context of the requirements for the strawman solar mission. The number of TES read-out channels is the same as IXO, so we can draw from the wealth of knowledge gained from studies of this mission.

Detector All the components necessary to validate the technology in a laboratory environment have been established and integrated, and it has been shown that they work together, and for this reason a technology readiness level (TRL) of 4 is assigned. Although Hydras of the necessary dimensions have not yet been validated, they are similar to many other Hydra designs, and the requirements do not appear challenging in comparison with what we expect is achievable. There exists an ongoing technology development program, and it is projected that this technology will be at TRL-5 in 2013 when close-packed 40x40 arrays of single pixels will be tested, and 80x80 arrays of Hydras will be demonstrated.[8]

Read-out Although the IXO read-out has been assigned a technology readiness level of 4-5, the requirements for a solar mission dictate that code division multiplexing will be necessary, and the technology readiness level of this read-out is 3-4. However, because of its similarity to time division multiplexing, we expect that the technology readiness level will advance quickly to TRL-5 by 2013 under existing funded programs.

FPA While the Astro-H FPA is currently at TRL-8, just as for IXO, a solar instrument will require a different design. Thus the FPA is assigned a TRL of 3-4. What is required for a higher TRL-level is some investment in support technology development, but there do not appear to be any major technical hurdles.

Cryogenics The technology readiness level for the cryogenics is already at an advanced stage due to the development of the Astro-H cryogenics, which, in its current form, will be the baseline for a solar instrument. The cryocooler system and the adiabatic demagnetization refrigerator are currently at TRL-5 and will be at TRL-6 in 2012.

4. FUTURE MISSION CONCEPTS WITH MICROCALORIMETERS

Microcalorimeter technology is making rapid progress to the point where it should be seriously considered for future solar missions. Under current technology development efforts, the technical readiness level of this technology will be at or beyond TRL-5 by 2013, and will be ready for a mission that begins development in the forthcoming decade. A scientific case for the application of this technology has broad community support,[2] and mission designs of various sizes that utilize its application are being proposed.[1,12,13] An astrophysics sounding rocket mission that will fly an array of SQUID-multiplexed TESs is scheduled for launch in 2011[37] and will demonstrate the potential of this technology as a high resolution spectrometer. Small Explorer scale missions such as SSE,[1] medium sized missions such as SSE+,[1] and large-scale missions[12,13] could potentially use a microcalorimeter to make a variety of measurements that will reveal unique plasma diagnostics measurements that could be the key to finally understanding the source of coronal heating.

REFERENCES

[1] "The Solar Spectroscopy Explorer Mission", Bookinder et al., A White Paper submitted to the heliophysics decadal survey "A Decadal Strategy for Solar and Space Physics" - Solar and Heliospheric Physics panel.

[2] "X-ray Diagnostics of Solar Plasmas from a Broadband Imaging High-Resolution Spectrometer" J.M. Laming et al., A White Paper submitted to the heliophysics decadal survey "A Decadal Strategy for Solar and Space Physics".

[3] "Thermal Detectors as X-ray Spectrometers", S.H. Moseley, J.C. Mather, D. McCammon , Journal of Applied Physics, 56, 5 1257-1262 (1984).

[4] "Transition-Edge Sensors", K.D. Irwin and G.C. Hilton in "Topics in Applied Physics", Vol **99**, pp. 63-149 (2005).

[5] "The Reconnection And Microscale (RAM) Solar Terrestrial Probe", J. Bookbinder, E. DeLuca, P. Cheimets, L. Golub, D. Hassler, C. Korendyke, P. Glenn, E. Silver. Proc. SPIE, vol. **4853** (2003), pp. 436-452.

[6] "The Advanced Technology Solar Spectroscopic Imager - a novel experiment employing a transition-edge sensor to probe the soft X-ray solar corona Paul Boerner, Dennis S. Martinez-Galarce, Kolo Wamba, Blas Cabrera, Steve Deiker, Kent Irwin, Troy W. Barbee(II), Phil C. Baker. Nuclear Instruments and Methods in Physics Research **A** 520 (2004) 372375.

[7] "Recent developments in transition-edge strip detectors for solar x-rays" A.J. Rausch, S.W. Deiker, G. Hilton, K.D. Irwin, D.S. Martinez-Galarce, L. Shing, R.A. Stern, J.N. Ullom, L.R. Vale Proceedings of the SPIE, Volume 7011, pp. 70113T-70113T-9 (2008).

[8] "High Spectral Resolution, High Cadence, Imaging X-ray Microcalorimeters for Solar Physics - Phase 2" A proposal accepted by the Solar and Heliospheric Physics Supporting Research and Technology program (SHP09), Dec. 2010-Sep. 2013.

[9] "The X-Ray Microcalorimeter Spectrometer for the International X-Ray Observatory" R.L. Kelley, S.R. Bandler, W.B. Doriese, Y. Ezoe, R. Fujimoto, G. Luciano, R. Den Hartog, J.-W. Den Helder, H. Hoevers, K.D. Irwin, Y. Ishisaki, C.A. Kilbourne, P. De Korte, J. Van Der Kuur, K. Mitsuda, T. Ohashi, L. Piro, F.S. Porter, K. Sato, K. Shinozaki, P. Shirron, S.J. Smith, Y. Takei, P. Whitehouse, N.Y. Yamasaki, LTD-13, AIP conf. proc. 1185 (2009) 757-762.

[10] "Multiplexed read-out of uniform arrays of TES x-ray microcalorimters suitable for Constellation-X" C.A. Kilbourne, W.B. Doriese, S.R. Bandler, R.P. Brekosky, A.-D. Brown, J.A. Chervenak, M.E. Eckart, F.M. Finkbeiner, G.C. Hilton, K.D. Irwin, N. Iyomoto, R.L. Kelley, F.S. Porter, C.D. Reintsema, S.J. Smith, J.N. Ullom, Proc. SPIE 7011 (2008) 701104.

[11] "High Spectral Resolution, High Cadence, Imaging X-ray Microcalorimeters for Solar Physics" S.R. Bandler, C.N Bailey, J.A. Bookbinder, E.E. DeLuca, J.A. Chervenak, M.E. Eckart, F.M. Finkbeiner, D.P Kelly, R.L. Kelley, C.A. Kilbourne, F.S. Porter, J.E. Sadleir, S.J. Smith, and R.K. Smith Proc. SPIE 7732 (2010) 773238.

[12] "Solar Eruptive Events (SEE) 2020 Mission Concept" R. P. Lin, S. Bandler, S. Christe, J. Davila, B. Dennis, G. Holman, R. Milligan A. Y. Shih, S. Kahler, M. Wiedenbeck, G. Doschek, G. H. Share, A. Vourlidas, J. Raymond, A. Caspi, M. McConnell, G. Emslie, A White Paper submitted to the heliophysics decadal survey "A Decadal Strategy for Solar and Space Physics"

[13] "The Reconnection And Microscale (RAM) Solar Terrestrial Probe" L. Golub et al., A White Paper submitted to the heliophysics decadal survey "A Decadal Strategy for Solar and Space Physics"

[14] "A Concept White Paper for a New Solar Flare Instrument Designed to Determine the Plasma Parameters in the Reconnection Region of Solar Flares at Flare Onset" G. Doscheck. A White Paper submitted to the heliophysics decadal survey "A Decadal Strategy for Solar and Space Physics"

[15] "The High Resolution Solar-C International Collaboration: Probing the Coupled Dynamics of the Solar Atmosphere" A White Paper submitted to the heliophysics decadal survey "A Decadal Strategy for Solar and Space Physics"

[16] "Various Optimisations of TES Arrays for X-ray Astrophysics" C.A. Kilbourne, S.R. Bandler, A.-D. Brown, J. Chervenak, E. Figueroa-Feliciano, F. Finkbeiner, N. Iyomoto, R.L. Kelley, F.S. Porter, S.J. Smith, Journal of Low Temperature Physics 151 1-2 (2008) 223.

[17] "High-density arrays of x-ray microcalorimeters for Constellation-X " C.A. Kilbourne, S.R. Bandler, A. Brown, J.A. Chervenak, E. Figueroa-Feliciano, F.M. Finkbeiner, N. Iyomoto, R.L. Kelley, F.S. Porter, T. Saab, J. Sadleir, J. White. Proc. of the SPIE 6266, 572 (2006).

[18] "Optimal filtering, record length, and count rate in transition-edge-sensor microcalorimeters", W.B. Doriese, J.S. Adams, G.C. Hilton, K.D. Irwin, C.A. Kilbourne, F.J. Schima, J.N. Ullom, AIP Conference proceedings **1185**, pp. 450-453 (2009).

[19] "Development of Embedded Heatsinking Layers for Compact Arrays of X-ray TES Microcalorimeters", F.M. Finkbeiner, C.N. Bailey, S.R. Bandler, R.P. Brekosky, A.-D. Brown, J.A. Chervenak, M.E. Eckart, R.L. Kelley, D.P. Kelly, C.A. Kilbourne, F.S. Porter, J.E. Sadleir, S.J. Smith Presented at the 2010 Applied Superconductivity Conference and submitted to the special issue of the IEEE Transaction on Applied Superconductivity.

[20] "Position-sensitive low-temperature detectors", E. Figueroa-Feliciano, Nucl. Inst. and Meth. A, 520 (2004) 496-501.

[21] "Development of position-sensitive transition-edge sensor X-ray detectors" S.J. Smith, S.R. Bandler, R.P. Brekosky, A.-D. Brown, J.A. Chervenak, M.E. Eckart, F.M. Finkbeiner, N. Iyomoto, R.L. Kelley, C.A. Kilbourne, F.S. Porter, J.E. Sadleir, E. Figueroa-Feliciano, IEEE Transactions on Applied Superconductivity 19(3) (2009) 451-455.

[22] "Development of arrays of position-sensitive microcalorimters for Constellation-X" S.J. Smith, S.R. Bandler, R.P. Brekosky, A.-D. Brown, J.A. Chervenak, M.E. Eckart, E. Figueroa-Feliciano, F.M. Finkbeiner, R.L. Kelley, C.A. Kilbourne, F.S. Porter, J.E. Sadleir, Proc. SPIE 7011 (2008) 701126.

[23] J. Bookbinder et al., astro-ph/1003.2847, (2010).

[24] "Characterization of a prototype SCUBA-2 1280 pixel submillimetre superconducting bolometer array", Adam L. Woodcraft et al., Proc. SPIE 6275, 62751F (2006).

[25] "Time-division superconducting quantum interference device multiplexer for transition-edge sensors", P.A.J. de Korte et al., Rev. Sci. Instr., 74, 3807 (2003).

[26] "In-focal-plane SQUID multiplexer", K.D. Irwin et al., Nucl. Instr. and Meth. A, 520, 544 (2004).

[27] "Prototype system for superconducting quantum interference device multiplexing of large-format transition-edge sensor arrays", C.D. Reintsema et al., Rev. Sci. Instr. 74, 4500 (2003).

[28] "Progress toward kilopixel arrays: 3.8 eV microcalorimeter resolution in 8-channel squid multiplexer", W.B. Doriese et al., NIMPA, Volume 559, Issue 2, Pages 808-810 (2006).

[29] "Code-division multiplexing of superconducting transition-edge sensor arrays" K D Irwin,et al., Supercond. Sci. Technol. 23 034004 (2010).

[30] "A High Spectral Resolution Observation of the Soft X-Ray Diffuse Background With Thermal Detectors", D. McCammon, et al, ApJ, Vol. 576, Iss 1, 188 (2002).

[31] "The next-generation microcalorimeter array of XRS on Astro-E2", C.K. Stahle, C.A. Allen, K. R. Boyce, R.P. Brekosky, G.V. Brown, J. Cottam, E. Figueroa-Feliciano, M. Galeazzi, J.D. Gygax, M.B. Jacobson, R.L. Kelley, D. Liu, D. McCammon, R.A. McClanahan, S.H. Moseley, F.S. Porter, L.E. Rocks, A.E. Szymkowiak, J.E. Vaillancourt, Nucl. Inst. & Meth. A, **520** 466 (2004).

[32] F.S. Porter, et al., "Laboratory astrophysics using a spare XRS microcalorimeter", Proc. SPIE, 4140 (2000) 45.

[33] "SPT-CL J0546-5345: A Massive z ¿1 Galaxy Cluster Selected Via the Sunyaev-Zel'dovich Effect with the South Pole Telescpe", M. Brodwin et al., ApJ 721:90-97, (2010).

[34] "The TopHat Experiment: A Balloon-borne Instrument for Mapping Millimeter and Submillimeter Emission", R.F. Silverberg et al., ApJS 160 59 (2005).

[35] "Cosmic-ray flux detected by an IR bolometer operated on board of a stratospheric aircraft", M. Lanfranchi et al., Infrared Physics & Technology, Volume 40, Issue 5, p. 379-386 (1999).

[36] "Herschel Space Observatory - An ESA facility for far-infrared and submillimetre astronomy", G.L. Pilbratt et al., Astronomy & Astrophysics, 518, L1 (2010).

[37] "Science with Micro-X: The TES microcalorimeter X-Ray Imaging Rocket", E. Figueroa-Feliciano, et al., Proc. SPIE, Space Telescopes and Instrumentation II: Ultraviolet to Gamma Ray; Martin J. Turner, Guenther Hasinger; Eds. Volume 6266, pp. 62660A (2006).

White Paper: Using Laboratory Experiments to Study Solar Corona Physics

P. M. Bellan

MC 128-95 Caltech, Pasadena CA 91125; pbellan@caltech.edu

1 Recommendation

It is recommended that the Solar and Space Physics Decadal Survey include the statement:

"Because of recent demonstrations that certain laboratory plasma experiments can simulate essential three dimensional MHD properties of solar coronal loops, the next decade of heliospheric research should include advanced laboratory plasma experiments designed specifically to address important outstanding issues relevant to the solar corona."

In support of this recommendation this white paper lists important current issues in solar corona physics, reviews recent lab experiments at Caltech that have addressed some of these issues, and concludes with an example of a future experiment designed to address a specific and important solar corona issue.

2 Issues in solar corona physics

1. The coronal heating problem - why does the solar corona blatantly violate the second law of thermodynamics by having a temperature over two orders of magnitude hotter than the photosphere?
2. The coronal loop uniformity problem - why is the cross-section of a coronal loop approximately constant along the length of a loop?
3. The triggering problem - why do coronal loops erupt suddenly? Are eruptions triggered by some slight internal change or by an external event?
4. The power supply problem - what is the energy source for the non-potential magnetic fields in the solar corona? Twisted magnetic fields have more internal energy than non-twisted (i.e., potential) fields satisfying the same boundary conditions. It is conventionally assumed in numerical magnetohydrodynamic (MHD) simulations that the twisting of coronal loops results from differential motion at the photospheric surface but recent measurements (Pevtsov *et al.*, 2003) show that actual photospheric footpoint motion is insufficient to account for the observed twist.
5. The particle acceleration problem - what accelerates particles to high energies in an eruption? Solar coronal physics is traditionally modeled using ideal MHD which has the essential feature that no electric field component exists parallel to the magnetic field. As a consequence of this feature, MHD is fundamentally incapable of predicting the generation of energetic particles. In stark contrast to this, observations show that a considerable fraction of the stored magnetic energy in twist is transformed into energetic particles and x-rays.
6. The scale height and particle source problem - why is the plasma density in coronal loops greater than the surroundings? If the coronal loop magnetic fields were potential or force-free, there would be no force other than gravity. Isothermal loops would therefore be in hydrostatic equilibrium in which case the plasma density in the loop would decay exponentially with altitude as $\exp(-z/H)$ where $H = \kappa T/mg \sim 3 \times 10^7$ m is the gravitational scale height. However, actual loops are uniformly filled with plasma to altitudes exceeding $3H$. The density at the apex of these high loops is thus more than 20 times greater than predicted by hydrostatic equilibrium models. This indicates the presence of strong upward forces that overwhelm gravity; neither force-free nor potential magnetic fields can explain such forces.

3 Lab experiments as a means for studying solar corona physics

A great similarity exists between solar corona magnetic structures and a magnetic fusion confinement plasma concept known as the spheromak (Bellan, 2000). The author is neither the first nor the only person to note this similarity which has been exploited in theoretical models (Kliem and Torok, 2006), numerical simulations (Gibson and Fan, 2008), and in the interpretation of ground and space-based observations (Nandy *et al.*, 2003). However, the author has gone a step further than merely noting a similarity. The author has purposely adapted spheromak technology to create lab experiments designed to be explicitly relevant to solar physics. This adaptation consists of locating the vacuum chamber wall far from the plasma to mimic the solar situation and changing the symmetry of boundary conditions to conform to the solar situation. As with solar plasmas, plasma dynamics in the experiments is subject to the constraint that $\mathbf{E} + \mathbf{U} \times \mathbf{B} = 0$, i.e., magnetic flux is frozen into the frame of the plasma. Thus as with solar plasmas, dynamics is determined by the MHD equation of motion giving the evolution of the plasma velocity as a function of the magnetic field while simultaneously the induction equation gives the magnetic field evolution as a function of the plasma velocity.

The laboratory method has its share of advantages and disadvantages and, like any other new idea, care is required to avoid overselling it. Advantages include the ability to replicate the topology and dynamics of observed solar structures in a highly reproducible manner, the ability to diagnose all physical parameters in principle, the ability to control initial conditions, and the ability to observe how MHD physics connects to non-MHD physics such as generation of energetic particles and x-rays. The use of a real plasma means that the laboratory simulation is inherently self-consistent and correct. The main shortcoming is that the lab plasma is not an exact scale model of a solar phenomenon so while what is observed in the lab can give useful insights into solar behavior, the lab plasma results cannot be simply scaled up. Furthermore, setting up a successful experiment involves significant challenges such as designing and building apparatus that provide unambiguous reproducible results, developing suitable diagnostics, and relating measurements to theoretical models in a useful way.

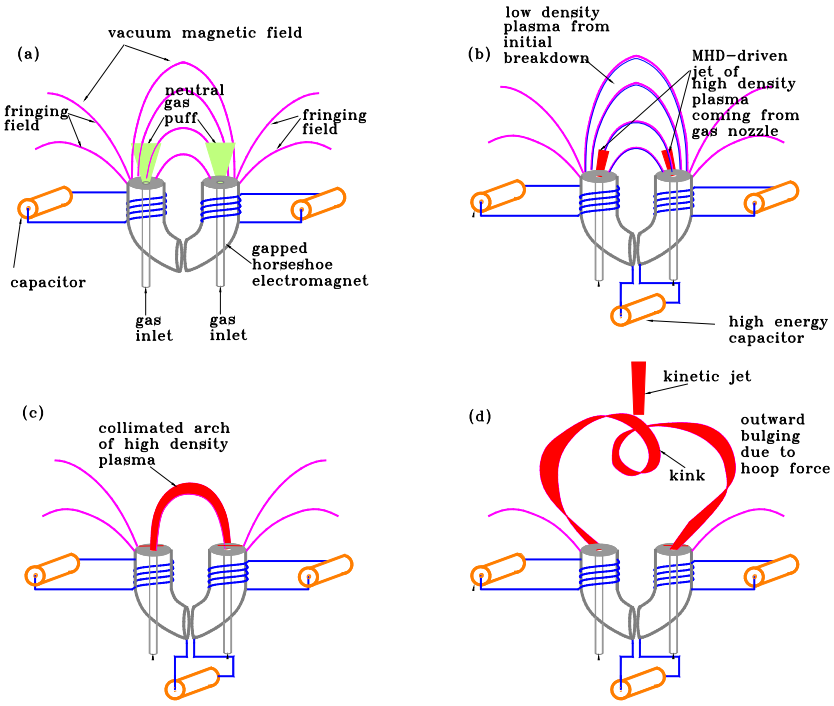


Figure 1: Sequence: (a) potential field established, neutral gas injected, (b) breakdown, upflows, (c) collimation, (d) expansion, kink, possible kinetic jet

The controlled nature of the experiments means that proposed models can be subjected to stringent tests. In particular, because basic parameters such as magnetic field strength and electric current can be adjusted, the credibility of a model can be quickly tested. This experimental approach to solar physics gives an intuitive 'hands-on' feel for the various physical phenomena. Laboratory experiments can also test numerical MHD simulations which at least should be able to simulate the MHD aspects of a lab experiment.

The author initiated the experiments described here in 1995 and has been continuing since. Several other groups have begun similar experiments in the last five years and so laboratory simulation of solar physics can now be considered to be an emerging discipline rather than just the effort of a single group. Other groups who have laboratory simulations of solar phenomena underway include the University of Bochum, Germany (Soltwisch *et al.*, 2010), the LAPD group at UCLA (Tripathi and Gekelman, 2010), the MRX group at Princeton (Williams *et al.*, 2008), and the Naval Research Lab (Dahlburg *et al.*, 2010).

4 Outline of basic method

Figure 1 sketches the typical sequence of Caltech experiments while Fig. 2 shows images of actual configurations. In the first step of the sequence, a quasi-static 'horseshoe' electromagnet (Fig.1a) produces an initial arch-shaped potential magnetic field with field strength of the order of 1-3 kG at the footpoints. Next, sufficient neutral gas to enable breakdown is puffed via high-speed electromagnetic pulsed gas valves into the region between the magnet poles. After the gas puff, a high-energy capacitor bank charged to several kilovolts is connected across the magnet poles [see Fig.1(b)] and breaks down the neutral gas in about 0.1 μs to form a low density plasma as indicated in Fig. 1(b).

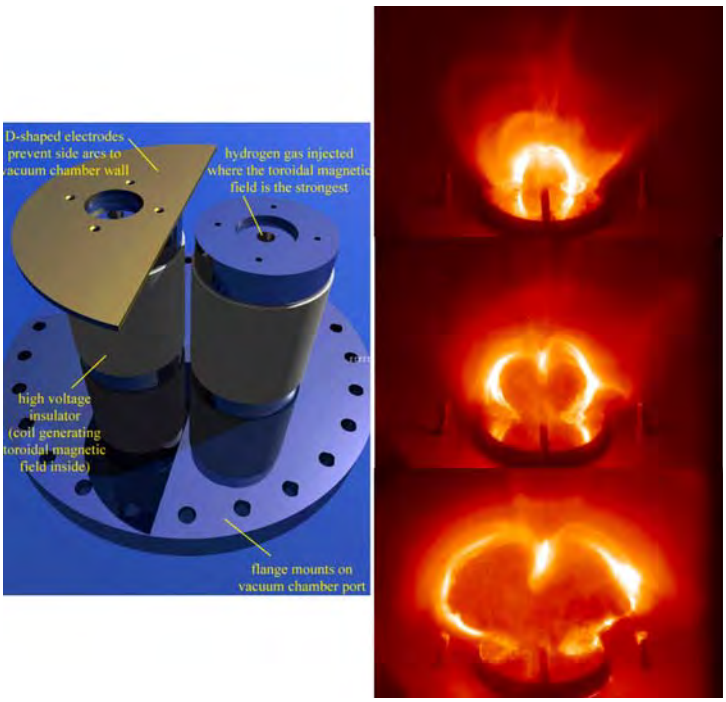


Figure 2: Left: configuration for creating single coronal loop (D-shaped electrode removed on RHS). Right: typical experimental lab simulations of single coronal loop

The initial current flowing through the tenuous plasma produced at breakdown is small and follows the arched vacuum field lines [Fig. 1(b)]. As the high-voltage capacitor discharges, this current increases to a peak value of 40-80 kiloamps in about 5-10 μs. The MHD $\mathbf{J} \times \mathbf{B}$ force associated with the current is observed to drive upflowing Alfvenic velocity plasma jets from both footpoints [see Fig. 1(b)]. These fast upflowing plasma jets quickly collide at the apex of the arched loop and fill the flux tube with plasma. The flux tube collimates as the two jets collide and then writhes to form a kink-like dip in the middle.

All the while, the magnetic hoop force causes an expansion of the major radius of the arch [see Fig. 1(d) and Fig. 2(right)]. The jet upflows and collimation were both unexpected effects that instigated development of a theoretical model (Bellan, 2003). This model motivated further experiments (You *et al.*, 2005) and these experiments provided verification for the flow/collimation model.

The sequence shown in Figs. 1 and 2(right) occurs only if neutral gas is provided at the nozzle; the gas is ionized at the nozzle and so provides a plasma source at the nozzle. Plasma density in the bright loop results from the flux loop being filled with plasma by the jet upflows. The magnetic configuration thus depends on a mass flux boundary condition (i.e., plasma ingestion) as well as on the normal magnetic field and the normal electric current boundary conditions.

5 Significant results to date

A single loop system as described in Fig. 1 was constructed. This experiment showed (i) physics long thought to occur only on the sun could, in fact, be reproduced in the laboratory, (ii) the loop self-collimated almost instantly, (iii) the loop axis writhed in a helix, (iv) the surface projection was sigmoidal, (v) the major radius increased with time [see Fig. 2(right)], and (vi) plasma upflow from the electrodes affected magnetic field behavior.

5.1 Strapping field

A second, independent magnetic field system was arranged to provide a horizontal magnetic field $\mathbf{B}_{ext}$ passing through the simulated coronal loop (see Fig. 3 top). The vertically directed $\mathbf{J}_{loop} \times \mathbf{B}_{ext}$ force could, depending on the polarity of $\mathbf{B}_{ext}$, aid or inhibit the tendency of the loop major radius to expand. Studies (Hansen and Bellan, 2001) of this multiflux system showed that loop eruption could be slowed or halted by application of a suitably strong $\mathbf{B}_{ext}$ with appropriate polarity (see Fig. 3 bottom).

5.2 Spider legs

The early stages of a configuration used for astrophysical jet simulations involved an arcade of eight coronal loops (You *et al.*, 2005) arranged like the legs of a spider (see Fig. 4). These loops have the same helicity and merge to form a coaxial jet which then undergoes a dramatic kink instability (Hsu and Bellan, 2003). As seen in Fig. 4, these spider legs are very collimated in contrast to the dipole potential field existing before their formation.

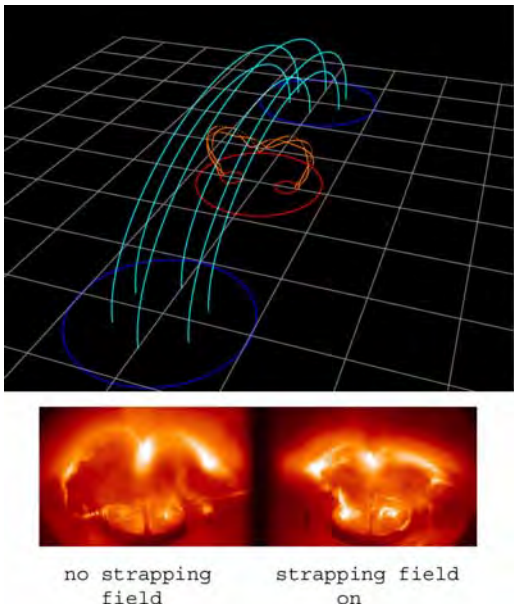


Figure 3: Strapping magnetic field (top, turquoise) inhibits upward expansion of loop (bottom right)

5.3 Tracking flows using two gas species

High speed movies indicate that the morphology of the simulated coronal loop depends on the mass flux of plasma upflowing from the footpoints; these observations motivated the 'gobble' model (Bellan, 2003). Since mass flux is an important behavior which is not generally appreciated and which has caused a certain amount of controversy, significant efforts have been devoted to check whether the prediction of plasma upflow from footpoints is indeed real. Flow from footpoints has been measured using Doppler spectroscopy and strong blue shifts are observed (Tripathi

et al., 2007) indicating suprathermal mass upflows.

A conclusive demonstation of mass upflowing from footpoints is provided as follows: By using two distinct gases the optical line emission of the plasma originating from the two footpoints is arranged to be distinct, i.e., the plasma is color-coded according to its originating footpoint if in fact it originates from a footpoint. By placing atomic line filters in front of the camera the two different colors are resolved and so the trajectories of footpoint-originating plasma is established, e.g., red plasma upflows from one footpoint while green plasma upflows from the other.

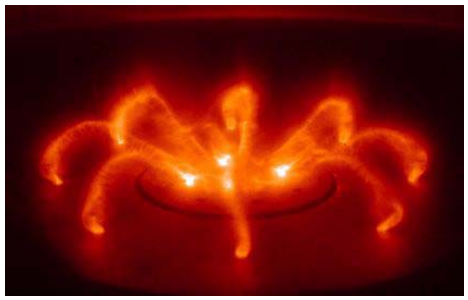


Figure 4: Spider legs: eight arched plasma-filled magnetic flux tubes

Experiments using two gases Stenson and Bellan (Stenson and Bellan, 2008) have clearly shown these upflows from each footpoint, with the flow velocity dependent on the gas mass. Figure 5 shows results where hydrogen (red) is supplied at the upper footpoint and nitrogen (green) at the lower footpoint. The hydrogen travels faster than the nitrogen and the loop is constituted by the identifiable flows from the two footpoints.

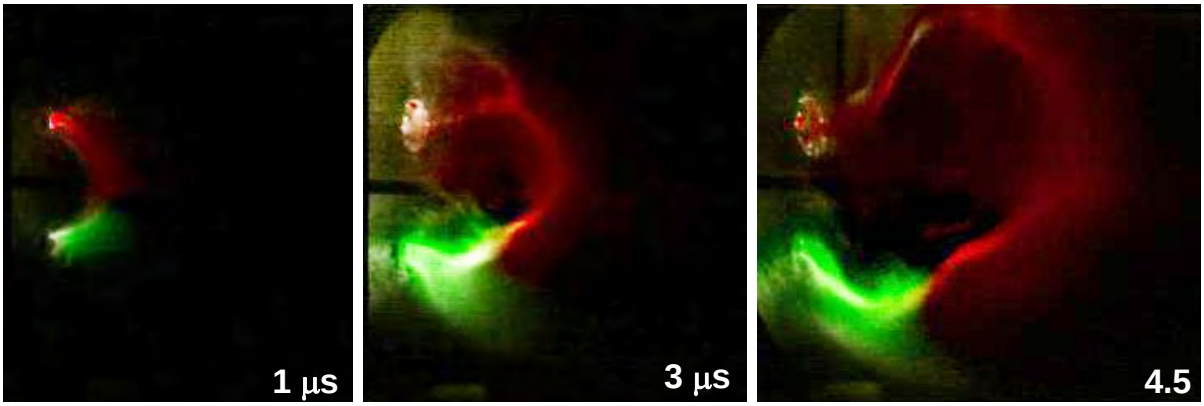


Figure 5: Two-gas experimental results: Composite photo made by superimposing line-filtered images of hydrogen plasma (top, red) and nitrogen plasma (bottom, green). Hydrogen gas is supplied at the top footpoint and nitrogen at the bottom footpoint (Stenson and Bellan, 2008).

6 Impact to date

The experiments described above have made an impact on both observational and theoretical solar physics. As examples, Meshalkina et al. (Meshalkina *et al.*, 2009) have cited our merging experiments in interpreting observed solar eruptions, while Kliem (Kliem and Torok, 2006) and Olmedo (Olmedo and Zhang, 2010) have used results from our experiments in the development of the torus instability model for eruptive instability.

7 The next decade

The experiments described above have demonstrated the feasibility of laboratory experiments and their relevance to solar corona research. Future experiments will exploit the infrastructure/expertise that

has been developed to investigate reconnection in three dimensions, x-ray production, Alfven wave propagation, triggering of instabilities, effect of magnetic nulls, jet flows, and plasma heating mechanisms. Improved diagnostics will measure the 3D evolution of the magnetic, electric, and velocity fields, the particle velocity distribution, and mechanisms by which magnetic energy is converted into particle energy. Figure 6 sketches a 3D magnetic reconnection experiment now under construction. This experiment has been motivated by discussions with solar observer S. Martin who has proposed that long solar filaments (prominences) form by repeated reconnections as shown in Fig. 6.

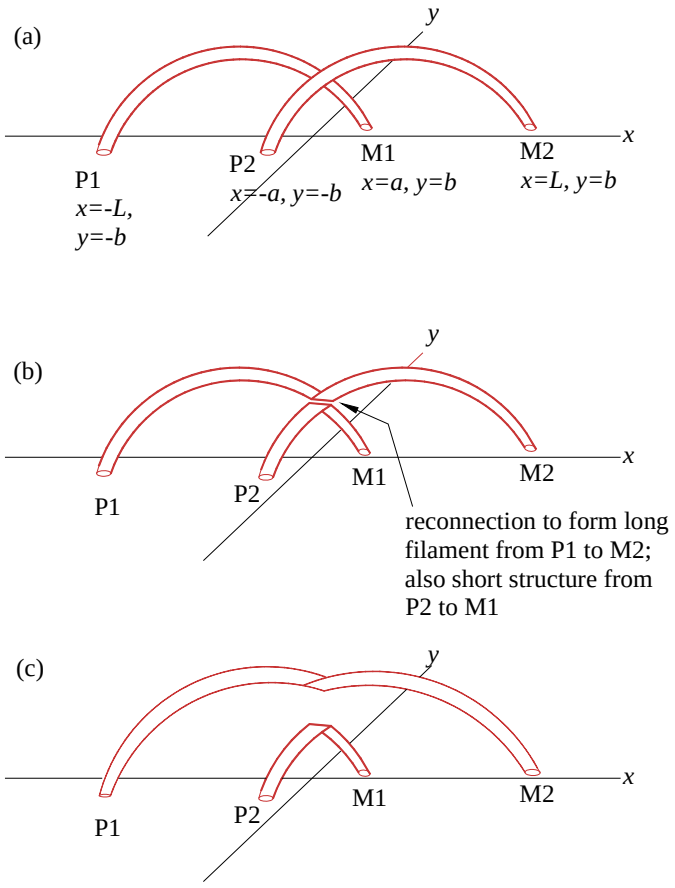


Figure 6: Sketch showing layout and expected sequence of 3D magnetic reconnection experiment now under construction. This sequence is believed to be the mechanism by which the long filamentary structures in prominences starting with much shorter structures.

Acknowledgement: The work described here has been supported by USDOE, NSF, and AFOSR.

References

P. M. Bellan. *Spheromaks: a practical application of magnetohydrodynamic dynamos and plasma self-organization*. Imperial College Press, London, 2000.

P. M. Bellan. Why current-carrying magnetic flux tubes gobble up plasma and become thin as a result. *Physics of Plasmas*, 10(5):1999–2008, 2003. Part 2.

J. Dahlburg, W. Amatucci, and et al. Exploiting laboratory and heliophysics plasma synergies. *Energies*, 3:1014, 2010.

S. E. Gibson and Y. Fan. Partially ejected flux ropes: Implications for interplanetary coronal mass ejections. *Journal of Geophysical Research-Space Physics*, 113(A9), 2008.

J. F. Hansen and P. M. Bellan. Experimental demonstration of how strapping fields can inhibit solar prominence eruptions. *Astrophysical Journal*, 563(2):L183–L186, 2001. Part 2.

S. C. Hsu and P. M. Bellan. Experimental identification of the kink instability as a poloidal flux amplification mechanism for coaxial gun spheromak formation. *Physical Review Letters*, 90(21), 2003.

B. Kliem and T. Torok. Torus instability. *Physical Review Letters*, 96(25), 2006.

N. S. Meshalkina, A. M. Uralov, V. V. Grechnev, A. T. Altyntsev, and L. K. Kashapova. Eruptions of magnetic ropes in two homologous solar events of 2002 June 1 and 2: a key to understanding an enigmatic flare. *Publications of the Astronomical Society of Japan*, 61(4):791–803, 2009.

D. Nandy, M. Hahn, R. C. Canfield, and D. W. Longcope. Detection of a Taylor-like plasma relaxation process in the sun. *Astrophysical Journal*, 597(1):L73–L76, 2003. Part 2.

O. Olmedo and J. Zhang. Partial torus instability. *Astrophysical Journal*, 718:433, 2010.

A. A. Pevtsov, V. M. Maleev, and D. W. Longcope. Helicity evolution in emerging active regions. *Astrophysical Journal*, 593(2):1217–1225, 2003. Part 1.

H. Soltwisch, P. Kempkes, F. Mackel, H. Stein, J. Tenfelde, L. Arnold, J. Dreher, and R. Grauer. Flarelab: early results. *37th EPS Conference on Plasma Physics*, 2010. abstract I3.409.

E. V. Stenson and P. M. Bellan. Dual-species plasmas illustrate MHD flows. *IEEE Transactions on Plasma Science*, 36(4):1206–1207, 2008. Part 1.

S. K. P. Tripathi and W. Gekelman. Laboratory simulation of arched magnetic flux rope eruptions in the solar atmosphere. *Phys. Rev. Letters*, 105, 2010. Art. No. 075005.

S. K. P. Tripathi, P. M. Bellan, and G. S. Yun. Observation of kinetic plasma jets in a coronal-loop simulation experiment. *Physical Review Letters*, 98, 2007. Art. No. 135002.

N. Williams, E. Oz, M. Yamada, H. Ji, S. Dorfman, B. McGeehan, and J. Schroeder. Experimental study of solar flares in MRX. *American Physical Society, 50th Annual Meeting of the Division of Plasma Physics, November 17-21, 2008,*, 2008. abstract JP6.039.

S. You, G. S. Yun, and P. M. Bellan. Dynamic and stagnating plasma flow leading to magnetic-flux-tube collimation. *Physical Review Letters*, 95(4), 2005.

Solar MAgnetized Regions Tomograph (SMART) Mission

By P.N. Bernasconi, and Nour-Eddine Raouafi
Johns Hopkins University / Applied Physics Laboratory

Science Questions addressed by the SMART mission

I. How do magnetic fields evolve with height in the solar atmosphere?

Information on the height-evolution of solar magnetic fields presently relies exclusively on the extrapolation of photospheric or chromospheric magnetic fields, rather than measurements. The photospheric and chromospheric magnetograms are used as the required lower boundary condition. Different assumptions result in various sophistication levels of the extrapolation of the same boundary magnetogram: current-free field (Schmidt 1964), linear force-free (LFF) field (Alissandrakis 1981; Gary 1989), and nonlinear force-free (NLFF) field (Schrijver et al. 2006, 2008; Metcalf et al. 2008 and references therein). NLFF extrapolation provides more accurate field topologies but they are computationally demanding.

Further inconsistencies arise when photospheric magnetic fields are used without preprocessing. These inconsistencies are due to the non-force-free nature of the magnetic field at the photospheric level (Georgoulis and LaBonte 2004). As such, they cannot be used in force-free field extrapolations. This problem is tackled by "preprocessing" the photospheric fields to bring them closer to a force-free equilibrium (Wiegelmann et al. 2006, 2008). The realism of the preprocessing methods is, however, still questionable (Schrijver et al. 2008).

The photosphere being non-force-free, chromospheric magnetic fields are believed to be much closer to a NLFF equilibrium (Metcalf et al. 1995—see, however, Socas-Navarro 2005a for a challenging finding). As a result, there should be a layer between the photosphere and the chromosphere where the transition from forced and force-free fields occurs. We call this layer the *magnetic transition region* (MTR), not to be confused with the radiative/thermal solar transition region. The evolution of solar magnetic fields through and above the MTR is completely unknown. Extrapolations have little to offer in this respect, given their problems and crude spatial sampling. Investigating this layer - where thermodynamic mechanisms of heating and acceleration chromospheric and low-coronal plasma take plane- require systematic multi-height and nearly simultaneous vector field measurements (see e.g. Figure 1). This approach has never been attempted due to the lack of dedicated instruments.

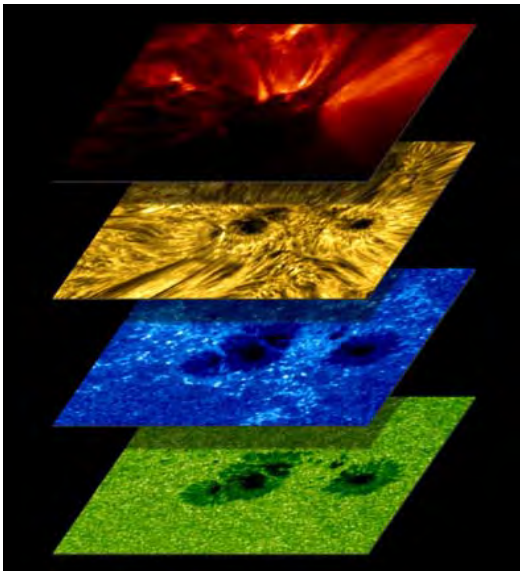


Figure 1: The proposed vector magnetograph will measure the variation of the solar magnetic field vector with altitude from the photosphere to well into the chromosphere, covering also the

Present vector magnetograms are predominantly photospheric, although chromospheric vector magnetography has recently seen some advances (Metcalf et al. 2005; Socas-Navarro 2005b; Lagg 2007). Coronal magnetography is still in its infancy. **The SMART mission proposes to probe the atmosphere between the photosphere through the MTR and into the chromosphere through near-simultaneous, multi-height measurements of the vector magnetic fields.**

II. When, and how much magnetic energy is released in solar eruptions?

Understanding the physical processes leading to transient solar activity (flares/CMEs) is extremely important for space-weather studies and predictions. The lack of significant breakthrough in this field is mainly due to the lack of measurement of the magnetic field at atmospheric layers where most of the free energy that powers solar eruptions resides. Most solar activity takes place well above the photosphere, where most available measurement capability exists. Regnier and Priest (2007) provided evidence for non-even distribution of the magnetic free energy in an active region and showed that it is predominantly accumulated within a few Mm to a few tens of Mm above the lower photospheric boundary (see also Figure 2).

These chromospheric altitudes are exactly the ones at which the proposed vector magnetograph will deliver measurements of the magnetic field vector. The availability of such measurements will allow accurate calculation of the magnetic free energy in a straightforward fashion by means of, say, the magnetic Virial theorem (Molodensky 1974; Aly 1984). These measurements are also very important for constraining extrapolation techniques and consequently the determination of coronal magnetic fields. Measurements before and after an eruption will determine the amount of magnetic energy released during the eruption. This task is so sensitive that it cannot be accomplished using seeing-dependent, ground-based vector magnetograms (see Metcalf et al. 2005 for an attempt).

In addition, high-quality measurements of chromospheric magnetic fields that are not too dependent on photospheric line-tying may establish well-defined, conspicuous eruption-associated magnetic patterns (on the photosphere, generally subtle magnetic field changes have been found after the eruptions; see Sudol and Harvey 2005). Such a finding would (i) significantly advance our understanding of solar magnetic eruptions and (ii) possibly lead to the identification of pre-eruption magnetic patterns that are telltale as to the imminent triggering of solar eruptions. Currently, all these capabilities are strongly desired but can only be envisioned as future accomplishments.

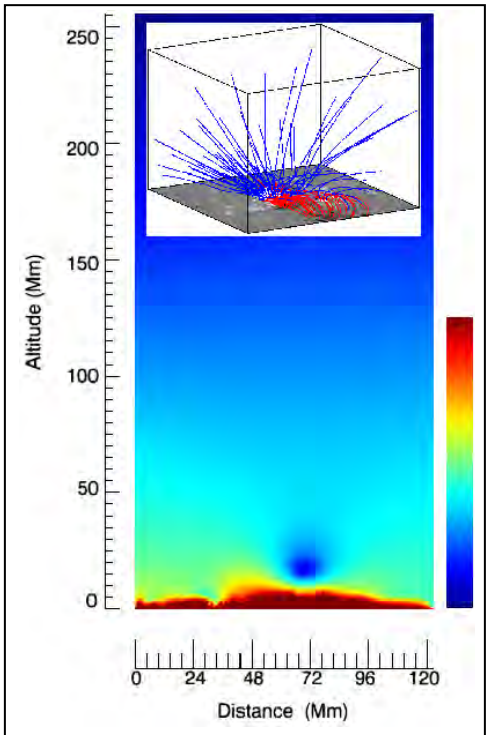


Figure 2: Spatial distribution of the free magnetic energy in a cross-section of the NLFF extrapolated magnetic fields in NOAA AR 10930 (inset). The color code corresponds to the logarithm of the free energy. About 98% of this energy resides

III. What heats the solar chromosphere and corona?

Ever since the discovery of the million-degree solar corona by Grotrian and Edlen (Grotrian 1939; Saha 1942) coronal heating remains a mystery. Even more puzzling is the fact that the

corona cannot be maintained without a chromosphere. Chromospheric heating and mass supply are very challenging problems that need to be addressed in order to advance our understanding of the solar corona.

Recent findings link the supply of chromospheric mass to the "leakage" of p-mode resonant acoustic oscillations through upward flows and spicules (De Pontieu et al. 2004). It has been established later that the damping of acoustic oscillations is decisively insufficient to provide the required power (Fossum and Carlsson 2005; Carlsson et al. 2007). Alfven waves (Tomczyk et al. 2007) and small-scale dissipation of magnetic free energy (the *nanoflare* concept of Parker 1988) may play a role in the heating of the chromosphere. Lites et al. (1998) showed that photospheric magnetic fields are not correlated with chromospheric brightenings, contrary to the findings of Sivaraman and Livingston (1982). We now know that the photospheric fields below the MTR have quite different properties compared to those of the chromospheric fields above the MTR. Therefore, a possible magnetic heating discovery in the chromosphere may await the flight of our proposed magnetograph.

Doppler velocity and magnetic field measurements must be taken with cadence <30 s in order to isolate any wave motions, and this speed would be achievable only with an imaging magnetograph that can switch rapidly from one spectral line to another. Our measurement strategy is to combine images of the vector magnetic field, line intensity, and velocities in the photosphere and MTR over a large field of view (>100 Mm). High resolution (~100 km) and high sensitivity (30 G, transverse) are called for in the Heliophysics Roadmap.

SMART Mission Concept

SMART is a space based solar observatory either in Geosynchronous or Sun synchronous orbit aiming at producing 3D tomographic magnetic field measurements of the solar active regions and the quiet Sun. It will map the solar vector magnetic field at different heights from the photosphere through the chromosphere. The measurement of the polarization Stokes vector (I, Q, U, & V) of several magnetic-sensitive spectral lines will be carried out at 5 to 7 positions across the profiles of every line. Several spectral lines in the red and near infrared wavelength range will be selected for this mission. The filter measurements at a single wavelength yield 2D spatial information both in intensity and polarization at one height in the atmosphere. The 3D map is built by combining the 2D observations from the different spectral lines. The spatial resolution will be from 0.15 arcsec in the photosphere to about 0.25 arcsec in the chromosphere, which is sufficient to resolve single flux-tubes both in the photosphere and in the chromosphere. The cadence will be on the order of 5 to 8 minutes per full tomographic scan, which is also sufficient to capture dynamic processes in the solar atmosphere.

SMART Instrument Description

The SMART instrument is composed of a light-weight 1-m solar telescope feeding an imaging vector magnetograph that uses two Lithium Niobate ($LiNbO_3$) Fabry-Perot etalons for the wavelength selection, a polarization analyzer based on two liquid crystal variable retarders and a linear polarizer, and two 4Kx4K CCD cameras, one for high resolution imaging and vector magnetic field mapping at 0.05 arcsec/pixel, and the second one for full-disk imaging and global vector magnetic field mapping at 0.5 arcsec/pixel resolution. A summary of the instrument characteristics is given in Table I.

Table I: SMART Strawman instrument characteristics

Major System	**Description**					
Gregorian Telescope	1-m diameter light weight					
	Focal plane fields of view	200 x 200 arcsec 0.05 arcsec/pixel				
		33 x 33 arcmin, 0.5 arcsec/pixel (full-disk)				
	Diffraction Limit	λ [nm]	630	850	1000	1100
		Resolution	0.15”	0.21”	0.25”	0.27”
Spectral filter	$LiNbO_3$ dual Fabry-Perot etalon					
	WL range	630 – 1000 nm				
	Free spectral range	2.6 – 7.4 nm				
	FWHM bandpass	0.08 – 0.22 nm				
Polarimeter	2 Liquid crystal variable retarders + 1 linear polarizer					
	Records Stokes (I,Q,U,V) in 4 seconds with 4 consecutive exposures					
Science cameras	2 4Kx4K CCD optimized for Red and NIR					

The two cameras with different fields of view have complementary purposes. The narrow angle camera (200x200 arcsec FOV with 0.05 arcsec pixels) will focus on following the temporal evolution of the vector magnetic fields in single active regions with a spatial resolution sufficient to observe single flux tubes and how they evolve from the photosphere to the chromosphere. The wide angle, full-disk camera will map the vector magnetic fields at a lower resolution (i.e., 0.5 arcsec/pixel) over the entire solar disk to provide context data to the narrow FOV images and for synoptic mapping. Both cameras will simultaneously take measurements at the same wavelength and for the same polarization state. The full disk vector magnetograms will be complementary to the vector magnetograms captured by HMI onboard SDO since they will have nearly the same spatial resolution but SMART will map the vector magnetic fields at different heights in the solar atmosphere.

A selection of the possible spectral lines that the instrument will probe is listed in Table II. For each line images of the full Stokes vector polarization are measured at 5 to 7 different position across the line. It will require about 1 minute to scan one spectral line and about 8 minutes to probe all the lines listed. For faster cadence the number of lines scanned can be reduced to only 3, enough to obtain a 3D representation of the evolution of the magnetic fields from the photosphere to the chromosphere.

Table II: List of candidate spectral lines that SMART may probe

Spectral line [nm]	630.1	630.2	656.3	108.3	849.8	854.2	868.8	769.9
Element	Fe I	Fe I	Hα	He I	Ca II	Ca II	Fe I	K I
Lande *geff*	1.67	2.5	1.05		1.07	1.1	1.66	1.33
Height [km]	330	260	~2000	~2000	~1100	~1000	~500	~450

Mission Implementation

The spacecraft carrying the 1-m telescope and attached vector mangetograph should fly in Sun-synchronous or Geosynchronous orbit allowing continuous viewing the Sun. A Geosynchronous orbit is preferable for reasons of higher data return. The mission baseline should be of at least 5 years to allow observing at least one solar half cycle, with launch timing preferably either at the minimum or at the maximum of solar activity. We have not performed a cost estimation but a good guess would be that such a mission may be in the vicinity of $500M

including launch vehicle and phase E costs. Table III summarizes the key mission profile elements.

Table III: Summary of the SMART mission profile

Orbit	Sun synchronous or Geosynchronous
Strawman Payload	1 m telescope with vis/NIR vector magnetograph
Launch timing	Beginning or maximum of solar cycle
Baseline mission duration	5 years
Observations duty cycle	5 to 7 minutes per full 3D mapping 24 h /day
Spacecraft	Single S/C 3-axis stabilized
Telemetry	TBD
Mass	TBD
Cost	~ $500 M

SMART will be an important addition to the Great Heliophysics Observatory

The SMART mission is meant to be an essential complement to the SDO mission. The HMI instrument onboard SDO is providing moderate spatial resolution, high temporal cadence vector magnetic field information at one single height in the solar atmosphere, namely the photosphere. SMART with its full-disk imager will provide vector magnetic field maps with similar spatial resolution, somewhat lower temporal cadence, but at different heights in the atmosphere providing the necessary information to accurately reconstruct the global vector magnetic fields over the atmospheric layer where the magnetic field transitions from being plasma dominated to dominating the plasma, without solely relying on models. The high spatial resolution imager will provide unprecedented extremely detailed information of the spatial and temporal evolution of single flux-tubes in active regions and quiet Sun not only on a plane to the solar surface but also in height.

Such a mission will be an essential addition to the suite of instruments constituting the Great Heliophysics Observatory.

References

Alissandrakis, C. E 1981, *On the Computation of Constant Alpha Force-Free Magnetic Field*, A&A 100**,** 197

Aly, J. J. 1984, On Some Properties of Force-Free Magnetic Fields in Infinite Regions of Space, ApJ 283, 349

Carlsson, M., Hansteen, V. H., DePontieu, B., McIntosh, S., Tarbell, T. D., Shine, D., Tsuneta, S., Katsukawa, Y., Ichmimoto, K., Suematsu, Y., Shimizu, T., and Nagata, S. 2007, *Can High Frequency Acoustic Waves Heat the Quiet Sun Chromosphere?*, Pub. Astron. Soc. Japan 59, S663

De Pontieu, B., Erdelyi, R., and James, S. P. 2004, *Solar Chromospheric Spicules from the Leakage of Photospheric Oscillations and Flows*, Nature 430, 536

Fossum, A. and Carlsson, M. 2005, *High-frequency acoustic waves are not sufficient to heat the solar chromosphere*, Nature 435, 919

Gary, G. A. 1989, *Linear Force-Free Magnetic Fields for Solar Extrapolation and Interpretation*, ApJ Suppl. Series 69, 323

Georgoulis, M. K., and LaBonte, B. J. 2004, Vertical Lorentz Force and Cross-Field Currents in the Photospheric Magnetic Fields of Solar Active Regions, ApJ 615, 1029

Grotrian, W. 1939, *Zur Frage der Deutung der Linien im Spektrum der Sonnenkorona*, Naturwissenschaften 27, 214

Lagg, A. 2007, *Recent Advances in Measuring Chromospheric Magnetic Fields in the He I 10830 A Line*, Adv. Space Res. 39, 1734

Lites, B. W., Rutten, J. R., and Berger T. E. 1999, *Dynamics of the solar chromosphere. II. Ca II H-2V and K_2V grains versus internetwork fields*, ApJ 517, 1013

Metcalf, T. R., Leka, K. D., and Mickey, D. L. 2005 *Magnetic Free Energy in NOAA Active Region 10486 on 2003 October 29*, ApJ 623, L53

Metcalf, T. R., Jiao, L., McClymont, A. N., Canfield, R. C., and Uitenbroek, H. 1995, *Is the Solar Chromospheric Magnetic Field Force-Free?*, ApJ 439, 474

Metcalf, T. R., DeRosa, M. L., Schrijver, C. J., Barnes, G., vanBallegooijen, A. A., Wiegelmann, T., Wheatland, M. S., Valori,G., and McTiernan, J. M. 2008, *Nonlinear Force-Free Modeling of Coronal Magnetic Fields. II. Modeling a Filament Arcade and Simulated Chromospheric and Photospheric Vector Fields*, Sol. Phys. 247, 269

Molodensky, M. M. 1974, Equilibrium and Stability of Force-Free Magnetic Field, Sol. Phys. 39, 393

Parker, E. N. 1988, *Nanoflares and the Solar X-ray Corona*, ApJ 330, 474

Regnier, S. and Priest, E. R. 2005, *Self and Mutual Magnetic Helicities in Coronal Magnetic Configurations*, A&A 442, 345

Saha, M. N. 1942, *The Solar Corona*, Nature 149, 524

Schmidt, H. U. 1964, On the Observable Effects of Magnetic Energy Storage and Release Connected With Solar Flares, in The Physics of Solar Flares (ed.: Hess, W. N.), NASA SP-50, Washington: NASA, p.107

Schrijver, C. J., DeRosa, M. L., Metcalf, T. R., Liu, Y., McTiernan, J., Regnier, S., Valori, G., Wheatland, M. S., and Wiegelmann, T. 2006, *Nonlinear Force-Free Modeling of Coronal Magnetic Fields Part I: A Quantitative Comparison of Methods*, Sol. Phys. 235, 161

Schrijver, C. J. et al. 2008, *Nonlinear Force-free Field Modeling of a Solar Active Region around the Time of a Major Flare and Coronal Mass Ejection*, ApJ 675, 1637

Sivaraman, K. R. and Livingston, W. C. 1982, *CA II K2V Spectral Features and their Relation to Small-Scale Photospheric Magnetic Fields*, Sol. Phys. 80, 227

Socas-Navarro, H. 2005a, *Are Electric Currents Heating the Magnetic Chromosphere?*, ApJ 633, L57

Socas-Navarro, H. 2005b, *The Three-dimensional Structure of a Sunspot Magnetic Field*, ApJ 631, L167

Sudol, J. J., and Harvey, J. W. 2005, *Longitudinal Magnetic Field Changes Accompanying Solar Flares*, ApJ 635, 647

Tomczyk, S., McIntosh, S. W., Keil, S. L., Judge, P. G., Schad, T., Seeley, D. H., and Edmondson, J. 2007, *Alfvén Waves in the Solar Corona*, Science 317, 1192

Wiegelmann, T., Inhester, B., and Sakurai, T. 2006, *Preprocessing of Vector Magnetograph Data for a Nonlinear Force-Free Magnetic Field Reconstruction*, Sol. Phys. 233, 215

Wiegelmann, T., Thalmann, J. K., Schrijver, C. J., DeRosa, M. L., and Metcalf, T. R. 2008, *Can We Improve the Preprocessing of Photospheric Vector Magnetograms by the Inclusion of Chromospheric Observations?*, Sol. Phys. 247, 249

Advanced Computational Capabilities for Exploration in Heliophysical Science (ACCEHS) --- a Virtual Space Mission

A. Bhattacharjee,[1] K. Schrijver, W. Abbett, F. Cattaneo, L. Chacon, C. Cohen, W. Daughton, P. Erickson, L. Fisk, K. Germaschewski, J. Giacalone, T. Gombosi, M. Hesse, J. Huba, M. Hudson, V. Jordanova, P. Judge, H. Karimabadi, G. Lapenta, A. Lazarian, Y. Lin, J. Linker, H. Liu, J. Luhmann, M. Miesch, M. Moldwin, N. Omidi, T. Onsager, M. Opher, M. Oppenheim, J. Raeder, A. Ridley, P. Riley, R. Schunk, M. Velli, M. Wiltberger, and G. Zank

Executive Summary

In the dynamically complex, nonlinearly coupled domains of heliophysics, the efficient use of computers is as important as access to state-of-the-art *in situ* and remote-sensing instrumentation. Computers help us explore where we cannot sense, build a comprehensive view of very sparsely sampled environments, provide forward modeling where inversions fail, and put the heliophysical domain into a controllable laboratory setting. Computers are also are critical in processing the terabytes of data coming from both real-world and virtual-world experiments.

With the development of computer hardware supported by large economic interests in industry, we can leverage these investments with a fraction of their true cost to rapidly move beyond the decades-old legacy codes we now mostly work with, improve the foundations and capabilities of the underlying physical models, integrate across physical domains, and reach out across wide ranges in scales. This is societally relevant as we work towards more realistic forecasting of space and terrestrial weather and climate, capitalizing on advances in computer architecture and data storage to increase the capability of our workforce.

As discussed in a community-wide workshop that is the foundation for this concept paper, multiple critical problem areas in heliophysics are ready to be moved forward through advanced computational capabilities, with benefits across multiple sub-disciplines and with value to society. Examples of such transformational projects include:

- Generation and emergence of active regions from the convective envelope of the Sun;
- The evolution of the ambient solar corona and its coupling to the inner heliosphere;
- Understanding and parameterizing the kinetic solar wind from the inner to the outer edges of the heliosphere;
- Acceleration and transport of energetic particles in the heliosphere;
- Impact of severe storms on the geospace environment;
- Accurate models of the radiation belts;
- Ionosphere-thermosphere-magnetosphere (ITM) coupling;
- Prediction of communication outages due to ionospheric density irregularities and turbulence.

Recommendation: We recommend that NASA, perhaps in partnership with NSF and other agencies, should lead by establishing a new peer-reviewed program in which critical-mass groups of heliophysicists, computational scientists, and applied mathematicians can be brought together to address transformational science questions that support its flight missions and advance heliophysics. The level of synergistic collaboration of such groups, requiring $5-10M per project spread over several years, exceeds what is currently supported within GI, TR&T, HTP, or SR&T program envelopes. This new program should eventually support up to five groups with steady funding for periods of up to five years at an approximate level of $2M per year each. These groups should primarily pursue a scientific problem, advancing through reviewed milestones to design and use tools for theory, modeling, and data assimilation that will exploit the capabilities of state-of-the-art supercomputers. In doing so, they should also support the training of postdoctoral scientists and graduate students in advanced computational capabilities.

While the concept underlying this paper has taken on special urgency in recent years, the 2003 Decadal Survey Report emphasized the important role of computation in its "Coupling Complexity Research Initiative". That Report identified some key issues that remain as valid today as they were nearly a decade ago, including: (i) the importance of clearly articulated support and funding lines for model development, (ii) the need for computer hardware, which should be treated like hardware for experimentation, (iii) the necessity of science questions driving computational initiatives, and (iv) the importance of supporting data assimilation and exploration techniques in synergy with computational model development.

1. Background

Computational modeling, simulation, and data analysis have been among the most important drivers of scientific discovery during the last three decades. This progress has been enabled by remarkable leaps in computing technologies, producing parallel computers of great power and speed which have been brought to bear on increasingly sophisticated software and efficient algorithms. It is reasonable to anticipate a near-future transition from the present-day terascale and petascale systems to the exascale, which will empower the science community further to undertake challenges that can be potentially transformational.[2]

The heliophysics science community is experiencing a rapid and radical transformation because of vast increases in the sophistication of instruments and in data volumes. Progress in our understanding of heliophysical processes requires that data analysis be combined with computational modeling, numerical simulation, and data-assimilation programs. The effectiveness of our community-wide theory, modeling, and data assimilation and analysis efforts depends critically on the development of innovative numerical algorithms and their use on high-performance computing platforms, as we see happen in other scientific disciplines (such as astrophysics, high-energy and nuclear physics, plasma and fusion science, and climate prediction and change), which are well-positioned to exploit fully the power of new computing technologies. We cannot rely on the slow diffusion of the fruits of the efforts made by other scientific and engineering disciplines, but need to actively work on advancing our discipline's capabilities to meet the demands for scientific breakthroughs, the design of next-generation space missions, and to tap into the pool of new and young talent, eager to bring the power of new computing technologies and methodologies to bear upon heliophysical science challenges.

A community-wide Workshop on Advanced Computational Capabilities for Exploration in Heliophysical Science (ACCEHS), was held on August 16-18, 2010 at NCAR in Boulder, Colorado (http://www.hao.ucar.edu/ACCEHS/). The Workshop brought together over 80 scientists in heliophysical science, as well as experts from the climate and computer science communities. In what follows, we describe the key findings of the Workshop, cast in the form required by a concept paper for the Decadal Survey, recommending possible strategies for action summarized above. A more detailed Report for the ACCEHS Workshop is currently being written.

2. Key Challenges and Opportunities in Heliophysical Science Disciplines

The concept underlying this paper touches on all three elements of the present Decadal Survey: Atmosphere-Ionosphere-Magnetosphere Interactions, Solar Wind-Magnetosphere Interactions, and Solar and Heliospheric Physics. The primary goal of this section is to articulate challenges or questions where targeted investment in advanced computational capabilities has the potential to transform heliophysical science. Our goal is *not* to assemble a comprehensive list, and certainly not to imply a prioritization. The topics identified in this Section should be viewed as examples of challenges that can take heliophysical science to the next and higher level of discovery.

2.1 Generation and emergence of active regions from the convective envelope of the Sun

Magnetic activity on the Sun originates from hydro-magnetic dynamo action in its highly turbulent convective envelope. A unified, comprehensive understanding of the diverse range of magnetic activity exhibited by the Sun, from small-scale flux elements in the photosphere to global patterns of magnetic activity such as the sunspot cycle, remains an unfulfilled challenge. In particular, understanding how sub-surface magnetic flux emerges into the solar atmosphere, energizes the corona, shapes the heliosphere, and regulates space weather is an essential prerequisite in the effort to understand solar variability and its impact on the Earth and other planets of our solar system.

Modeling the generation and emergence of magnetic flux on the scale of active regions requires a realistic physical description of the complex boundary layers that straddle the Sun's convection zone. Near the base of the convection zone, large thermal, rotational, and magnetic gradients promote the generation, accumulation, and subsequent destabilization of magnetic structures that can buoyantly rise to the photosphere. As these structures pass through the visible surface and expand into the solar corona, they traverse a thin boundary layer where the physical environment changes drastically. Plasma densities drop by ten to twelve orders of magnitude, radiation and electron thermal conduction replace convection as the primary means of energy transport, and magnetism overtakes gas pressure as the dominant energy reservoir. The physics of this transition is particularly complex in the solar chromosphere, where shifting ionization states and non-local radiative processes become energetically important. Similarly, the physics of the boundary layer at the base of the convection zone is complicated by internal and interfacial wave

modes and tachocline instabilities.

The challenge, then, is to develop the theoretical and computational techniques necessary to model these physically, spatially, and temporally disparate regimes in a way that retains the essential physics well enough to interpret current and future ground and space-based observations. Currently, global convection simulations exhibit differential rotation and large-scale dynamo action but do not have sufficient spatial resolution to capture the formation and destabilization of concentrated flux structures. Local flux emergence simulations adequately capture the radiative MHD of the upper convection zone but rely on idealized initial conditions and simplified, spatially-limited models of the chromosphere and corona. A unified, higher-fidelity model is within reach but will require sophisticated, coupled, numerical algorithms that can fully exploit next-generation computational resources.

2.2 The evolution of the ambient solar corona and its coupling to the inner heliosphere

Many of the essential questions of solar physics are reflected in those about the structure of the solar corona. What mechanism(s) heat the corona? How is the fast solar wind accelerated? What is the origin of the slow solar wind? An understanding of coronal structure is not only important in its own right, but is implicit to understanding the geomagnetic effects of CMEs.

Phenomena in the solar corona and solar wind occur at many different scales. Observations of smaller scale phenomena give us clues to the fundamental processes that heat the solar corona and accelerate the solar wind. Models of the processes that drive these phenomena (e.g., reconnection, wave heating, and acceleration) are highly idealized but allow us to explore the viability of different mechanisms. Large-scale observations of the corona show us the consequences of the small-scale behavior –coronal streamers, coronal holes, the slow and fast wind. Global models of the corona and wind can capture much of this behavior, but must use empirical prescriptions or parameterizations of the smaller-scale physics to obtain realistic results – these constraints are imposed by the lack of resolution available to incorporate more basic physical ideas. These disjoint treatments make it difficult to test and improve theoretical models.

To understand which dominant physical mechanisms shape and power the solar corona and solar wind, global models must incorporate the small-scale processes as consistently as possible so that we can directly test theories with observations. Current state-of-the-art models cannot resolve the effects of random photospheric motions, which shuffle the footpoints of the coronal magnetic field and allow closed field lines in the streamer belt to interact with nearby open field lines in the coronal holes. Will significant reconnection occur? Can the resulting plasma release account for properties and extent of the slow solar wind? The introduction of these dynamics require ~10-100 times the number of grid points to resolve the subsequent scales. Breakthrough calculations (10^9 grid points and beyond) are becoming tractable as computers transition to megacore architectures. A major model development program is urgently required to take advantage of this opportunity, because present-day codes cannot be simply extended to the new state-of-the-art machines. Extensions to the physics of the models are also necessary to accomplish these groundbreaking simulations. For example, multi-fluid calculations with ionization equilibrium equations for the different species will allow the composition of the solar wind to be predicted from the models and tested against observations. Kinetic physics may very well be necessary to completely understand solar wind heating and acceleration, and may eventually need to be incorporated via coupling of kinetic simulations to MHD models.

2.3 Understanding and parameterizing the kinetic solar wind from the inner to the outer edges of the heliosphere

Although our well-developed and widely applied MHD 'fluid' concept of the solar wind has served us for decades as a reliable means of describing the macroscopic features of the interplanetary medium, it is inadequate to fully understand the details of the underlying behavior of the ions and electrons that make up the solar wind that the Earth is constantly exposed to. Numerical simulations of the kinetic processes involved in solar wind heating/acceleration and the wave-particle interactions that pervade the heliosphere are currently confined to local or highly simplified global treatments and cannot capture the effects of the boundary conditions and radial evolution that are so key to understanding the solar wind. Large-scale kinetic simulations of solar wind physics, accommodating more aspects of both ion and electron distribution functions in realistic coronal and interplanetary field geometries have the potential to revolutionize common working paradigms of heliophysics. For example, they can lead to improved parameterized descriptions of solar wind heating in the increasingly realistic global heliospheric simulations. They can lead to more accurate interpretation and use of ion and electron thermal anisotropies

and halo electron strahl.

At the farthest edges of the heliosphere, the solar wind interacts with the local interstellar medium where the distribution of both the plasma and neutrals can be highly anisotropic and non-Maxwellian, leading to significant effects on the global heliosphere. Moreover, from observations returned by Voyager and IBEX spacecraft mapping the heliosheath and the global structure, it is clear that the structure of the magnetic field plays a crucial role in this region. These observations reinforce the need to resolve disparate physical scales and kinetic processes. An example is the behavior of the heliospheric current sheet close to the heliopause. In the heliosheath, the sector regions approach each other as the solar wind slows down. The current state-of-the-art relies on idealized models that largely neglect the complexities of the heliospheric current sheet and the intrinsic time dependence of the solar wind. The demands of modeling extremely fine scales within a 3D time-dependent MHD code over very long times, couple with the need to model the kinetic physics of neutral hydrogen and plasma processes over several solar cycles makes this one of the most challenging computational problems in space physics.

2.4 Acceleration and transport of energetic particles in the heliosphere

Understanding how solar energetic particles (SEPs) are accelerated and transported in the heliosphere is a long and outstanding problem that is at the heart of space weather and cannot be attacked without a larger effort than is currently employed. Presently, only portions of the entire problem are being addressed through separate modest numerical simulations, e.g., energy release and magnetic reconnection at the Sun, shock formation, particle acceleration by interplanetary shocks and magnetic reconnection, and particle transport through the heliosphere. Coupling these simulations, e.g., incorporating particle acceleration into a simulation of coronal mass ejection eruption, has proven to be difficult and generally beyond the currently available resources (both in terms of computational resources and in terms of funding of a team of experts for a significant amount of time).

Similar to the problem of understanding the solar wind in the heliosphere, the acceleration of SEPs requires incorporating kinetic physics and a large range of physical scales into numerical simulations. A prime example is the question of how thermal, or very low-energy particles, that move slowly with respect to the shock are accelerated by the shock to high energies, is largely ignored in current modeling. This problem involves scales that range from the gyroradius of the thermal particles (~100 km at 1 AU) to the scattering mean-free path of the highest energy particles (~> 1 AU for ~GeV protons at 1 AU). The kinetic physics is important because the shocks that accelerate the particles move slowly enough that they are affected by the microstructure within the shock layer. Current models either ignore the shock microstructure and consider the acceleration to the highest energies only, or include the microstructure but are limited in the size of the model and therefore are limited in terms of the maximum energy achieved.

2.5 Impact of severe storms on the geospace environment

The magnetosphere represents the furthest extent of Earth's environment into the surrounding plasma and electromagnetic fields of interplanetary space. It stands between the solar wind and the ionosphere-atmosphere system and thereby controls the flow of mass, momentum, and energy. The magnetosphere spans a huge volume that is highly under-sampled, due to the small number of satellites making in-situ measurements and limited remote sensing opportunities. Understanding the magnetosphere is further complicated by the dominance of coupled, nonlinear physical processes that cover orders of magnitude in temporal and spatial scales. Conditions in the magnetosphere can change in a matter of minutes from quiet-time to storm-time, and storms can last from hours to days. Large geomagnetic storms can cause serious damage to technological systems like satellites, increase the radiation exposure of astronauts, and disrupt communication and power systems. Predicting geomagnetic storms (especially severe storms) is one of the biggest challenges in magnetospheric physics.

Given the complex and multi-scale nature of the magnetosphere and its coupling to the solar wind and the ionosphere-atmosphere system, understanding its dynamics requires state-of-the-art models, computational facilities, and access to comprehensive data sets. The two most promising approaches to understanding the global behavior of the magnetosphere are MHD/multi-fluid and hybrid simulations. MHD simulations have been in use for more than 3 decades and advances in hardware and software technology have resulted in their ongoing improvement. These models, however cannot accurately describe the inner magnetosphere because they do not include the energy-dependent drifts of the pressure bearing ring current particles. Robust coupling of MHD models with kinetic ring current models is difficult to achieve at present due to the required higher temporal and spatial resolution. Hybrid models treat the ions

kinetically and allow coupling of microscopic and macroscopic processes, but due to their considerably larger computational expense, hybrid simulations currently utilize a system size that is smaller than the magnetosphere. Both MHD/multifluid and hybrid models need to include ionospheric ion outflow models that will allow coupling between the magnetosphere and the ionosphere under different solar wind driving conditions. Advances in hardware technology and supporting computational libraries will allow future development of near-real-time predictive capabilities for geomagnetic storms.

2.6 Accurate models of the radiation belts

The dynamic variability of radiation belt electrons over orders of magnitude poses a challenge to modelers and a threat to our increased dependence on satellite systems vulnerable to these changes. A thorough understanding of the mechanisms involved in their access, transport, trapping, acceleration, and loss can be achieved only with a system level approach, including coupling between adjacent regions. Trapped MeV electron flux is highly variable, peaking during the declining phase from solar maximum, with strong enhancements around the ~ 11 year maximum in sunspot number. The low energy (eV) plasmasphere population which co-rotates with the Earth is continuously refilled by the topside ionosphere and stripped away by changing magnetospheric convection, typically around L ~ 4. Our understanding of the coupling of the solar wind drivers to the dynamics of the low energy plasmasphere population, and its effects on MeV electron fluxes, requires advanced modeling of wave-particle interactions near the plasmapause, responsible for determining the depth of penetration of MeV electron fluxes into the altitude range of, for example, GPS spacecraft. Recent observations of coherent large amplitude whistler waves in this region responsible for electron loss and local acceleration have been limited by time resolution of earlier measurements. Advances in spacecraft data acquisition at higher time resolution present a data storage and processing challenge similar to that faced by modelers.

Local micro-scale processes such as the generation of plasma waves and their effect on particle dynamics must be included in global models. Major computational challenges are thus to develop models that couple self-consistently the plasma and the fields across various regions of the magnetosphere, and that include both large-scale and micro-scale physics. This could be achieved by coupling global MHD and/or multi-fluid models with kinetic or full particle models, which each address different regions or physical processes. Including ionospheric outflow as well as solar wind entry in such models is essential, as such outflow/inflow repopulates the plasmasheet, the source of both energetic electron and ion components. Inclusion of plasmasphere dynamics in global models is essential, as the plasmapause has been shown to be a significant boundary affecting both radial transport and localized heating and loss of radiation belt electrons and ring current ions. Thus, particle energies span approximately six orders of magnitude over which particle dynamics must be modeled in a combined fluid and particle approach. Cross-coupling of particle populations over such a broad energy range, combined with the physical scale of the coupled system and range of timescales between global processes and localized acceleration and pitch angle scattering present major computational challenges. The development of data assimilation models will allow the integration of ground-based and space-borne data sets with models, soon to be augmented by the first spacecraft mission, RBSP, dedicated to radiation belt studies in two solar cycles.

2.7 Ionosphere-thermosphere-magnetosphere (ITM) coupling

It has become increasingly clear over the past several years that coupling processes between the lower atmosphere, thermosphere, ionosphere, and magnetosphere have a much larger impact on microscale to mesoscale behavior than previously understood. The major computational challenges are to develop accurate and robust numerical models for each region, and to seamlessly couple these models into a single, self-consistent unified model. The development of whole atmosphere models in recent years reflects these needs. Other examples include frameworks that link together models of the magnetosphere, ionospheric electrodynamics, and upper atmosphere.

An example of neutral atmospheric dynamics on the behavior of the ionosphere is the phenomenon of sudden stratospheric warming. During these events, lower atmospheric waves, such as planetary, tidal, and gravity waves, undergo significant variation and produce changes in the stratosphere, mesosphere, and thermosphere that significantly impact the electrodynamics of the ionosphere. Current modeling capabilities are inadequate to fully explain and predict these behaviors, in many cases due to the need for expanded resolution and improved boundary dynamics. Increased resolution and improved physics is critical to correctly quantify large scale waves (i.e., tides and planetary waves) and to resolve mesoscale

waves (e.g., gravity waves). These waves have important implications for ionospheric variability and space weather, but not reproducible by current global models.

2.8 Predict navigation and communication outages due to ionospheric density irregularities and turbulence

A classic example of this challenge is equatorial spread F (ESF), during which the equatorial ionosphere becomes Rayleigh-Taylor unstable: large (tens of km) electron density bubbles develop and rise to high altitudes (1000 km or greater at times). Attendant with these large-scale bubbles is a spectrum of density irregularities that can extend to wavelengths as short as 10 cm. Understanding and modeling ESF is important because of its impact on space weather: the associated electron density irregularities can cause radio wave scintillation that degrades communication and navigation systems.

The first major computational problem is to seamlessly couple different spatially overlapping, physics models. Large-scale (hundreds of m) ionospheric processes are well described by fluid theory. However, a kinetic description of the plasma is needed to model small-scale wave structures (~100 m) and the turbulent cascade of energy down to tens of cm. The computational algorithms to allow self-consistent coupling of fluid and kinetic models, which need to be developed, are crucial to describe the onset and decay of scintillation-causing irregularities. The second major computational problem is modeling the electrodynamics of the IT system. All global models of the IT system assume that the geomagnetic field lines are equipotentials; this reduces the potential equation to two dimensions. To accurately describe the self-consistent coupling of large scale to small- scale density irregularities, it is necessary to have a fully electrodynamics model. This requires the development of a fully three-dimensional solver that is robust, efficient, and uses parallel processing on a non-uniform grid. The third major computational problem is to couple computational models over a large range of spatial and temporal scales. The nominal spatial grid scales for current IT system models are 2.5-5.0 degrees in latitude and longitude (i.e., 300-500 km) and 2-10s km in altitude. The nominal time scale ranges from ~10s to ~10 min. To capture scintillation-producing ionospheric irregularities in a computational model, the grid resolution has to be reduced to ~1 m, which is about three orders of magnitude smaller than is currently feasible.

3. Computational Needs and Opportunities

Progress on some of the most exciting scientific challenges in heliophysics is awaiting the development of new computational techniques and the full utilization of petascale computing. Looking across the various sub-fields, the basic requirements for new tools and approaches share many common themes. Researchers seek to understand systems with vast disparities in spatial and temporal scales. This requires accurate models governing the large-scale evolution, which may depend on a kinetic processes. New petascale machines and other advanced hardware such as general-purpose graphical processing units (GPU) offer the potential for tremendous progress on these multi-scale problems. The ability to fully exploit these computers will permit calculations (100-1000) times larger than previous efforts. However, the spatial and temporal scale separations are so large that simply adapting existing algorithms to the new hardware is not sufficient. New breakthroughs will also require improvements in multi-scale algorithms and adapting the modern advancements in algorithms. In addition, new approaches are needed to deal with the large quantity of data generated by both simulations and observations. We have identified four basic categories of computational tools to address these needs:

3.1 High performance solvers and advanced algorithms for fluid simulations

Fluid simulations play a critical role in all sub-fields of heliophysics, including neutral fluids in the atmosphere/thermosphere, electrostatic fluid models in the ionosphere, and MHD models in magnetospheric, interplanetary and solar physics. These range from 1D regional models (such as solar wind acceleration), to 2D codes (such as magnetotail reconnection models) to fully 3D global codes (such as global magnetosphere and heliosphere codes). Most of these codes were written by scientists with deep understanding of the physics but limited background in modern algorithms, and before massively parallel machines were widely available.

Efforts to develop modern fluid simulation capabilities should include three key components: improved physics and closures, advanced solvers and the ability to fully utilize existing and future computer hardware. Depending on the specific science focus, the optimum mixture may differ. For example, how to properly capture kinetic effects within fluid calculations or describe the effects of turbulence remain key issues for accurately modeling magnetic reconnection, but improvements in solver technology and/or petascale computers are also very important for performing the high-resolution

simulations required to test improved fluid closures and subgrid methods. Most heliophysics simulations are physically far under-resolved, and would benefit greatly modern adaptive mesh refinement (AMR) methods and implicit time stepping. However, in order to simulate physically relevant large systems, the parallel and algorithm scalability is crucial. For explicit codes, parallel scalability is typically quite good using simple domain decomposition, and remains good for enhanced methods like block-structured AMR. While more difficult, recent work has demonstrated that implicit MHD algorithms showing such optimal algorithmic properties are possible. These issues are also important for electrostatic fluid treatments in the ionosphere, where there is a critical need for highly scalable 3D Poisson solvers on non-uniform meshes.

3.2 Advanced kinetic simulation approaches and algorithms

While fluid models have been very successful in modeling the macroscopic structure of many phenomena, there are many important science questions in heliophysics that require a kinetic description. A range of different kinetic simulation approaches have been employed including fully kinetic (ions and electrons), hybrid (fully kinetic ions and fluid electrons) as well as gyrokinetic models. Most kinetic simulations still employ a uniform mesh, with simple explicit algorithms. On the positive side, these simple well-tested algorithms can achieve excellent performance and scaling on modern computer architectures. Recent 3D fully kinetic simulation studies of magnetic reconnection have employed over ~1 trillion computational particles running on 10^5 computational cores, with an additional factor of ~10 anticipated in the next few years. These capabilities are leading to a range of new insights into the physics of reconnection, and may be extremely useful for addressing basic physics issues in shocks and turbulence. However, for many problems the spatial and temporal scale separations are so large, that simple explicit approaches (with uniform mesh) will always be limited due to rigid stability constraints and the long-time scale accuracy of the solution. While investments are clearly needed to help researchers exploit new architectures, there is an equal need for targeted investments in algorithms and multi-scale techniques. This includes, asymptotic-preserving discrete numerical formulations, non-uniform structured mesh, AMR, and implicit time stepping and/or discrete-event multi-stepping techniques. The scientific payoff will be high if researchers can effectively combine the raw power of petascale computing with modern advances in algorithms.

3.3 Computational framework: coupling physics across disparate scales

Development of a general simulation approach to handling multi-physics, multi-scale problems described by different models (codes) remains an outstanding research topic in computational physics. Problems in heliophysics will likely require software frameworks that include a superstructure layer that drives the coupled-model application, and an infrastructure layer that provides utilities and data structures for model developers. There are an abundance of heliophysics applications that will benefit from new investments in these capabilities. As one example, researchers have already successfully coupled ionosphere/thermosphere dynamics to global magnetospheric MHD codes. With petascale computers, global hybrid calculations are becoming increasingly realistic, but will still need to be coupled to ionospheric models. While hybrid calculation can properly treat the structure of ion-scale transition layers and shocks, neither approach (hybrid or MHD) can correctly capture the microphysics of magnetic reconnection. Future efforts in computational frameworks may allow researchers to couple more complete kinetic descriptions within the thin current sheets, and thus move towards a more predictive understanding of the global dynamics.

3.4 Analysis tools, parallel visualization and data mining and assimilation

Both simulations and space observations are producing increasingly larger data sets that are challenging for researchers to study with existing tools and approaches. New missions such as the Solar Dynamic Observatory have a data rate of 1.4 terabytes per day, while recent 3D kinetic simulations using the newest petascale computers are generating ~40 terabytes in a single run. Extracting new science from these enormous data sets is becoming progressively more difficult. Researchers are drowning in the over abundance of data, and legacy tools and approaches have failed to keep pace. The largest simulations are preformed at a few leadership class facilities around the country, and the resulting data sets are now far too large to move. To utilize these new machines, researchers need parallel tools to remotely visualize and manipulate the data, and to easily allow access to teams of researchers working in a collaborative fashion.

Although there have been efforts in recent years in extending simulation capabilities to petascale computers, less attention has been paid to the data analysis and visualization of the resulting large data sets.

In terms of spacecraft data, researchers have introduced data mining techniques and have worked with experimentalists in their analysis. This effort has met with considerable success and is rapidly gaining adoption. However, funding is required for further expansion of these techniques, as well as the integration of these new methods with the existing data analysis software used by experimentalists (e.g., autoplot, VO, etc.).

References

1. *Contact information*: Space Science Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH 03824, e-mail: amitava.bhattacharjee@unh.edu.
2. See the "Scientific Grand Challenges Workshop Series" on extreme scale computational science, organized by the Office of Science, Department of Energy (http://extremecomputing.labworks.org/index.stm).

The Solar Spectroscopy Explorer Mission

Jay Bookbinder[1]§, J. Adams[3], D. Alexander[8], M. Aschwanden[12], C. Bailey[3], S. Bandler[3], S. Bradshaw[8], N. Brickhouse[1], J. Chervenak[3], S. Christe[3], J. Cirtain[3], S. Cranmer[1], S. Deiker[12], E. DeLuca[1], G. Del Zanna[13], B. Dennis[3], G. Doschek[2], M. Eckart[3], F. Finkbeiner[3], A. Fludra[15], P. Grigis[1], R. Harrison[15], L. Ji[1], C. Kankelborg[14], V. Kashyap[1], D. Kelly[3], R. Kelley[3], C. Kilbourne[3], J. Klimchuk[3], Y.-K. Ko[2], J. M. Laming[2], E.. Landi[6], M. Linton[2], D. Longcope[14], J. Mariska[2], D. Martinez-Galarce[12], H. Mason[13], D. McKenzie[14], R. Osten[17], G. Peres[10], A. Pevtsov[16], K. Phillips[4], F. S. Porter[3], D. Rabin[3], C. Rakowski[2], J. Raymond[1], F. Reale[10], K. Reeves[1], J. Sadleir[3], D. Savin[11], J. Schmelz[9], R. K. Smith[1], S. Smith[3], R. Stern[12], J. Sylwester[5], D. Tripathi[13], I. Ugarte-Urra[7], P. Young[7], H. Warren[2], B. Wood[2]

[1]Smithsonian Astrophysical Observatory, [2]Naval Research Laboratory, [3]NASA Goddard Space Flight Center, [4]Mullard Space Science Laboratory, UCL, [5]Space Research Centre, Polish Academy of Sciences, [6]University of Michigan, [7]George Mason University, [8]Rice University, [9]University of Memphis, [10]University of Palermo, [11]Columbia University, [12]Lockheed Martin, [13]DAMTP, University of Cambridge, [14]Montana State University, [15]STFC Rutherford Appleton Laboratory, [16]National Solar Observatory, [17]Space Telescope Science Institute

EXECUTIVE SUMMARY

The Solar Spectroscopy Explorer (SSE) concept is conceived as a scalable mission, with two to four instruments and a strong focus on coronal spectroscopy. In its core configuration it is a small strategic mission ($250-500M) built around a microcalorimeter (an imaging X-ray spectrometer) and a high spatial resolution (0.2 arcsec) EUV imager. SSE puts a strong focus on the plasma spectroscopy, balanced with high resolution imaging – providing for break-through imaging science as well as providing the necessary context for the spectroscopy suite. Even in its smallest configuration SSE provides observatory class science, with significant science contributions ranging from basic plasma and radiative processes to the onset of space weather events. The basic configuration can carry an expanded instrument suite with the addition of a hard X-ray imaging spectrometer and/or a high spectral resolution EUV instrument – significantly expanding the science capabilities. In this configuration, it will fall at the small end of the medium class missions, and is described below as SSE+. This scalable mission in its largest configuration would have the full complement of these instruments and becomes the RAM (Reconnection And Microscale) mission. This mission has been designed to address key outstanding issues in coronal physics, and to be highly complementary to missions such as Solar Probe Plus, Solar Orbiter, and Solar-C as well as ground-based observatories.

The rationale for a scalable concept was to allow the Decadal committee to select the most appropriate size mission for additional study, once budget constraints are understood. The nominal mission design life is three years, with consumables sized for 10 years. The observatory design is built on studies performed over the last decade by SAO, GSFC, and several commercial spacecraft providers, and has strong heritage in all components from previous space flight missions. This submission is one of a trio of white papers discussing how new instrumentation (Bandler et al. 2010) can provide breakthroughs in heliophysics (Laming et al. 2010).

1. SUMMARY OF SCIENCE CONCEPT

The Solar Spectroscopy Explorer (SSE) addresses fundamental and timely questions in solar physics. Solar physics, like much of heliophysics, is a mature science. No single instrument can comprehensively address the key science questions outlined in the 2009 Roadmap or the 2003 Decadal Survey. SSE's scalable approach identifies the smallest instrument complement that can address a set of strategic science questions and outlines how additional instrument(s) expand the capabilities of the mission. The reader is referred to the numerous white papers in particular that discuss the breadth of science opportunities achievable via high-resolution X-ray spectra and imaging (e.g. Laming et al. 2010, Golub et al. 2010, Lin et al 2010, Smith et al. 2010, Osten et al. 2010, Kashyap et al. 2010) as well as instrument white papers on the XMS (Bandler et al., 2010b), FOXSI (Christe et al. 2010), and FACTS (Korendyke et al. 2010).

§Contact information: jbookbinder@cfa.harvard.edu; 617-495-7058

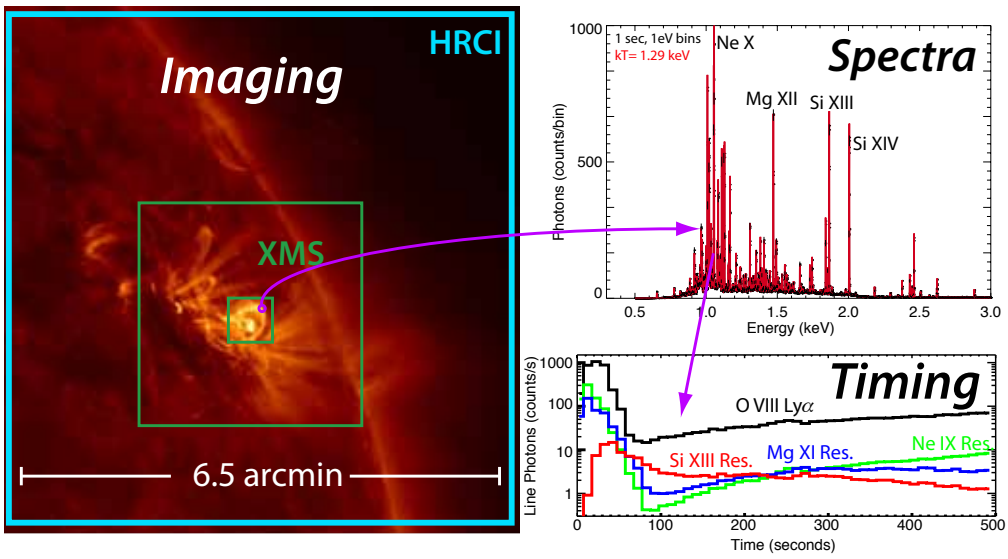


Figure 1 –SSE combines imaging, spectra, and timing, as it can image regions at arcsecond resolution over a wide field of view, creating high-resolution spectra at each point and tracking the evolution of individual emission lines as a function of time. Shown here is a post-CME active region observed at 171Å with SDO/AIA, with the FOV of the core & extended XMS array (in green) as well as the HRCI (in blue) overlaid. A 1s spectrum from a 100 sq. arcsec extraction region is shown, along with one possible evolution of each line's flux over a 100 minute observation. The FOXSI FOV (10×10 sq. arcmin) extends beyond the image.

The corona radiates most of its energy in the soft X-rays, emitting resolvable lines that provide a wealth of spectral diagnostics. These soft X-ray emission lines tell the story of how the plasma is heated, where the plasma is heated, and for how long the plasma has been heated via measurements of the temperature, density, and time-dependent ion populations as well as departures from Maxwellian velocity distributions. These will be combined with measurements of flow and turbulent velocities from line profiles and line shifts. These are the measurements that are necessary to understand the physics of magnetic reconnection, coronal heating, flare and CME energetics, and solar wind acceleration (Table 1).

2. ADVANCING SOLAR PHYSICS IN THE NEXT DECADE

Unlike previous missions, SSE will provide excellent imaging, spectral, and timing information simultaneously. One of the two core instruments, the X-ray Microcalorimeter Spectrometer (XMS), will enable measurements with unprecedented spectral (<2 eV) and temporal (milliseconds) resolution over a broad spectral range (0.2-10 keV). Unlike narrow-band or single-line imagers, all of the coronal plasma can be seen, enabling comprehensive studies of its evolution – critical information as the corona is dynamic, necessitating that all of the diagnostic lines be observed simultaneously for a complete picture. The high temporal resolution of the XMS will characterize flare energy distributions with unprecedented accuracy, in a manner that has been applied successfully for other stars (Kashyap et al. 2002). These are known to be distributed as a power-law over many orders of magnitude (Hudson 1991); photon counting enables sophisticated analyses (developed for low signal stellar data) that explore a variety of physics at different timescales and energy ranges.

This X-ray spectroscopy capability is complemented by the second core instrument: the high resolution coronal imager (HRCI), a normal incidence system at the 193Å (Fe XII, XXIV; log T=6.1, 7.3) and 131Å (Fe XX; Log T=7) passbands that provides high cadence and high spatial resolution (0.1" pixels) imaging over a ~7×7 sq. arcmin FOV (~14×14 sq. arcmin goal). This is a factor of five better imaging resolution than currently available on AIA. The HRCI allows us to follow the evolution of the fine-scale coronal sub-structure, whose scale is based on spectroscopic observations that indicate that ~10% of the observed area is filled with emitting plasma (see Golub et al. 2010 and Doschek 2010). The associated topological and dynamical contributions from the HRCI are shown in Table 1. The HRCI is particularly sensitive to low amplitude transverse oscillations that may result from small impulsive events or waves that originate in the lower atmosphere.

Table 1 – Science Flowdown for SSE & SSE+ Mission Options

Science			Observatory Requirements				
Topic	**Target**	**Physical Process**	**Spatial Resol.**	**Spectral Resolution**	**Timing**	**Measurement**	**Instrument**
			arcsec	**ΔE@E (keV)**	**seconds**		
Magnetic Reconnection	Active Regions	Heating	1	0.002@1	10	$T_e(s),n_e(s)$	XMS
		Flows	1	0.002@1	10	$v(T_e,s,t)$	XMS,HRCI
		Particle Acceleration	10	1@15-30	~0.001	$dn_e(s)/dE$	FOXSI (Pri)
			1	0.002@1	1	$dn_e(s)/dE$	XMS (Sec)
		Dynamics/Topology	0.2	–	1	v(s,t),I(s,t),connectivity	HRCI
FIP Fractionation/ Coronal Heating	Active Regions & Coronal Loops	Heating	1	0.002@1	10	$T_e(s),n_e(s)$	XMS
		Flows	1	0.002@1	10	$v(T_e,s,t)$	XMS,HRCI
		Dynamics/Topology	0.2	–	2	v(s,t),I(s,t),connectivity	HRCI
		Abundances	1	0.002@1	10	Ion fractions	XMS
CME Energetics	Behind-/Off-Limb Flares	Heating	5	0.01@6	5	$T_e(s),n_e(s)$	XMS
		Dynamics/Topology	0.2	–	1	v(s,t), I(s,t)	HRCI
Solar Flares	Flaring Active Regions	Heating	1	0.002@6	0.1	$T_e(s),n_e(s)$	XMS
		Flows	1	0.002@6	10	$v(T_e,s,t)$	XMS,HRCI
		Particle Acceleration	10	1@15-30	~0.001	$dn_e(s)/dE$	FOXSI (Pri)
			1	0.002@6	1	$dn_e(s)/dE$	XMS (Sec)
		Dynamics/Topology	0.2	–	1	v(s,t),I(s,t),connectivity	HRCI
Solar Wind Acceleration	Off-Limb Coronal Holes	Heating	1	0.002@1	30	$T_e(s),n_e(s)$	XMS
		Flows	1	0.002@1	30	$v(T_e,s,t)$	XMS,HRCI
		Dynamics	0.2	–	1	v(s,t), I(s,t)	HRCI
		Abundances	1	0.002@1	10	Ion fractions	XMS

Notes: $T_e(s)$ & $n_e(s)$ are electron temperature & density as a function of position, while $v(T_e,s,t)$ is the velocity as a function of electron temperature and/or position & time. $dn_e(s)/dE$ is the electron energy distribution; this can be measured with FOXSI directly, or indirectly with spectra from XMS. I(s,t) is the EUV intensity.

This core mission can be extended in a number of ways. The most natural addition is a hard X-ray imaging spectrometer (FOXSI; see Christe et al. 2009), whose focusing optic allows for more detailed studies of the non-thermal electron population. The energetic contribution of non-thermal particles in reconnection events tightly constrains models of the reconnection process. The combination of the XMS (which follows the influence of the non-thermal electrons on the ion populations and indirectly derives their distribution) with FOXSI (which directly images the non-thermal electrons and thus their distribution functions) provides a unique capability for understanding coronal heating, electron acceleration, and CME acceleration. This three instrument configuration (XMS, HRCI, FOXSI) is termed SSE+. An alternative SSE+ implementation would replace FOXSI with an EUV spectrometer (such as FACTS; Korendyke et al. 2010) which provides several exciting capabilities: spectroscopic analysis of transition regions and low coronal lines, better velocity resolution (~5 km/s). These capabilities would require rastering a slit across the FOV and so sacrifices the time resolution needed to follow rapid events.

Significant science contributions from SSE range from understanding basic plasma and radiative processes in the corona to providing fundamental measurements of the onset of space weather events. Working together, the instruments comprising SSE/SSE+ will allow us to investigate the propagation of flare and CME disturbances throughout the solar atmosphere, and pinpoint the location and height of magnetic reconnection, energy deposition, and solar wind acceleration.

3. IMPLEMENTATION

The core SSE mission is conceived as a small strategic mission (i.e., with a cost in the range of FY10 $250M - $500M) that will be placed into a geosynchronous orbit using either an Atlas-V or a Falcon 9. The LV selection is driven by the weight of the calorimeter and requirement for a geosynchronous orbit (set after trade studies that were driven by data rates). As a result, there is significant excess launch capacity that can be exploited by augmenting the instrument suite and moving the mission to the low-cost end of the medium-sized category (SSE+). We describe both implementations below.

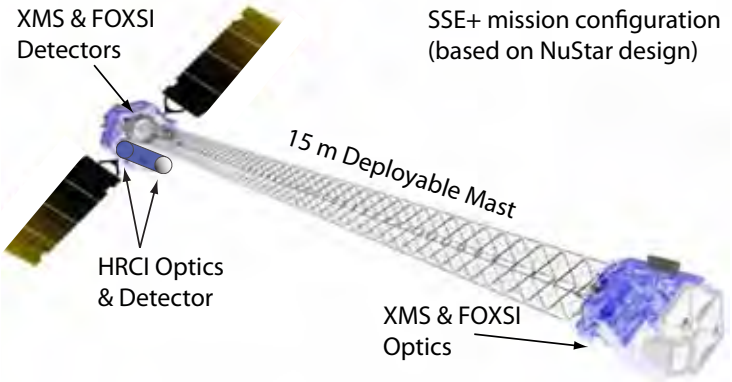


Figure 2 – Schematic of the SSE+ mission; core SSE omits FOXSI.

3.1 SSE Instrument Module

The SSE payload (see cartoon in Fig. 2) consists of (1) a grazing incidence (GI) Flight Mirror Assembly (FMA) mounted on the end of the extendable optical bench (EOB) module; (2) the XMS detector mounted at the focal plane; and (3) the HRCI that consists of a co-aligned self-contained AIA-like telescope.

The FMA design provides effective area of ~2 cm^2 at 1.25 keV, and ~1.5 cm^2 at 6 keV. The XMS (Bandler et al. 2010a) provides high spectral resolution (~2 eV over the 0.2-10 keV bandpass), non-dispersive imaging spectroscopy over the central array with 1 arcsec resolution and a field of view of 36×36 sq. arcsec. An extended array expands the FOV to 3×3 sq. arcmin over the same bandpass with slightly lower energy resolution (≤5 eV). Since the devices are photon counting, each X-ray photon arrival is time-tagged with microsecond precision. Count rates up to 1000 cps can be accommodated with a slight degradation in spectral resolution. The dynamic range of the XMS is increased to over 100,000 with a filter wheel such as used on the Hinode XRT. A adiabatic demagnetization refrigerator (ADR) and a mechanical cryocooler provide cooling to 50 mK without expendable cryogen – so there is no intrinsic limitation to the mission lifetime.

Table 2 – Resource Summary for SSE & SSE+

Payload	Mass	Power	Data Rate
	(kg, CBE)	(W, CBE)	(mbps)
XMS	300	650	22.8
HRCI	50	50	22.4
S/C (incl. EOB)	350	150	1
Total CBE	700	850	46.2
Reserves (30%)	210	255	13.9
SSE Mission Totals	910	1105	60
S/C capability		1500	130
LV throw to GEO	2000		
Margins	**55%**	**26%**	**54%**
FOXSI	100	60	1
SSE+ CBE	800	910	47.2
SSE+ Reserves	240	273	14.2
SSE+ Mission Totals	1040	1183	61.4
SSE+ Margins	**48%**	**21%**	**53%**

The HRCI is a higher-resolution version of an AIA telescope consisting of a 0.4m diameter Ritchey-Crétien optic that – similar to AIA – is coated in segments to allow for multiple EUV channels (nominal wavelengths are: 194Å and 131Å). As with TRACE and AIA, a guide telescope provides the image stabilization signal to a steering secondary optic. The detector for HRCI is a 4k×4k (8k×8k goal to increase the FOV) CCD. The most significant change from AIA is the plate scale (0.1 arcsec instead of 0.6 arcsec pixels). The expected cadence of the HRCI is ~1s.

3.2 Extendable Optical Bench (EOB) Module

The EOB is the portion of the metering structure which is extended on orbit. It consists of an ADAM mast, almost identical to that on the NuSTAR SMEX mission. High precision deployment

accuracy/repeatability was proven with the 60 m ADAM used in space on the NASA's Shuttle Radar Topography Mission, allowing the XMS and FOXSI to have very modest focus mechanisms. A simple metrology system (RHESSI or NuSTAR heritage) will track and correct motions for the two instruments that use the EOB – FOXSI and XMS – to enable accurate attitude reconstruction.

3.3 SSE+ Instrument Module

The addition of a hard X-ray imager is included in the SSE+ payload; the FOXSI instrument is used as a specific example. FOXSI extends the SSE+ mission response to ≥60 keV with 8 arcsec FWHM using nested grazing incidence optics and depth-graded multilayers to provide ~180 cm^2 of effective area at 40 keV; the FOXSI optic has the same focal length as the XMS optic, and is mounted to the same EOB. The FOXSI instrument uses 256×256 CdTe or CdZnTe detectors that provide 500 micron (6 arcsec) pixels, an energy resolution of 1 keV from 5-60 keV, and a FOV of 10×10 sq. arcmin (limited by optics vignetting). Background suppression is achieved with an active anticoincidence shield surrounding five sides of the imager, with heritage from NuSTAR and Astro-H.

3.4 Spacecraft Module

SAO – together with GSFC and several commercial spacecraft providers – developed preliminary concepts that concluded that the spacecraft could be built with technologies that are fully mature today – the SSE/SSE+ concept will therefore not push spacecraft capabilities. All subsystems utilize established hardware with substantial flight heritage, including the Ka-band communications system developed for SDO. Most components are "off-the-shelf."

3.5 Mission and Science Operations

An initial concept has been developed that describes the mission operations approach and the supporting ground data systems. Key aspects are (1) flight and science operations will be conducted from a joint SSE Science and Operations Center, staffed 5x8; (2) all instruments observe simultaneously; (3) there are daily command uploads (similar to AIA or XRT operations); (4) adequate on-board storage and downlinks avoid the need for on-board flare triggers that have proven difficult to implement. The instrument FOVs are sufficiently large that neither external flare triggers nor repointing will be required (see Fig. 1).

With the data volumes and downlink capabilities listed in Table 2 SSE (or SSE+) can observe continuously 24 hours/day. With the rate listed XMS will be downloading photons at an average rate of 300 photons/pixel/s and HRCI would be taking full 4k×4k images with a cadence of ~1s. The daily data volume is about 40% of SDO's. A typical day of SSE data would be (1) a tracked solar active region with 1s spectra (Fig. 1) at every XMS pixel and 131Å and 193Å images every ~2s, and (2) a full scan of the solar disk at full spectral resolution of the XMS and (3) a full scan of the solar disk at 131Å and 193Å at HRCI resolution, thus creating a rich set of data products that can be used to address a wide array of science problems. These full sun scans will take less than an hour per day.

A wide range of quicklook data products can be produced: 1k×1k (binned) HRCI movies, line-specific (e.g. Fe XVII) images from XMS, light curves in different spectral lines or passbands at each XMS pixel, ion-dependent velocity maps at each XMS pixel, and (with SSE+) hard X-ray light curves from FOXSI.

4. MISSION COST AND SCHEDULE

Best-estimate costs for SSE were assembled using various approaches for the different instruments and configurations, including parametric cost modeling, analogy, grass-roots and – given the level of mission maturity – educated estimates. Cost estimates

Table 3: Instrument Costs and Methodologies

Instrument	Methodology	Est. Cost
XMS	Bottom-up based on detailed WBS and MEL (IXO, Astro-H); analogy (Astro-E, Astro-E2); PRICE-H models (IXO project); PRICE-H (GSFC IDL study)	$100M*
HRCI	Analogy (AIA telescopes on SDO; IRIS telescope)	$30M
FACTS	parametric; analogy (EIS)	$85M
FOXSI	Bottom-up; analogy (NuSTAR)	$40M

**Does not include $8M for optical bench (carried with S/C in mission costs).*

Table 4: Costs by WBS

		Estimated Cost ($M, FY10)		
WBS	**Description**	**SSE (Small)**	**SSE+ (Medium)**	**RAM (Large)**
1,2,3*	PM/SE/SMA	33	42	67
4	Science	13	20	26
5	Payload			
	XMS	100	100	100
	HRCI	30	30	75**
	FOXSI	–	40	40
	FACTS	–	–	85
6	Observatory	93	93	93
7	Mission Ops	15	17	19
8	Launch Serv.	75	75	75
9	GDS	5	6	7
10*	System I&T	9	12	20
11	EPO	3	4	5
Reserves (30% of total)§		89	106	158
Total		**465**	**544**	**769**

** Wrap factors for WBS1,2,3&10 scale w/complexity*
*** Ultra-high imager rather the standard HRCI*
§ Excludes LV (WBS7) and GO program (in WBS4).

have been generated for every element within the mission Work Breakdown Structure (WBS). For the XMS significant cost credibility exists since it draws heavily on the Astro-H and IXO calorimeter approaches which have been extremely well studied. The EOB and its metrology system are both costed with the S/C. Table 3 summarizes the methodologies used to cost the SSE instruments. These do not include Phase E related costs (i.e., ground system development, operations, etc.) nor science except what is required to design and test the instruments; these other efforts are carried in the appropriate WBS element (see Table 4).

Table 4 summarizes the overall mission cost estimates in the scalable configurations, including Phase E costs. Note that a first order "complexity factor" was accounted for by increasing the wrap percentages for WBS1,2,3 and WBS 10 as the number of instruments increases, and that WBS7 and 9 also increase as the mission scope increases. For a payload complement for SSE+ we assumed FOXSI ($40M) instead of FACTS ($85M), though either spectrometer could be flown and a science-based trade study should be done.

4.1 SSE schedule

An overall SSE mission schedule (in its small strategic mission configuration) was generated to determine an NET launch date, and is provided in Fig. 3, assuming a FY13 start. The critical path (which flows through the XMS detector system and then to Observatory I&T) is highlighted on the schedule. Nominal start and delivery dates for all major elements are shown. The schedule supports a September 2018 launch readiness review, with 62 months for development through on-orbit checkout (Phases B, C, and D). This includes a total of 6 months of schedule slack on the critical path which slightly exceeds the NASA guidelines.

From the science requirements perspective, there is no strong driver on SSE/SSE+ relative to the solar cycle. The science objectives of the mission are studies of the quiet and active solar corona, as well as X-ray flares. Regardless of the launch date, the nominal three-year lifetime ensures that an adequate portion of the solar cycle containing numerous flares will be observed.

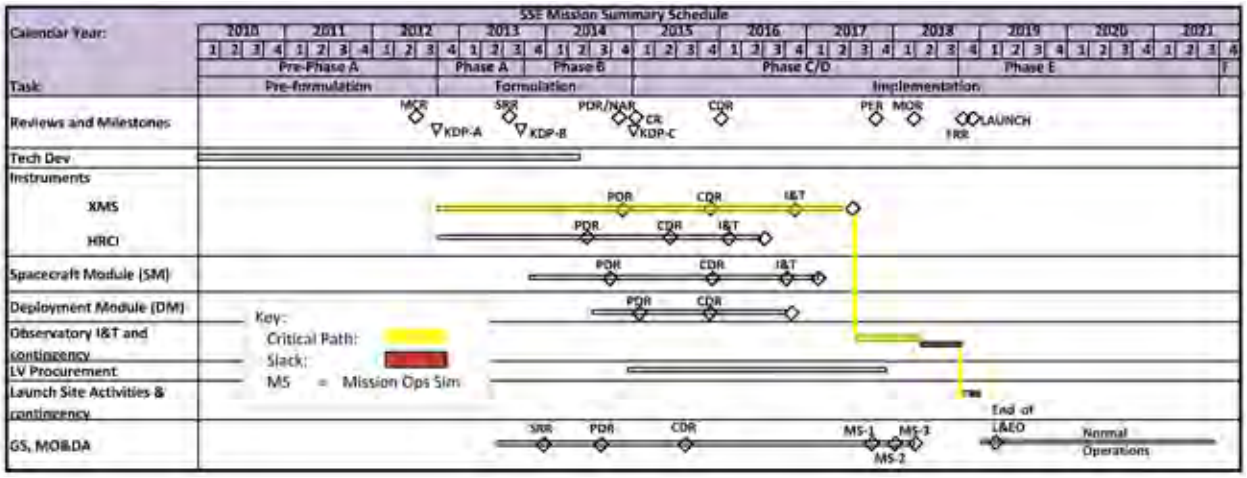
The schedule reflects the ability to capitalize on the modular nature of SSE. The three main observatory modules (Instrument Module, Spacecraft Module, and EOB Module) will be developed and qualified in parallel, and then delivered for final observatory Integration and Test (I&T). Time has been allocated as appropriate for each of the processes for solicitation, selection, and contract awards. Because the critical path flows through the XMS, and the other instruments being considered for SSE+ and RAM have shorter development times, neither the NET launch date nor the critical path should be heavily affected by adding these elements, though some additional time (~1 month) during Observatory I&T may be required due to the increased complexity.

The modularity of SSE/SSE+ also enables potential cost-sharing with ESA, ESA member states, or JAXA with clearly delineated interfaces, though our cost estimates reflect a NASA-only approach.

5. DECADAL EVALUATION CRITERIA

Identified as a high priority or requirement in previous studies or roadmaps: The core SSE science has been highly recommended by the 2006 & 2009 Heliophysics roadmaps. SSE is an evolution of the

Figure 3 – Schedule for SSE mission development with nominal launch in late 2018



RAM mission (Bookbinder et al. 2003; Golub et al. 2010 White Paper) concept that has been studied by SAO since 2000; RAM was recommended in the 2006 Heliophysics Roadmap.

Makes a significant contribution to more than one of the Panel themes: SSE is primarily a basic science mission. The rich SSE mission data products (images, spectra, and timing data) will motivate and challenge researchers in theory, modeling and laboratory plasma physics and in atomic physics.

Contributes to important scientific questions facing solar and space physics today: The fundamental question in space physics is the nature of instabilities in magnetized plasmas. The storage and release of energy in the plasma is controlled by the onset of the instabilities. The SSE X-ray spectra give the definitive measurement of the thermal state of the plasma as a function of space and time, while the SSE HRCI shows the magnetic configuration that confines, heats and moves the plasma.

Complements other observational systems or programs available: SSE/SSE+ is highly complementary to other planned and conceived space missions with its balanced emphasis between coronal spectroscopy and imaging. When SSE is operational several significant ground-based facilities will be online. Foremost is the Advance Technology Solar Telescope (ATST), whose 4m telescope will primarily study plasma processes in the photosphere and chromosphere. The magnetic flux that creates and defines the solar corona necessarily passes through these high beta (and beta~1) regions. ATST and SSE are complementary in their spectroscopic capabilities. The Frequency Agile Solar Radio Telescope (FASR; Gary et al. 2010) directly images the evolution of coronal electrons in flares and coronal mass ejections, while SSE will measure the properties of the ionic plasma (see Fig. 1), providing a natural complementarity for shock and flare studies. COSMO is an innovative program to measure the magnetic field directly in the corona above the solar limb, used to quantifying the energetics of CMEs.

Has an appropriate degree of readiness (technical, resources, people): The only element of the SSE that requires development is the XMS (see §3 of Bandler et al. 2010b). The technology development plan for the XMS has it achieving TRL5 by 2013, well ahead of an SSE new start.

Fits with other national and international plans and activities: The SSE mission will provide the observations necessary to address fundamental science questions outlined by the 2009 Heliophysics Roadmap. SSE will provide complementary science to the heliospheric and solar missions Solar Probe Plus and (potentially) Solar Orbiter (ESA) and Solar-C (JAXA). SSE will give the most comprehensive description of the coronal plasmas' thermal state along field lines that pass near Solar Probe Plus. The knowledge of the state of the plasma at the sun is critical for interpreting the evolution of the plasma as it enters the heliosphere. SSE and Solar Orbiter will simultaneously observe the corona from different view points. Solar-C (Plan B) is a high resolution mission focused on the chromosphere/corona interface. SSE and SSE+ provide complementary coronal observations to Solar-C extending the full complement of spectroscopy from the UV - chromosphere through the soft and hard X-ray corona.

6. REFERENCES

Note: White papers submitted to the heliophysics decadal survey "A Decadal Strategy for Solar and Space Physics" - Solar and Heliospheric Physics panel identified in italics.

[1]Bandler, S. et al. 2010a SPIE 732, 77328
[2] *Bandler, S. et al. 2010b*
[3] Bookbinder, J. et al. 2003, SPIE 4853, 436
[4] *Christe, S. et al. 2010*
[5] Christe, S. et al. 2009, SPIE, 7437E, 4
[6] *Doschek, G. 2010*
[7] *Gary, D. et al. 2010*
[8] *Golub, L. et al. 2010*
[9] Hudson, H., 1991, Sol.Phys., 133, 357
[10] Kashyap et al. 2002, ApJ, 580, 1118
[11] *Korendyke, C. et al. 2010*
[12] *Laming, J. M. et al. 2010*
[13] *Lin, R. et al. 2010*
[14] *Smith, D. et al. 2010*
[15] *Osten, R. et al. 2010*

7. ACRONYMS

ADAM – Able Deployable Articulated Mast
AIA – Atmospheric Imaging Assembly
ADR – Adiabatic Demagnetization Refrigerator
CBE – Current Best Estimate
CCD – Charge Coupled Device
EOB – Extendable Optical Bench
ESA – European Space Agency
EUV– Extreme UltraViolet
FACTS – Fine scale Advanced Coronal and Transition Region Spectrograph
FMA – Flight Mirror Assembly
FOV – Field of View
FOXSI – Focusing Optics X-ray Solar Imager
GI – Grazing Incidence
HRCI – High Resolution Coronal Imager
IXO – International X-ray Observatory
NET – No Earlier Than
NuSTAR – Nuclear Spectroscopic Telescope Array
RAM – Reconnection And Microscale
S/C – SpaceCraft
SMEX – Small Explorer
SR&T – Supporting Research & Technology
SSE – Solar Spectroscopy Explorer
SSE+ – SSE with XMS, HRCI, and FOXSI
SSOC – SSE Science and Operations Center
TRACE – Transition Region and Coronal Explorer
TRL – Technology Readiness Level
WBS – Work Breakdown Structure
XMS – X-ray Microcalorimeter Spectrometer

A Science White Paper in response to the 2012 Heliophysics Decadal Survey "Solar & Heliospheric Physics" and "Solar Wind-Magnetospheric Interactions"

An Experimental Plasma Dynamo Program for Investigations of Fundamental Processes in Heliophysics

Benjamin Brown (U Wisc & CMSO), Cary Forest (U Wisc & CMSO),
Mark Nornberg (U Wisc & CMSO), Ellen Zweibel (U Wisc & CMSO),
Fausto Cattaneo (U Chicago & CMSO) and
Steven Cowley (Imperial College & CCFE, England)

Plasma experiments in laboratory settings offer the opportunity to address fundamental aspects of the solar dynamo and magnetism in the solar atmosphere. Experiments are currently under construction that can investigate the self-generation of magnetic fields and related processes in large, weakly magnetized, fast flowing, and hot (conducting) plasmas. These and future experiments will probe questions that are of crucial importance to heliophysics in the solar interior, atmosphere and wind. Uniquely, laboratory plasma experiments coupled with theoretical explorations can serve to calibrate the simulation codes which are being used to understand the solar dynamo, magnetic reconnection and flares in the solar atmosphere, the nature of CMEs, and the interactions between planetary magnetospheres and the solar wind. Laboratory plasma experiments are likely to contribute new understanding complementary to the traditional observational and modeling approach normally used by space physicists.

We argue here that ground-based laboratory experiments have direct connections to NASA based missions and NSF programs, and that a small investment in laboratory heliophysics may have a high payoff. We will use the Madison Plasma Dynamo Experiment (MPDX)[1] as an example, but advocate here for broad involvement in community-scale plasma experiments.

Fundamental plasma processes in heliophysics and connection to NASA missions and NSF programs The 22-year solar cycle stands out as one of the most remarkable and enigmatic examples of magnetic self-organization in nature. The Sun's cycles of magnetic activity profoundly affect our modern technological society. Unsurprisingly, solar magnetism is a fundamental focus of current and future NASA missions.

Magnetic fields that emerge at the solar surface as sunspots are built by dynamo action in the solar convection zone. In most solar dynamo models, organized fields are built in the tachocline, an interface layer deep in the Sun at the bottom of the solar convection zone. These magnetic structures then become buoyantly unstable and rise through the turbulent convection to emerge at the surface. After emergence, magnetic reconnection in the solar atmosphere can lead to flares and coronal mass ejections (CMEs) which have substantial impacts on the heliosphere and on Earth's

[1]This work builds upon excitement in recent years of using liquid metals to study dynamos and will extend these studies to more astrophysically relevant parameters. Use of a plasma for such experiments allows the magnetic Reynolds number (the dimensionless parameter governing self-excitation of magnetic fields) to be approximately a factor of 10 larger than in liquid metal experiments. These experiments will be the first to investigate self-excited dynamos in a plasma, the state of matter that makes up most naturally occurring astrophysical dynamos.

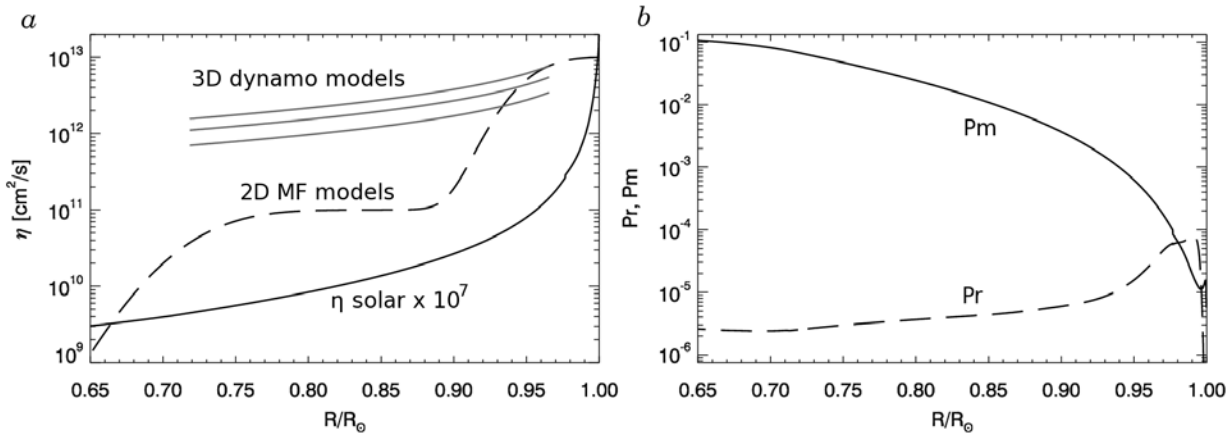


Figure 1: (a) Radial profiles of magnetic diffusivity η used in various dynamo models. Shown in grey lines are three representative profiles from 3D convective solar dynamo simulations using the anelastic spherical harmonic (ASH) code [e.g., 1–3]. Also shown is the double-step profiles used in 2D mean-field Babcock-Leighton models that have been used to predict the current solar cycle [4]. The molecular diffusivity for a hydrogen plasma at solar conditions is also shown, multiplied by 10^7 for display purposes. (b) The Prandtl number $\mathrm{Pr} = \nu/\kappa$ (dashed) and magnetic Prandtl number $\mathrm{Pm} = \nu/\eta$ (solid) for a hydrogen plasma at solar conditions. Simulations use values of order unity.

magnetosphere in particular. These processes of magnetic flux creation, emergence, and reconnection form many of the driving questions behind the current Solar Dynamics Observatory (SDO), Hinode, the Stereo mission, the very successful Solar and Heliospheric Observatory (SOHO), and the upcoming ground-based NSF Advanced Technology Solar Telescope (ATST). Understanding the origins of the heliospheric magnetic field and the evolution of CMEs in the heliosphere is fundamental to the future Solar Sentinels mission and plays important roles in the future Solar Probe and Interface Region Imaging Spectrograph (IRIS) mission as well. Indeed, the evolution of solar magnetism forms the driving questions behind the Living With a Star (LWS) program at NASA: namely, *how and why does the Sun vary, how does the Earth respond, and what are the impacts to humanity?*

As a community, we have made significant progress in understanding solar magnetism through a combination of theoretical treatments and observational techniques, but many key components of the solar dynamo remain poorly understood. In particular, the processes behind magnetic buoyancy and the generation of magnetic field through turbulent correlations occur deep in the solar interior and are difficult to constrain with either helioseismic observations or numerical simulations. Magnetic reconnection in the solar atmosphere is likewise poorly understood at present. Our understanding of interactions between the Sun and the Earth have grown tremendously and are greatly facilitated by in situ observations of the solar wind in the near-Earth environment by missions like the Advanced Composition Explorer (ACE) and the Magnetospheric Multi-scale Mission (MMM), but many details of interactions between coherent magnetized plasma structures like CMEs with the Earth's magnetosphere remain unclear.

A fundamental limitation facing theoretical explorations of heliophysically relevant plasma processes is the vast separation in parameter space between the natural phenomena and the capabilities of even the largest super computers. As one example, cutting edge global-scale solar dynamo simulations run as part of the NASA High-End Computing (HEC) program[2] and as part of the NSF

[2]using for example the Pleiades and Columbia supercomputers

Partnerships in Advanced Computing Infrastructure Teragrid program[3] typically employ turbulent diffusivities which are more than ten orders of magnitude larger than the molecular diffusivities of the solar plasma (Figure 1*a*). This also applies to the mean-field models which have been used in attempts to predict amplitude and timing of the current solar cycle [e.g., 4, 5]. Yet these simulations still require substantial computational resources and represent a significant investment by NASA HEC and the NSF Teragrid.

It is deeply impractical to consider simulating global-scale dynamo action in the Sun using molecular diffusivities and doing so would likely require a century of further growth in computational resources. Instead, it is vital to formulate better estimates and models of the turbulent transport processes that occur in a plasma under solar conditions. As an example, even the ratio between how quickly magnetic fields are transported and mixed by turbulence compared to the transport of momentum by the same is poorly constrained. Molecular values of this magnetic Prandtl number Pm are of order 10^{-1}–10^{-5} (Fig. 1*b*); turbulent values are thought to be of order unity, but this has not been adequately explored for real turbulent transport in a stratified plasma. These problems are not unique to the solar interior, but are true for simulations of photospheric convection that couple to the chromosphere and corona [e.g., 6], and simulations that capture the interactions between the solar wind, CMEs and planet magnetospheres throughout the heliosphere and to the boundary with interstellar space itself.

The growth of computational resources has made it feasible to directly simulate the plasma regime explored by experiments like MPDX. A program of laboratory experimentation that is tightly coupled with significant numerical simulation efforts will lead to new insights into key plasma processes. In particular, these laboratory plasma experiments will likely achieve self-sustaining dynamo action, will explore the dynamics of magnetic buoyancy in a stratified plasma atmosphere, and will explore the interactions between an Earth-like magnetosphere and magnetotail with a solar-wind like plasma flow. In addition to revealing the fundamental underlying plasma processes at work in these systems, these experiments can anchor the computationally intensive simulations which are used to model the solar dynamo, the buoyant rise and emergence of magnetic structures in the solar interior, and interactions of the solar wind and CMEs with the Earth's magnetosphere.

The importance of simulations coupled to experiments Just as we cannot directly simulate the turbulent conditions of solar convection, we cannot construct laboratory experiments that capture the vast spectrum of scales present in the Sun. Some important physical ingredients acting in solar convection are very difficult to reproduce in experiments: in particular, it is very difficult to produce conditions that mimic the radial self-gravity of systems like the Sun, and experiments that do must generally be conducted at great cost in low-g environments [e.g., 7].

Laboratory experiments provide great opportunities for exploring key physical processes occurring in plasmas under solar conditions, but they cannot capture the integrated global-scale picture. Global-scale couplings between convection, magnetism and rotation are likely crucial to fully understanding the solar cycle as a whole and may be important for understanding eruptive flares and CMEs as well. At present, simulations are the only option for exploring the coupled global-scale system. Simulations can offer insight into processes that are difficult or impossible to directly measure in the laboratory (e.g., the full 3-D velocity field and magnetic field at all points in a volume at all instances in time) and can suggest profitable avenues for additional experimental exploration. But this is not a one-way process: the laboratory experiments validate the codes and inform those

[3]using supercomputers at Pittsburg Supercomputing Center (PSC), San Diego Supercomputer Center (SDSC), Texas Advanced Computing Center (TACC), National Institute for Computational Science (NICS), at the National Center for Supercomputing Applications (NCSA), and others

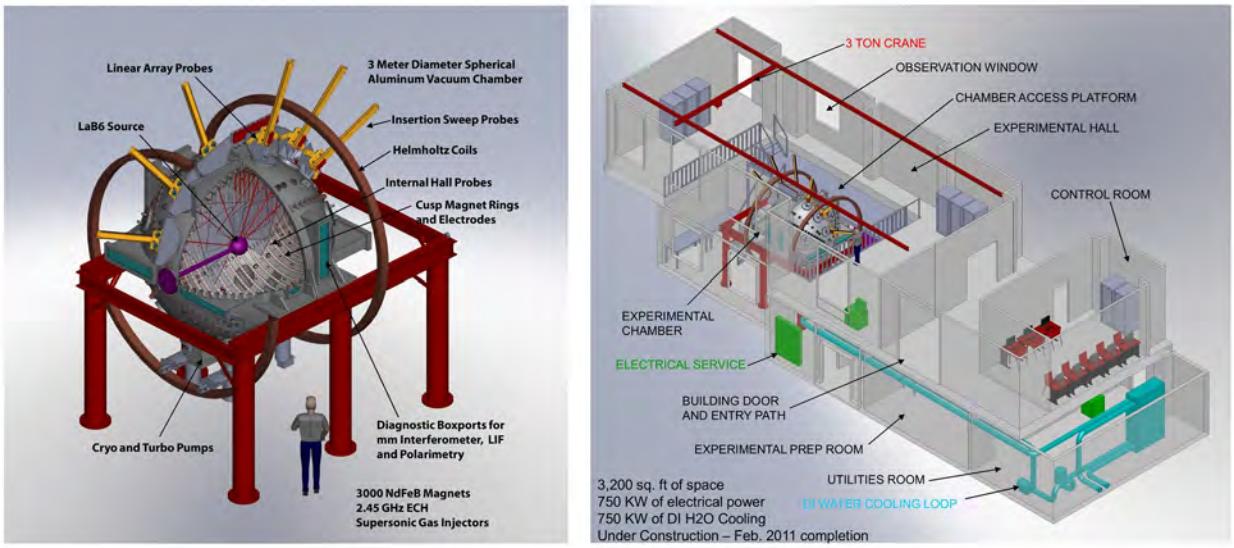


Figure 2: The Madison Plasma Dynamo Experiment (MPDX), with human figure to scale (left). Plasma Dynamo Laboratory Overview (right). The drawing shows the new Plasma Dynamo Facility under construction at the UW Madison. The main experimental hall is two stories high, with approximately 2000 ft^2 of floor space and houses the 3 m diameter Plasma Dynamo Experiment.

simulations in how to correctly capture processes that currently elude simple theoretical explanation, including turbulent transport, magnetic buoyancy instabilities, and the non-linear couplings that arise between flows and magnetism.

Ultimately, theoretical explorations, numerical simulations and laboratory experiments must work together in a tightly coupled fashion to achieve optimal results. But this is feasible and can yield large and rapid payoffs. In the MPDX experiment for example, this work has begun already for modeling of the dynamo and magneto-rotational instability scenarios using the NIMROD code, which solves the MHD equations with a number of two-fluid extensions. These simulations have indicated that sustained dynamo action is possible and in which parameter regimes this is most achievable. New simulations are exploring the possibility of buoyancy instabilities, including magnetic buoyancy, and suggest that these may be experimentally achievable. Direct comparisons between the numerical simulations and pathfinder prototype experiments are completing the code *validation* loop even before construction finishes on the main experiment (MPDX). Theoretical explorations of the simulations and the experiment are in turn providing the ultimate link to the plasma astrophysics.

Tightly linked programs of laboratory experimentation and numerical simulation, backed by significant theoretical analysis to understand both systems, offer opportunities for rapid advancement in our understandings of solar plasma physics.

Status of MPDX and future plasma experiments Compared to space missions, ground based laboratory experiments can be quite affordable and can be built on rapid timescales. As an example, the MPDX experiment has been funded for construction through the NSF Major Research Instrumentation program with at total device cost of $2.5M (part of which was $750k in cost-sharing from the UW). In 2009 a Plasma Dynamo Kick-Off Meeting, sponsored by the NSF Center

for Magnetic Self-Organization,[4] was held to help formulate the scientific program and to provide feedback on the design of the experiment. Funding began in 2009, and the 3 year project is on schedule: 2010 was spent on design and lab remodeling, in 2011 the machine will be assembled, and in 2012 the experiment should be ready for initial operations. An illustration of the experimental apparatus and the Plasma Dynamo Laboratory built around it is shown in Figure 2.

There are significant barriers to running such programs on a community scale. Using MPDX as an example again, after construction there is not yet a program in place to operate this experiment. This experiment is larger than a typical single-investigator NSF experiment and might be considered a *medium-scale* basic plasma or laboratory plasma astrophysics experiment as called for in the recent Plasma 2010 report from the National Research Council. The operations of such an experiment and the diagnostic requirements are similar in scale to university-scale experiments in fusion research, yet there is no mechanism for funding the operation of a community-based facility such as this. A major challenge is adequately funding the operations of such experiments and also the significant theoretical explorations and numerical simulation efforts which are vital to maximizing the scientific returns of such facilities.

Heliophysically relevant plasma experiments Future plasma experiments could directly address many plasma processes that are directly relevant to heliophysics and NASA. Most experimental plasma facilities at present are focused on other science, with substantial investments in exploring basic plasma processes and the application to fusion energy generation. Plasma experiments can now directly address several key components of the solar dynamo and of eruptive events in the solar atmosphere.

A variety of laboratory experiments are feasible in the near future that are clearly heliophysically relevant:

1. large-scale dynamo experiments (generation of the large scale solar magnetic field),
2. small-scale turbulent dynamo experiments (small scale structure of the solar dynamo and surface magnetism),
3. buoyancy driven stratified convection (solar convection zone and surface),
4. magnetic buoyancy instabilities in a stratified atmosphere (sunspot emergence and solar tachocline physics),
5. flow driven plasma turbulence in near-equipartition (solar wind),
6. flow driven plasma turbulence impacting a model magnetosphere (space weather, sun-earth connection),
7. magnetic field line stretching with potentially explosive, flow-driven reconnection (magnetotail, solar flares),
8. plasma instabilities at high β and low collisionality (solar wind spectrum).

A program of heliophysically relevant laboratory plasma experiments can provide important insight into these problems. Such a program will surely suggest additional unexpected avenues for research, as expertise is gained in experimental techniques, and in the theoretical and numerical explorations of such experiments.

[4] [http://cmso.info/html/meetings/wi_dec09.htm]

The proposed plasma experiments transcend the scope of *single user facilities* and instead require substantial *community-based facilities*. Such experiments should be supported at a larger scale than can be undertaken by individual universities and research groups. This support must extend past the mere construction of experimental facilities: instead, significant advances can be achieved by ongoing support for operations, for theoretical explorations, and for tightly coupled numerical simulations. These synergistic efforts are likely to maximize scientific output and may rapidly advance our understanding of plasma processes occurring in the solar interior, atmosphere and wind, whereas isolated efforts must continue to wrestle with intractable limitations. In summary, we advocate that heliophysics could greatly benefit from the support of community-scale, heliophysically relevant plasma experiments.

References

[1] Brun A S, Miesch M S and Toomre J 2004 *ApJ* **614** 1073–1098

[2] Brown B P, Miesch M S, Browning M K, Brun A S and Toomre J 2010 *ApJ* submitted

[3] Miesch M S, Brown B P, Browning M K, Brun A S and Toomre J 2010 *Preprint* `1009.6184`

[4] Dikpati M and Gilman P A 2006 *ApJ* **649** 498–514

[5] Charbonneau P 2010 *Living Reviews in Solar Physics* **7** 3:1–91

[6] Nordlund Å, Stein R F and Asplund M 2009 *Living Reviews in Solar Physics* **6** 2:1–116

[7] Hart, J E, Toomre, J, Deane, A E, Hurlburt, N E, Glatzmaier, G A, Fichtl, G H, Leslie, F, Fowlis, W W and Gilman, P A 1986, *Science* **234** 61

Heliospheric Physics Survey White Paper: Intermediate-scale MHD wind tunnel for turbulence and reconnection studies

Submitted: November 5, 2010

Prof. Michael Brown†, Dr. Timothy Gray, Prof. William Matthaeus
Swarthmore College/ University of Delaware
Department of Physics and Astronomy
500 College Avenue
Swarthmore, PA 19081
(office) 610-328-8648
(fax) 610-328-7895
(e-mail) doc@swarthmore.edu

I. PROJECT SUMMARY:

Laboratory plasma experiments are plagued by a lack of separation between the dissipation scale and the macroscopic system scale. Typical table-top plasma experiments ($B = 0.1\ T, T_i = 20\ eV$) have a proton gyro orbit of about 1 cm. As an example, the MHD plasma in the SSX reconnection device has a dissipation scale of $\sim 1\ cm$ (comparable to both the ion gyro-scale and the ion inertial scale) while the macroscopic scale is $10-40\ cm$. Even intermediate laboratory plasmas geared towards fusion (eg the MST RFP, $B = 0.5\ T, T_i = 1\ keV, r = 0.5\ m$) also have proton gyro orbit of about 1 cm so the dissipation scale and the macroscopic system scale are not widely separated. Meanwhile, the solar wind MHD plasma has a dissipation scale of $\sim 1\ km$ with a macroscopic scale approaching 1 AU ($150 \times 10^6\ km$).

The SSX device (Swarthmore Spheromak Experiment) at Swarthmore College features a $L \cong 1\ m$ long, high vacuum chamber in which we generate $n \geq 10^{20}\ m^{-3}, T \geq 20\ eV, B \cong 0.1\ T$ plasmas. Compact plasmas are generated by pulsed magnetized plasma guns typically at opposite ends of the device. The plasma is fully ionized and fully magnetized ($\rho_i < L$). Plasmas are accelerated to high velocity ($\leq 100\ km/s$) by the discharge current in the guns ($\leq 100\ kA$) and injected into a highly evacuated target volume called a flux conserver. We have been operating prototype MHD wind tunnel experiments in SSX, launching high-velocity magnetized plasmas at high velocity down a long evacuated cylinder.

We suggest that an intermediate scale MHD wind tunnel could be built for about $1M. The device would be 1 m diameter and 10 m long. The device would feature ultra-high vacuum quality and a suite of fast magnetic diagnostics. The separation of scales would be at least 2 orders of magnitude, much more than typical small-scale laboratory plasma experiments. Experiments on such a device would illuminate MHD turbulence processes such as observed in the solar wind.

II. TURBULENCE:

Turbulence in MHD fluids is characterized by large Reynolds numbers, both fluid $R_e = Lv/\nu$ and magnetic $R_m = \mu_0 Lv/\eta$ (where ν is a plasma viscosity, and η a resistivity), and activity over a range of spatial and temporal scales. In the solar wind, the best studied

MHD turbulence laboratory [1], fully-developed turbulence is observed over 4 decades in space (km to 1000's of km) and time (seconds to 1000's of seconds). In laboratory plasmas, dynamical scales are more limited [2]. In particular, the SSX reconnection experiment displays dynamical activity over about a decade spatially (from a few cm to just a few 10's of cm) and a few decades temporally (a fraction of a μs to 10's of μs).

A hallmark of turbulence [1] is the spectral transfer of energy in both spatial and temporal frequencies, characterized by magnetic energy spectra $E_B(k)$ and $E_B(f)$ respectively. The typical picture of turbulent dynamics is that energy is introduced into the system at large spatial scales (ie low spatial frequency k) and non-linearities drive the energy to ever-smaller scales where it is ultimately dissipated as heat. This flow in k-space establishes the energy spectrum ($E_B(k)$ for magnetic fluctuations, for example). For very large systems, with a substantial separation between the injection and dissipation scales, a self-similar power-law dependence can develop in the so-called "inertial range".

In 3D MHD systems, turbulence subject to certain constraints can lead to the emergence of large-scale structures [3]. It can happen that quantities "inverse-cascade" to generate self-organized, large-scale structures out of the small-scale turbulence [4, 5].

III. SSX PROTOTYPE EXPERIMENT:

Dynamics in SSX are typically initiated by the merger of two dense, high-magnetic flux spheromaks from either end of the flux conserver generated by magnetized plasma guns. A single high-velocity spheromak can also be injected from either end. Inflow speeds are on the order of 100 km/s. The flux conserver serves as a high-vacuum (10^{-8} $torr$), low magnetic field target chamber bounded by a highly conductive shell of different shapes. Plasma and magnetic field flow into the flux conserver and relax via reconnection and turbulence towards a final state. The final states are determined by flux conserver aspect ratio, the initial helicities, and eigenmodes of $\nabla \times \mathbf{B} = \lambda \mathbf{B}$, but the underlying dynamics are complex. Merging of two counter-helicity spheromaks in SSX initiates magnetic reconnection at the midplane that is fully 3D [6] and generates energetic particles [7, 8], flow [9], and heating [10, 11].

Recently, we have observed the relaxation of merged spheromaks to a helical $m = 1$ state in our prolate flux conserver geometry ($L = 0.6\ m, R = 0.2\ m$) [12, 13]. The final state has

a flat $\lambda = \mu_0 J/B$ profile and is well-fit by the minimum energy solution to $\nabla \times \mathbf{B} = \lambda \mathbf{B}$ (with $\lambda = 17.7\ m^{-1}$, a constant with units of inverse length). This state was driven by the merger of two identical spheromaks with the same magnetic helicity.

We have recently begun turbulent relaxation studies by driving a single magnetized plasma plume down a long MHD wind tunnel. The prototype experiment (see Figure 1) consists of a single source at one end of a 1 m long, 0.16 m diameter cylindrical flux conserving tube. The final, relaxed state is a double-helix geometry depicted in Figure 2. In a

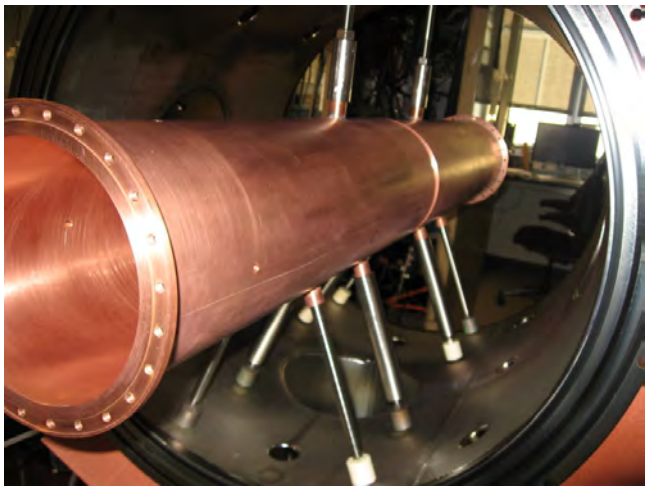


FIG. 1: Photo of the prototype SSX MHD wind tunnel flux conserver. An extension to the MHD wind tunnel is being designed with an exhaust region in the main chamber.

preliminary experiment, we have verified this basic structure with a 20-probe array placed along the axis of the device. In Figure 3, normalized components of the measured magnetic field, $B_x/|B|$ and $B_y/|B|$ are plotted as a function of z. The initial helicity of the injected magnetic field was right-handed and a right-hand twist is clearly evident in the final state. In Figure 4, we display the dynamics of relaxation to the final state by plotting the spatial mode structure as a function of time ($E(k_z, t)$). Note that early in time, there is mode activity at several spatial scales (spanning about a decade) but very rapidly, the activity relaxes into a single mode with $k_z = 1/\ell = 2.2\ m^{-1}$. Within the errors of our coarse prototype probe array, this is consistent with the predicted infinite Taylor double-helix state of $\mathbf{k}R = 1.234$ [14].

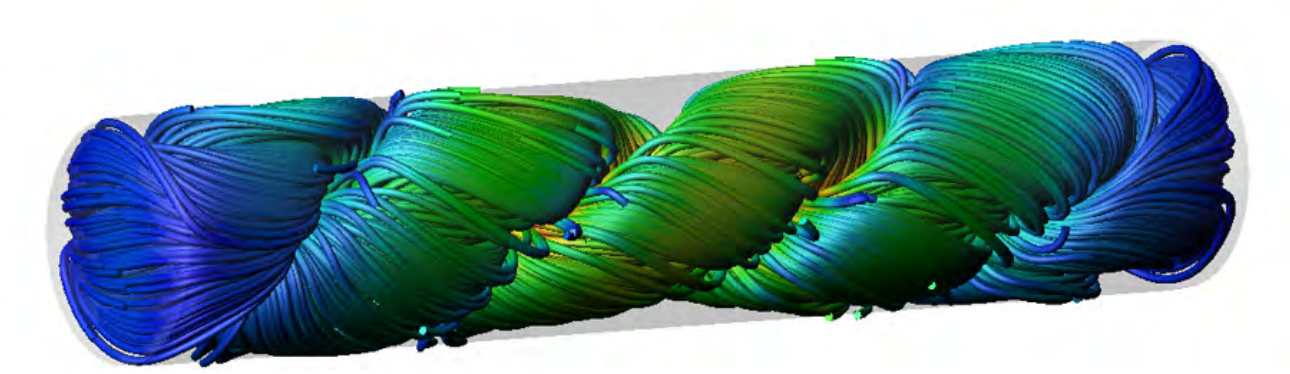


FIG. 2: Rendering of the 3D twisted helical magnetic structure of the lowest energy eigenmode of $\nabla \times \mathbf{B} = \lambda \mathbf{B}$ for the SSX MHD wind tunnel flux conserver (1 m long, 0.16 m diameter) using the VisIt package. Eigenmode structure calculated by G. Marklin of the PSI-Center using PSI-TET.

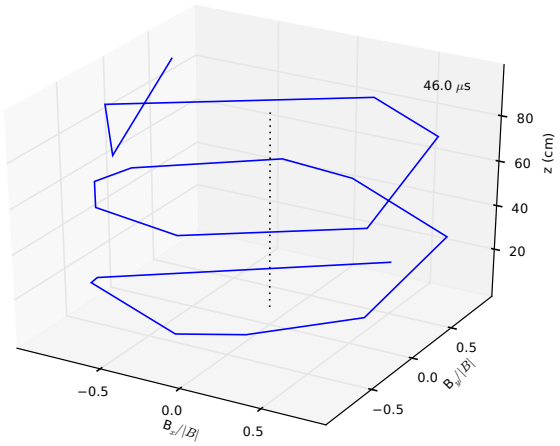


FIG. 3: Preliminary magnetic data from an axial magnetic probe array (20 channels) clearly showing the right-handed helical twist. These data were taken late in time ($t = 46\ \mu s$), after the plasma has fully relaxed. Note that the pitch is about $\ell = 0.45\ m$, so $k_z = 1/\ell \cong 2.2\ m^{-1}$.

IV. PROPOSED INTERMEDIATE SCALE EXPERIMENT:

We suggest that an interesting laboratory plasma facility would be an intermediate scale magnetized plasma wind tunnel with length about 10 meters and diameter 1 meter. Our experience suggests that with large scale fields of 1000 gauss and plasma temperatures of 10's of eV, the dissipation scale would be about 1 cm while the energy macroscale would

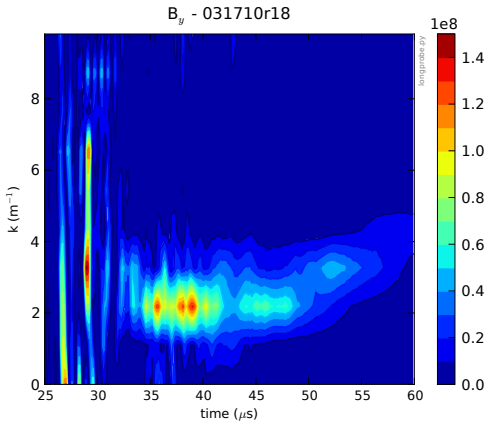


FIG. 4: Time evolution of the relaxation dynamics in k-space, showing the relaxation to the $k_z = 2.2\ m^{-1}$ Taylor double-helix state ($\mathbf{k}R = 1.234$). Note that early in time, there is activity over a decade in spatial scale but late in time, the activity has condensed to one mode.

be about 1 meter, two full decades of scale separation. The scale of the device would be approximately that of the LAPD device at UCLA. A key difference is that the MHD wind tunnel would be pulsed from a high flux, high energy source. The plasma source would be similar to that of the now closed SSPX spheromak device previously operated at LLNL.

We envision that the vacuum chamber would consist of up to 10 48-inch long, 48-inch diameter sections. The plasma-facing material would be highly conducting copper perhaps 0.25 inch thick wall, 1 meter diameter, and 10 meters long. This is 4800 pounds of copper. We have found that it is critical to bake and plasma-clean the flux-conserving boundaries in vacuum. Copper (or a tungsten coating) can adsorb up to 10^{20} H_2 molecules on its surface. A high velocity MHD wind can liberate the adsorbed gas, cooling the plasma and adding mass to the flow.

From our experience with the SSX prototype MHD wind tunnel, we estimate that 5-3000 liter per second cryo pumps are required. The plasma volume will be about 400 times that of our prototype wind tunnel so we estimate 40-500 μF, 10 kV capacitors (25 kJ each, 1 MJ total) would be required. A weak (100 Gauss) axial field would add some flexibility. Initial diagnostics would consist of a 100 magnetic probe array for high frequency fluctuation and correlation studies, a He-Ne interferometer system for density and density fluctuation studies, and ion Doppler spectroscopy for ion temperature studies.

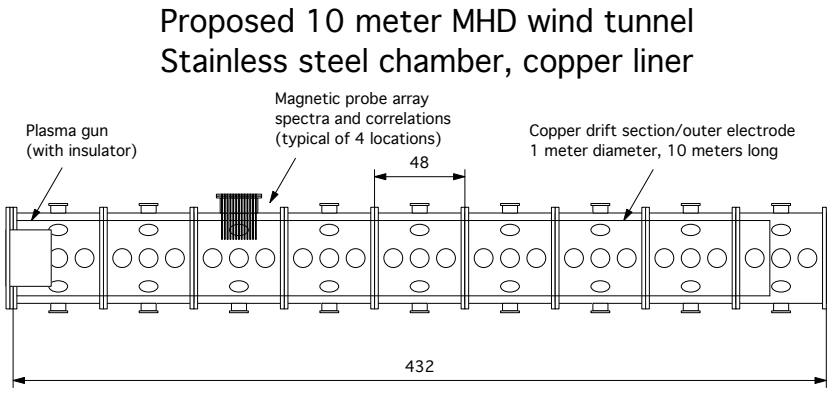


FIG. 5: Preliminary design of 10-meter MHD wind tunnel. The main vacuum chamber features nine 48-inch sections. The 10-meter copper drift tube is placed inside.

number	**description**	**total**
5	3000 liter per second cryo pumps, compressor, gate valve	\$100K
1	1 meter diameter, 10 meter long, 1 cm thick copper flux conserver	\$20K
10	48-inch long, 48-inch diameter vacuum sections @ \$40K	\$400K
40	500 μF, 10 kV capacitors, ignitron switches	\$100K
10	96 channel DTAQ digitizer system (fluctuation spectra)	\$100K
1	magnetized plasma gun (tungsten coated)	\$50K
1	He glow discharge system (power supply, gas flow, pump)	\$50K
1	100 gauss magnet system (axial guide field)	\$100K
	TOTAL	**\$920K**

[1] M. Goldstein, D. Roberts, and W. Matthaeus, Annual Review of Astronomy and Astrophysics **33**, 283 (1995), ISSN 0066-4146.

[2] W. Gekelman and R. Stenzel, Journal of Geophysical Research-Space Physics **89**, 2715 (1984).

[3] M. Brown, Journal of Plasma Physics **57**, 203 (1997), ISSN 0022-3778.

[4] U. Frisch, A. Pouquet, J. Leorat, and A. Mazure, Journal of Fluid Mechanics **68**, 769 (1975), ISSN 0022-1120.

[5] A. Pouquet, U. Frisch, and J. Leorat, Journal of Fluid Mechanics **77**, 321 (1976), ISSN 0022-1120.

[6] C. D. Cothran, M. Landreman, W. H. Matthaeus, and M. R. Brown, Geophys. Res. Lett. **30**, 1213 (2003).

[7] M. R. Brown, C. D. Cothran, M. Landreman, D. Schlossberg, and W. H. Matthaeus, Astrophys. J. **577**, L63 (2002).

[8] M. R. Brown, C. D. Cothran, M. Landreman, D. Schlossberg, W. H. Matthaeus, G. Qin, V. S. Lukin, and T. Gray, Physics of Plasmas **9**, 2077 (2002), URL `http://link.aip.org/link/?PHP/9/2077/1`.

[9] M. R. Brown, C. D. Cothran, and J. Fung, Physics of Plasmas **13**, 056503 (pages 8) (2006), URL `http://link.aip.org/link/?PHP/13/056503/1`.

[10] V. H. Chaplin, M. R. Brown, D. H. Cohen, T. Gray, and C. D. Cothran, Physics of Plasmas **16**, 042505 (pages 11) (2009), URL `http://link.aip.org/link/?PHP/16/042505/1`.

[11] Y. Lin, X. Y. Wang, M. R. Brown, M. J. Schaffer, and C. D. Cothran, Plasma Phys. & Controlled Fusion **50**, 074012 (2008).

[12] C. D. Cothran, M. R. Brown, T. Gray, M. J. Schaffer, and G. Marklin, Phys. Rev. Lett. **103**, 215002 (2009).

[13] C. D. Cothran, M. R. Brown, T. Gray, M. J. Schaffer, G. Marklin, and V. S. Lukin, Physics of Plasmas **17**, 055705 (pages 7) (2010), URL `http://link.aip.org/link/?PHP/17/055705/1`.

[14] J. B. Taylor, Reviews of Modern Physics **58**, 741 (1986), ISSN 0034-6861.

The importance of ground-based observations of the solar corona

J. Burkepile[1], S. Tomczyk[1], P. Nelson[1], A.G. deWijn[1], S. Sewell[1], D. Elmore[2], L. Sutherland[1], R. Summers[1], D. Kolinski[1], L. Sitongia[1]

[1] High Altitude Observatory, NCAR; [2] National Solar Observatory



Summary

The solar corona is a unique physical system that expands to fill interplanetary space providing a direct connection between the Sun and the space environment at Earth. The continual eruption of magnetic flux into the solar atmosphere creates stresses which give rise to various forms of solar activity and result in the global restructuring of the corona and heliosphere over the solar cycle. The highly conducting corona continually evolves as it processes solar magnetic fields into metastable states and eventually ejects them into the heliosphere as Coronal Mass Ejections (CMEs), which are the primary driver of major interplanetary disturbances at Earth. Understanding the processes responsible for the Sun's continuous release of plasma and energy and their impact at Earth and throughout the heliosphere are primary goals of programs including the NSF National Space Weather Program and the NASA Living With a Star. To address these goals, The National Center for Atmospheric Research (NCAR) operates the Mauna Loa Solar Observatory (MLSO) containing a suite of instruments designed to provide the community with observations for studying both short- (minutes-hours) and long-term (years) changes on the Sun from the photosphere, through the chromosphere and into the corona. Among these instruments is the MK4 white light K-coronameter, which records images of the low corona (1.12 to 2.9 solar radii), as illustrated in Figure 1. The MK4 instituted improvements over the earlier MK3 K-coronameter which operated from 1980 to 1999, while utilizing the same telescope, and optical path of the MK3. A new state-of-the-art ground-based COSMO K-coronagraph has recently been designed that will significantly improve the ability to detect and measure coronal structures, such as CMEs, from the ground at a small fraction of the cost of its space-based counterparts. The new instrument will provide extremely high time cadence (15 seconds) and will provide new white light views into the very low corona (down to 1.05 solar radii) where most CMEs form. These observations will provide exciting and novel information of CME formation and propagation and enhance the value of existing and new observatories on the ground (e.g. SOLIS, ATST, COSMO, FASR) and in space (e.g. SDO, Hinode, STEREO, SOHO, Rhessi, and ACE).

The importance of low coronal observations

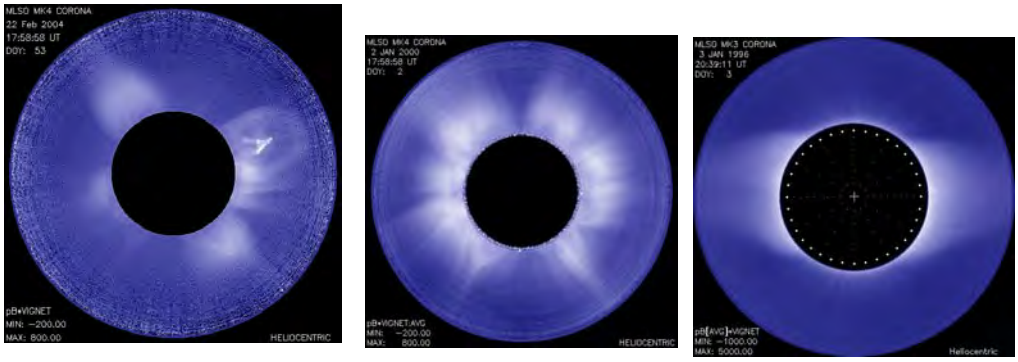


Figure 1. The corona recorded by the Mauna Loa K-coronameter. **Left:** An image of a CME taken on February 22, 2004 showing the classic 3-part structure (loop-caivity-prominence) of CMEs. **Center and Right:** These images show the evolution of the global corona over the 11-year solar cycle. The center image was taken during solar maximum on January 2, 2000 by the MK4 K-coronameter, the right image was acquired during solar minimum on January 3, 1996 by the MK3 K-coronameter.

Coronal activity is driven by the free energy contained in the twisted magnetic fields (the electric current systems) generated as the highly conducting plasma responds to the emergence of magnetic field and plasma from beneath the solar surface and to rotational and convective sub-photospheric motions. Strong flares are frequently associated with large, sudden releases of coronal plasma and magnetic field into the heliosphere, known as CMEs. CMEs are the primary drivers of major geomagnetic storms and may be an essential component of the solar dynamo by removing magnetic flux and helicity from the corona, thus aiding in the polarity reversal of the solar magnetic field every ~11 years (Low 1996). CMEs and other space weather events are driven by changes in the coronal magnetic field. White light observations do not provide direct detection of the magnetic field. Observations of the magnetic field are a critical component to understanding solar activity and are discussed in other papers (COSMO, ATST, FASR). However, heat conduction constrains the million degree coronal plasma to follow magnetic field lines so that the plasma acts as a tracer of the coronal magnetic field morphology. White light coronal observations record the plasma (density) structure and by inference, the magnetic topology of the corona (e.g. Ko et al. 2008) . CMEs can be unambiguously detected and tracked in white light. In fact, most of the basic properties of CMEs (e.g. location, width, speed, mass) have been determined from white light observations. Since the 1970s, space-based coronagraph observations from OSO-7, Skylab, SMM, SolWind, SOHO, and STEREO, have provided numerous discoveries about CMEs and the global structure of the corona. The Mauna Loa K-coronameters have provided valuable and complimentary imagery of the low corona in support of these and other space-based missions since 1980. Mauna Loa observations of the very low corona have contributed important information on CMEs including: 1) detection of intensity dimings (and thus density loss) in the corona due to the passage of a CME (Hansen et al. 1974); 2) determining that the largest CME acceleration occurs below 3 solar radii; 3)

provide measurements of the initial CME acceleration that result in more accurate CME start times (MacQueen et al. 1985), 4) have shown that most CMEs form in the first coronal scale height (Harrison et al. 1990); 5) identified that bright new helmet streamers are the newly closed magnetic fields which were opened during a CME (Hiei et al. 1993). These observations have revealed that the forces responsible for CME production and propagation occur very low in the solar atmosphere. Our understanding of CME formation, their relation to other forms of solar activity and the manner in which they drive geomagnetic responses has dramatically improved. Many discoveries have been achieved through combined observations, of the inner and outer corona, chromosphere and photosphere. These discoveries have reinforced the need for white light observations, particularly in the region of the very low corona.

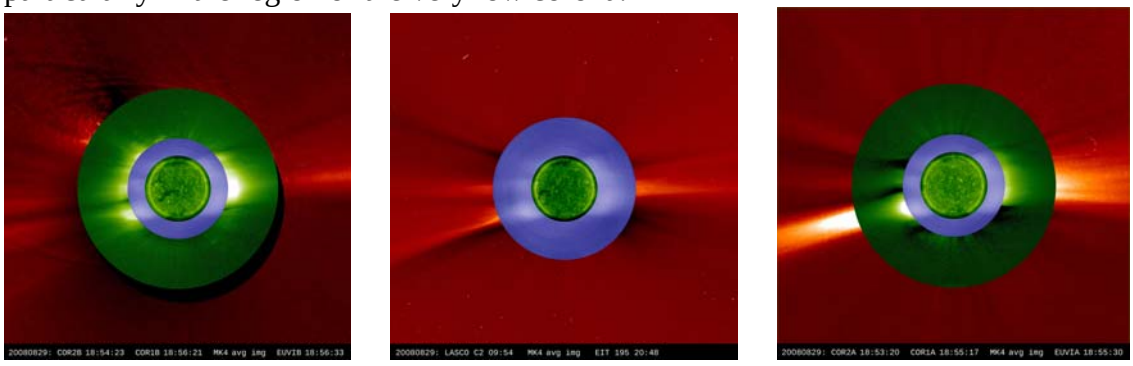


Figure 2. Composite images for the corona: Left: STEREO-B EUV (disk), MK4 (blue), COR1B (green) and COR2B (red). Center: EIT (disk), MK4 (blue) and LASCO C2 (red). Right: STEREO-A EUV(disk), MK4 (blue) COR1A (green) and COR2-A (red). MK4 provides unique observations of the very low corona that compliment space-based coronagraph fields-of-view.

The benefits of the new ground-based COSMO K-coronagraph

The Mauna Loa MK4 K-coronameter observes lower in the corona (1.12 solar radii) than any existing coronagraph. It furnishes unique observations of the density structure of the low corona used for studying features such as coronal mass ejections (CMEs), coronal cavities, helmet streamers, transient dimmings and polar plumes. Since 1980, MK3 and MK4 have observed over 1000 CMEs (examples are shown in Figs. 3 and 4) and have recorded density changes in the corona for over 3 solar cycles. This long-term record is unique and compliments the wider fields-of-view of space-based coronagraphs. MK3/4 observations have been combined with STEREO, Hinode, SOHO, SMM, SolWind, TRACE, and other space- and ground-based data to provide a more complete view of the corona and CME dynamics.

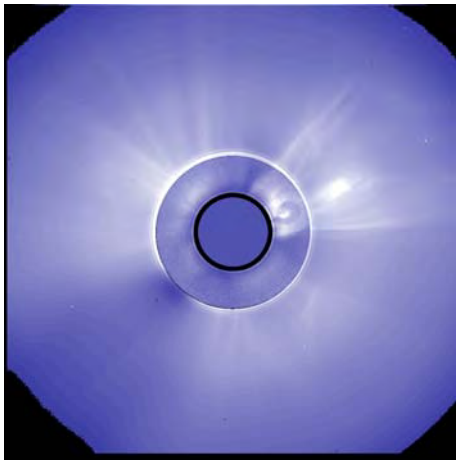
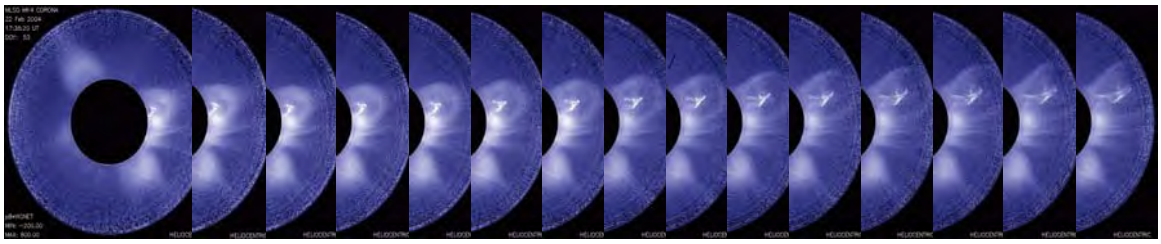


Figure 3. A CME observed on Aug 2, 2001 as seen in the low corona by MK4 and the outer corona by LASCO


Figure 4. A sequence of MK4 images, taken every 3 minutes, on February 18, 2003 recording a CME, showing the propagation of the CME through the low corona.

The MK4 is an internally occulted coronagraph. Externally occulted coronagraphs are ideal for viewing farther out into the corona and provide lower scattered light than internally occulted instruments. However, externally occulted instruments are not able to view very low in the corona unless the occulter lies at extremely large distances from the image plane. Space-based missions, such as the European Space Agency Proba-3, are invoking formation flying with one spacecraft acting as the occulter and a second spacecraft acting as the detector to view low in the corona. Such missions should provide high quality low coronal images but at great expense and for limited mission lifetimes. The primary source of unwanted scattered light for an internally occulted coronagraph comes from solar light falling directly on the objective lens. At pristine ground-based coronal sites such as Mauna Loa, the scattered light from the objective is much greater than the background sky light levels. Under these conditions, the scattered light from a ground-based coronagraph is the same as the scattered light for an internally occculted space-based coronagraph. ***There is no scattered light advantage in viewing the low corona from a space-based internally occulted corongraph.*** State-of-the-art ground-based coronagraphs, such as the COSMO K-coronagraph, will provide as high a quality signal-to-noise images as a space-based instrument given the same detector.

The Mauna Loa MK4 includes a 1-D detector with 4.86” pixels and the MK3 scanning device which rotates 360 degrees in 3 minutes building up an image of the low corona in polarization brightness (pB). It was installed at MLSO in 1998 replacing the MK3 detector and some optical components still utilizing the MK3 telescope, scanning apparatus and optical path. Its performance was an improvement over the MK3 but is constrained by the Mk3 scanning and optical system technologies designed in the 1970s. Improved detector technology and methods for removing systematic noise allow for the development of an entirely new telescope that will provide significant improvements in instrument performance. The COSMO K-coronagraph will take full advantage of these improvements providing exciting opportunities for discovery. Full funding for design and fabrication has been provided by NCAR and will include a 2-D high speed detector and a substantially better objective lens and hepa filter system for keeping the objective lens clean, which is the key to minimizing unwanted scattered light. The COSMO K-coronagraph will acquire images every 15 seconds over a field-of-view from 1.05 to 3 solar radii. It will provide the high-time cadence measurements needed to determine the rate of change of CME acceleration and lead to a better understanding of the forces driving CMEs that are needed to differentiate between the many competing models of

CME formation. A 15 second time cadence makes it possible to detect fast mode MHD coronal waves (as observed with EIT and STEREO EUVI) above the solar limb. It also provides better detection of precursor changes of CMEs that can lead to better prediction of CME onsets. The improved signal-to-noise (factor of 10 over MK4) can also provide observations of many halo CMEs from the ground. Currently MK4 is only able to detect the denser regions of the brighter halo events. Better signal will also provide significantly better density measurements in coronal holes, which is important for the study of solar wind acceleration and coronal heating and in identifying changes in the density structure preceeding the formation of a CME. The lower FOV will provide the first routine white light measurements into the 1st scale height of the corona for viewing CMEs as they form. The COSMO K-Coronagraph has other advantages over the MK4 including: many fewer moving parts, making it more robust and easier to maintain, improved calibration hardware and uniform spatial resolution. The new coronagraph is approximately the same size as the current MK4 and can fit on the existing spar at Mauna Loa, so additional costs for infrastructure are minimal.

Ground-based is cost effective

The COSMO K-Coronagraph has been fully funded by NCAR for $2.2 million, a small fraction of the cost of similar space-based telescopes. Some of the many advantages of ground-based observing include:

1) Instrumentation can be upgraded for very modest investments

2) Operating and maintaining ground-based operations are significantly lower than space-based platforms.

3) Facilities can be maintained for decades, providing a continual record of observations needed to understand changes occurring over solar cycle time periods

4) Telemetry costs are significantly cheaper than space-based, providing high time cadence data at a fraction of the cost of space-platforms.

One of the obvious advantages of space-based platforms are high duty cycles. However, ground-based networks, such as GONG, are able to deliver high duty cycles. An international network of low cost K-coronagraphs could provide nearly continuous observations of the low corona for space weather studies and forecasting, enhancing observations from space-based coronagraphs and other instrumentation.

Additional information about the COSMO K-coronagraph can be found at: http://mlso.hao.ucar.edu/mlso_newcoronagraph.html

Agency support for ground based observations

Agencies can enhance the benefits of ground-based observations through support of ground-based networks and new instrumentation, encouraging new instrument scientists, and support for research that includes ground-based observations.

References

Hansen, R. T., C. J. Garcia, S. F. Hansen, and E. Yasukawa. (1974), "Abrupt Depletions of the Inner Corona", *Publ.Astron.Soc.Pacific***, 86**, 500-515

Harrison, R. A., E. Hildner, A. J. Hundhausen, D. G. Sime, and G. M. Simnett. (1990), "The Launch of Solar Coronal Mass Ejections: Results from the Coronal Mass Ejection Onset Program", *Journal of Geophysical Research***, 95 (A2)**, 917-937

Hiei, E., A. J. Hundhausen, and D. G. Sime. (1993.), "Reformation of a Coronal Helmet Streamer by Magnetic Reconnection After a Coronal Mass Ejection", *Geophysical Research Letters***, 20 (24)**, 2785-2788

Ko, Y., J. Li, P. Riley, and J. C. Raymond. (2008), "Large-Scale Coronal Density and Abundance Structures and Their Association with Magnetic Field Structure", *Astrophys.J.***, 683**, 1168-1179

Low, B. C. (1996), "Solar Activity and the Corona", *Solar Physics***, 167**, 217-265

MacQueen, R. M. (1985), "Coronal Mass Ejections: Acceleration and Surface Associations", *Solar Physics***, 95**, 359-361

A Mission Whitepaper submitted to the 2013-2022 Decadal Survey in Solar and Space Physics

Stellar Imager (SI): developing and testing a predictive dynamo model for the Sun by imaging other stars

10 November 2010

Kenneth G. Carpenter (GSFC), Carolus J. Schrijver (LMATC), Margarita Karovska (CfA), Steve Kraemer (CUA), Richard Lyon (GSFC), David Mozurkewich (Seabrook Eng.), Vladimir Airapetian (GMU), John C. Adams (GSFC), Ronald J. Allen (STScI), Alex Brown (UCO/Boulder), Fred Bruhweiler (CUA), Alberto Conti (STScI), Joergen Christensen-Dalsgaard (U. Aarhus), Steve Cranmer (CfA), Manfred Cuntz (U. Texas/Arlington), William Danchi (GSFC), Andrea Dupree (CfA), Martin Elvis (CfA), Nancy Evans (CfA), Mark Giampapa (NSO/NOAO), Graham Harper (UCO/Boulder), Kathy Hartman (GSFC), Antoine Labeyrie (College de France), Jesse Leitner (GSFC), Chuck Lillie (NGST), Jeffrey L. Linsky (UCO/Boulder), Amy Lo (NGST), Ken Mighell (NOAO), David Miller (MIT), Charlie Noecker (BATC), Joe Parrish (Aurora Flight Systems), Jim Phillips (CfA), Thomas Rimmele (NSO), Steve Saar (CfA), Dimitar Sasselov (CfA/Harvard), H. Philip Stahl (MSFC), Eric Stoneking (GSFC), Klaus Strassmeier (AI-Potsdam), Frederick Walter (SUNY), Rogier Windhorst (ASU), Bruce Woodgate (GSFC), Robert Woodruff (LMSSC)

For additional information, please contact:

Dr. Kenneth G. Carpenter
Code 667, NASA-GSFC
Greenbelt, MD 20771
Phone: 301-286-3453, Email: Kenneth.G.Carpenter@nasa.gov

I. Summary of Science Mission Concept

The objective of the Stellar Imager (SI) mission is to accelerate the development and validation of a predictive dynamo model for the Sun and other magnetically active stars. In particular, we wish to understand the origins of variability in the Sun-Earth system and provide the foundations for more accurate long-term forecasting of solar/stellar magnetic activity and its impact on space weather, planetary climates, and life. To accomplish this, SI will:

- Characterize the patterns of surface magnetic activity for a large sample of Sun-like stars by obtaining time-resolved images with ~1000 resolution elements on their surfaces.
- Characterize the internal structure and differential rotation of these stars using asteroseismology with at least 30 resolution elements across the stellar diameter.
- Determine the dependence of dynamo action on mass, internal structure, flow patterns, and time by carrying out a population study of Sun-like stars. *This will enable testing of dynamo models over a few years of observations of many stars, instead of over many decades using only the Sun.*

Our understanding of the solar dynamo has been advanced significantly by combining solar observations with numerical experiments. Yet, none of the current models yields the Sun's average activity level; none of the self-contained numerical models has proven predictive capabilities; and the (quasi-)empirical models disagree strongly on forecasts for the next solar cycle. To make significant progress we really need to know what is important in making a dynamo go, i.e., the dependence of dynamo action on, e.g., mass, effective temperature, surface gravity, rotation rate, internal structure, flow patterns, and time. By choosing the correct stars for a population study we can vary one parameter at a time and see how that affects stellar activity.

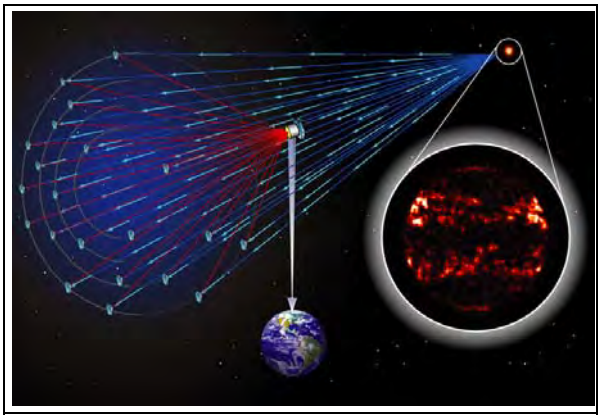


Fig. 1: SI - An array of 20-30 one-meter mirrors fly in precision formation to form a virtual parabola ~500m in diameter to enable sub-milliarcsecond spectral imaging of solar-type stars, e.g., as shown above in the light of CIV emission lines.

The Stellar Imager mission concept (Fig. 1) that will accomplish these goals is a space-based, UV/Optical Interferometer (UVOI). The required resolution dictates baselines ~500 meters, the imaging quality requires ~30 apertures, and the imaging rate requires apertures ~1.0 meter diameters. All three of these numbers can be decreased at the expense of a decreased number of targets and increased time to complete an image. The resulting instrument will have over 200x the spatial resolution of Hubble Space Telescope. It will enable 0.1 milli-arcsec (@2000Å) spectral imaging of stellar surfaces and, via spatially-resolved asteroseismology, measurements of the structure and differential rotation of stellar interiors.

The required technologies all have mid-level TRLs; the concepts are thoroughly tested, at least in the lab. Many have been deployed at ground-based astronomical observatories. ***However, to enable the SI mission in the late 2020's, significant technology development in the upcoming (2013+) decade is critical to qualify these technologies for space missions.*** The key technology needs include: 1) precision formation flying of many spacecraft and 2) closed-loop control of many-element, sparse optical arrays.

SI has been in the SEC/SSSC/Heliophysics Division Roadmaps since 2000 and is a Flagship "Landmark/Discovery Mission" in the 2005 Roadmap (Fig. 2). It was selected as a NASA Vision Mission (VM; "NASA Space Science Vision Missions" 2008, “Progress in Astronautics and Aeronautics”, vol. 224, pp. 191-227; AIAA, ed. M. Allen). The full SI Vision VM Study report, related science and technology whitepapers, and additional information and references can be found at: http://hires.gsfc.nasa.gov/si/. Here we first discuss the primary science goals of SI in further detail, and then outline the mission architecture, technology needs, and estimated costs.

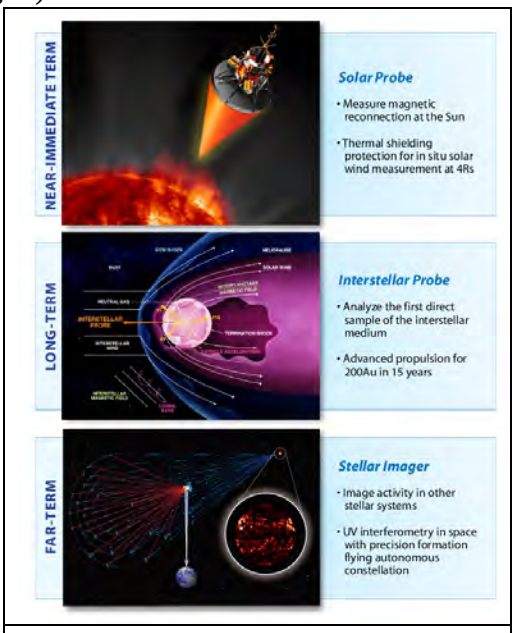


Fig. 2: Stellar Imager is identified as a “Landmark/Discovery Mission” in the 2005 SSSC (Heliophysics Division) Roadmap.

A. Advancing Solar/Space Physics by Improving our Understanding of Dynamos and Magnetic Activity in the Sun & Stars

A magnetic dynamo operates in all stars at least during some phases in their evolution. It regulates the formation of stars and their planetary systems, the habitability of planets, and the space weather around them. Dynamos also operate in objects including planets, accretion disks, and AGN. Although we know that flows in rotating systems are essential to such dynamos, there is no comprehensive dynamo theory from which we can derive the strength, the patterns, or the temporal behavior of stellar magnetic fields. Understanding the complex of non-linear couplings in a dynamo requires that we combine numerical studies and theory with observations of the evolving surface field of Sun and stars, of the average properties of magnetic fields in a large sample of stars, and of stellar internal flows. *These observations will enable us to constrain the source and sink properties of the surface magnetic field, by measuring the latitudinal and longitudinal flux-emergence patterns, the flux dispersal rate, and the differential and meridional flows. These properties are both important input parameters to flux-transport models and sensitive differentiators between dynamo models.*

For cool stars like the Sun, dynamo action persists from the very formation of the star throughout its existence as a fusion reactor. The Sun’s activity is modulated significantly from cycle to cycle, sometimes persistently for several decades. Activity decreased, for example, for decades in the 17th Century when Earth experienced the Little Ice Age. Observations of other cool stars are crucial to our understanding of solar/stellar dynamo action. This has taught us, for example, that convection is part of all solar-like dynamos, and that rotation regulates their strength. But we do not have a theory that explains why stars are as active as they are, why some show cycles and some do not, what causes the Sun’s cycles to differ from one to the next (e.g., the extra long duration of Cycle 23), and how a cyclic dynamo can restart after a Maunder-like (e.g. Little Ice Age) minimum. Hence, we cannot usefully forecast long-term space weather or reliably model the effects of stellar magnetism on the evolution of stars, planetary systems, planetary atmospheres, and thus the habitability of a planetary system.

Numerical modeling has taught us valuable lessons about dynamos, including the fact that they are highly non-linear and couple processes that occur over geometric extents as large as the full convection zones of stars down to the smallest convective scales. Our computer models are

therefore of necessity simplified, and hence require observational guidance from stars with a variety of properties. *Key to successfully developing a predictive dynamo theory is the realization that we need a population study of stars at high spatial resolution: we need to study the evolution of dynamo-driven activity in both latitude and longitude in a sample of stars like the Sun, and compare it to observations of young stars, old stars, binary stars, etc.* The potential for a breakthrough in our understanding lies in spatially-resolved imaging of the dynamo-driven emission patterns on this wide-ranging stellar sample. These patterns, and how they depend on stellar properties (such as convection, (differential) rotation and helicity, meridional circulation, and evolutionary stage/age), are crucial for dynamo theorists to explore the sensitive dependences on many poorly known parameters, to investigate bifurcations in a non-linear 3D dynamo, and to ultimately validate a model. *Direct UV/optical interferometric imaging* (0.1 mas observations in UV emission lines of magnetically active regions on stellar surfaces) *is the only way to obtain the required information on the dynamo patterns for stars of Sun-like activity.*

To address these science goals, the Stellar Imager is designed to provide UV/Optical sub-milliarcsec images and disk-resolved asteroseismology for a sample of stars similar to the Sun, as well for other cool stars with very different characteristics. SI will have access to an exciting array of distinct stars and stellar systems (see Fig. 3). An array diameter of 500 m is needed to resolve a medium-sized solar-type active region when observing a Sun-like star at ~4 pc. A km-sized array provides the necessary resolution out to ~8 pc. SI will, for the first time, enable imaging of magnetic activity of a variety of Sun-like stars (there are ~3 dozen F, G, K main-sequence (MS) stars within 8 pc) and many cooler M-type MS stars, including many nearby and more distant stars with shallow convective envelopes, fully-convective stars, close binary systems with dual active components magnetically coupled at a few stellar radii, compact RS-CVn-type binaries, mass-transferring Algol-type systems, symbiotic systems, and giant & supergiant stars. Imaging magnetically active stars and their surroundings will also provide us with an indirect view of the Sun through time, from its formation in a molecular cloud, through its phase of decaying activity, during and beyond the red-giant phase during which the Sun will swell to about the size of the Earth's orbit, and then toward the final stages of its evolution (see Fig. 4).

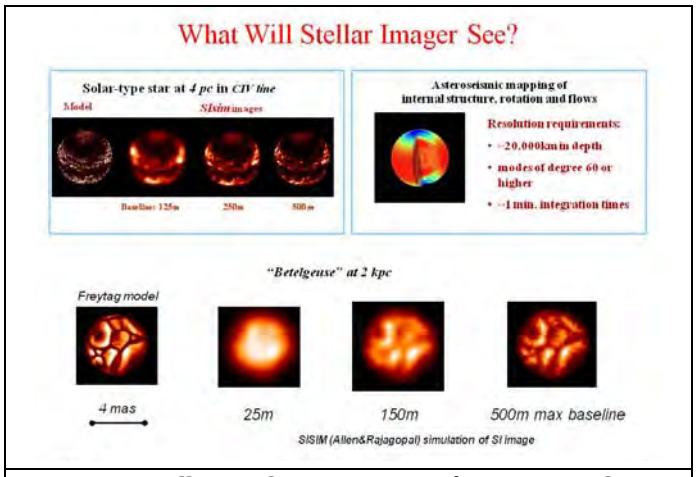


Fig. 3: SI will provide UV images of stars at resolutions only previously dreamed of, as seen in these simulations (using 30 mirrors in a non-redundant pattern with indicated max. baselines).

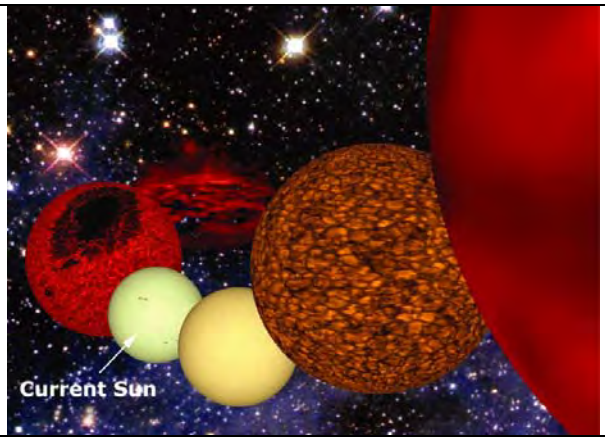


Fig. 4: Evolution of the Sun in time - SI will provide us a view by resolving stars representing each solar era.

B. Mission Architecture

NASA commissioned a "Vision Mission (VM) Study" of SI in 2004-2005, which developed a design that meets the performance requirements required to achieve the science goals described above. The baseline full-mission concept for SI was developed in collaboration with the GSFC Mission Design Lab (MDL) and Instrument Design Lab (IDL). The MDL worked on the overall design of a space-based Fizeau interferometer, located in a Lissajous orbit around the Sun-Earth L2 point. A variety of disciplines considered the implications of this general design, including power, guidance & navigation, flight dynamics, operations, communications, quality assurance, system engineering, etc. The IDL concentrated its efforts on the design of the beam-combining hub in the context of the selected overall architecture, again from a multiple-discipline viewpoint, and including accommodation of the MDL results. In addition to assisting in the development of the architecture, the Design Labs explored the technical feasibility of the mission and identified the technology developments needed to enable the mission in the late 2020's. In order to address its science goals, SI must have the following capabilities:

- **wavelength coverage: 1200 – 6600 Å**
- **access to UV emission lines** from Lyα 1216 Å to Mg II 2800 Å
  - Important diagnostics of most abundant elements
  - much higher contrast between magnetic structures and background
  - shorter baselines (UV saves 2-4x vs. optical, active regions 5x larger)
  - ~10-Å UV pass bands, e.g. C IV (100,000 K); Mg II h&k (10,000 K)
- **broadband, near-UV or optical** (3,000-10,000 K) for high temporal resolution spatially-resolved asteroseismology to resolve internal stellar structure
- angular resolution of 50 μas at 1200 Å (120 μas @2800 Å) to provide ~1000 pixels of resolution over the surface of nearby (4pc) dwarf stars, and more distant giant and supergiant stars.
- **spectral resolution** of R>100 (min) up to R=10000 (goal)
- **long-term (~ 10 year) mission**, to enable study of stellar activity cycles:
  - individual telescopes/hub(s) can be refurbished or replaced

The VM study defined a detailed flow down of requirements from science goals to data and measurements requirements to engineering implications to the key technologies needed to implement the mission. The baseline mission architecture shown in Fig. 5 was derived from these requirements. The selected design is a space-based (at Sun-Earth L2), UV-Optical Fizeau Interferometer with 30 one-meter primary mirrors, mounted on formation-flying "mirrorsats" distributed over a parabolic virtual surface whose diameter can be varied from 100m up to as much as 1000m, depending on the angular size of the target to be observed. Table 1 summarizes the mission and performance parameters of the baseline SI design.

C. Technology Drivers

The two major technology challenges (the "tallest poles") to building SI are:

- **precision formation-flying (PFF) of ~20-30 spacecraft** (including precision metrology over baselines of 100's of meters)
- **wavefront sensing and real-time autonomous analysis and closed-loop optical control of a many-element sparse array**

These technology challenges have been addressed, along with the less difficult ones, prior to and during the SI Vision Mission (VM) study and in GSFC Integrated Design Center (IDC) studies.

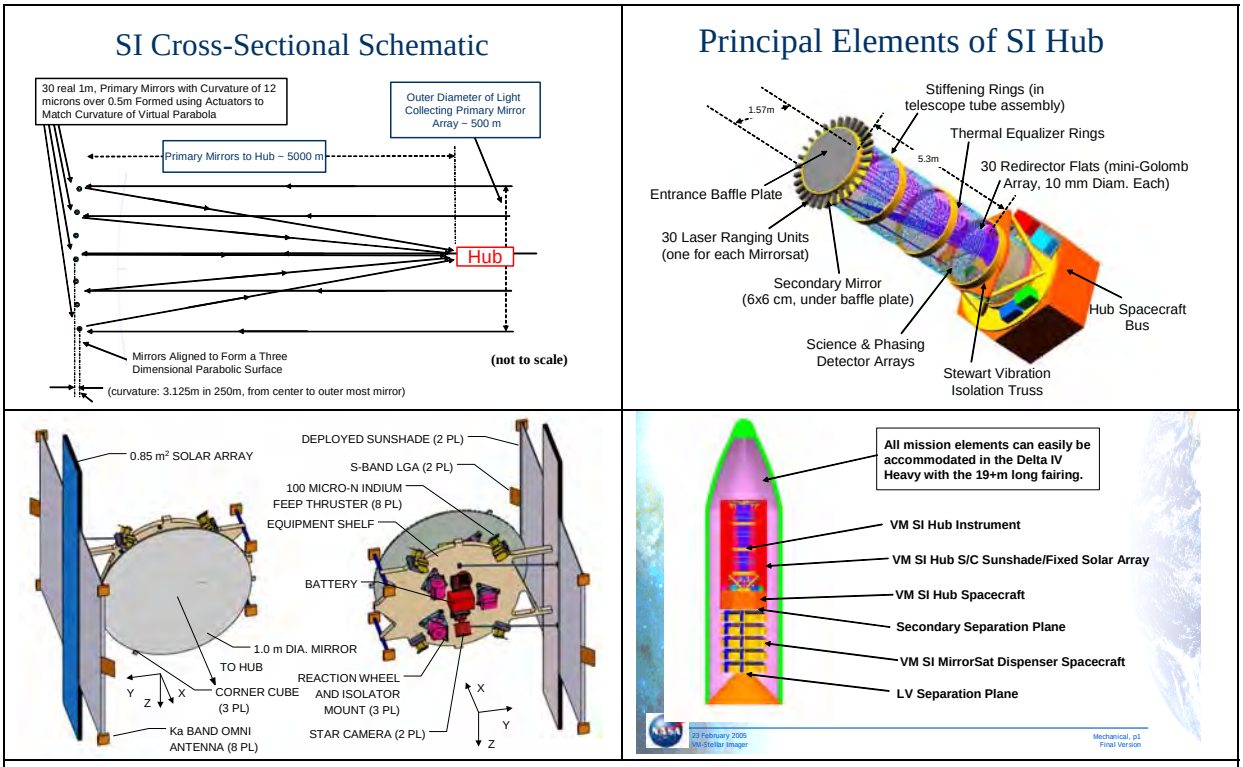


Fig. 5: An overview of the VM SI Design, the details of the hub & mirrorsats, and one launch option.

Table 1: Mission and Performance Parameters from the Vision Mission Study

Parameter	Value	Notes
Maximum Baseline (B)	100 – 1000 m (500 m typical)	Outer array diameter
Effective Focal Length	1 – 10 km (5 km typical)	Scales linearly with B
Diameter of Mirrors	1 - 2 m (1 m currently)	Up to 30 mirrors total
λ-Coverage	UV: 1200–3200 Å ; Opt: 3200–6600 Å	λ-Sensing in optical
Spectral Resolution	UV: 10 Å (emission lines) UV/Opt: 100 Å (continuum)	
Operational Orbit	Sun-Earth L2 Lissajous, 180 d	200,000x800,000 km
Operational Lifetime	5 yrs (req.) – 10 yrs (goal)	
Accessible Sky	Sun angle: $70^o \le \beta \le 110^o$	Entire sky in 180 d
Hub Dry Mass	1455 kg	For each of 2
Mirrorsat Dry Mass	65 kg (BATC design)	For each of 30
Ref. Platform Mass	200 kg	
Total Propellant Mass	750 kg	For operational phase
Total Mass to Orbit	4355 kg	
Class of Launcher	Delta IV Heavy – 19 m fairing	
Angular Resolution	50 μas – 208 μas (@1200–5000Å)	Scales linearly ~ λ/B
Typical time to image star	< 5hr for solar type; < 1 day for supergt,	
Imaging time resolution	10 – 30 min (10 min typical)	Surface imaging
Seismology time res.	1 min cadence	Internal structure
# res. pixels on star	~1000 total over disk	Solar type at 4 pc
Minimum FOV	> 4 mas in single exposure	Larger via mosaic
Minimum flux detectable at 1550 Å	5.0×10^{-14} ergs/cm^2/s integrated over C IV lines	10 Å bandpass
Precision Formation Fly.	s/c control to mm-cm range	
Optical Surfaces Control	Actuated mirrors to μm-nm range	
Phase Corrections	to λ/10 Optical Path Difference	
Aspect Control/Correct.	3 μas for up to 1000 sec	Line of sight mainten.

Credible and feasible approaches to the successful development of all these technologies were derived during the course of those studies and are documented in the SI VM Final Report, but these must be funded and implemented during the 2013 decade to enable the technologies to be ready for flight in the following decade. ***All of these technologies exist – it is simply a matter of extending such work to improve performance and/or operational scale.*** Various ground-based experiments, e.g., the MIT SPHERES, GSFC FFTB, and JPL FCT formation flying testbeds and GSFC Fizeau Interferometer (large optical array phasing/synthetic imaging) Testbed, can be used to mature these technologies in preparation for validation on a small, space-based experiment (for PFF) and on ground-based optical interferometers such as MROI (for optical array phasing).

II. Cost Estimates

We estimate the total cost, over the full decade, of the ground-based technology work described above, at ~$37M. Costs are estimated based on experience-to-date with the efforts noted above. Flight validation of PFF technologies is also required by other potential missions and would build on the results of the Orbital Express (DARPA) and PRISMA (Swedish Sp. Corp.) missions and the upcoming PROBA-3 (ESA) and F-6 (DARPA) missions. If an additional mission becomes necessary, it is expected to be funded by cross-enterprise Strategic Technology Development initiatives, with additional cost-sharing possible from DoD and/or ESA sources. ***An endorsement by the Decadal Survey of the SI science and mission as worthy of flight in the following (2023) decade would enable us to pursue funding in the 2013 decade for the technology development, mission design, and international coordination***.

Over the last 8 years we have performed a number of SI design studies in the GSFC IDC, including studies of the overall mission, the beam-combining hub, and the science payload. These studies involved experts from a full range of technical disciplines, including power, guidance & navigation, flight dynamics, operations, communications, quality assurance, system engineering, etc., and cost estimation. These studies explored the technical feasibility of the mission, identified the technology developments needed to enable the mission in the 2024+ timeframe, and provided cost estimates as described below. We have used both Parametric and Analogous methods to estimate the cost of the SI mission. Parametric Cost Modeling at GSFC includes the use of PRICE-H (Hardware) and/or SEER-H (Hardware) for mission hardware cost estimating. The Analogous cost methodology is based on historical data and involves comparison and extrapolation to like items or efforts in previous missions and design efforts.

We first considered the costs of the mirrorsats and beam-combining hub, using both Price-H and the JSC mass-based Advanced Missions Cost estimator (http://cost.jsc.nasa.gov/AMCM.html). Costs based on Price-H from the GSFC IDC sessions were used and inflated to 2009$. The mirrorsats were studied during 2001, the hub spacecraft in 2004, and the payload in 2005. Second, the spacecraft/vehicle level mass-based cost estimator from the JSC web site (http://cost.jsc.nasa.gov/SVLCM.html) was used assuming the parameters shown in Table 1. We also compare in Table 2 the cost estimations for the mirrorsats and hub/payload. Inflating the JSC 2004$ estimates to 2009$ gives $1.5B which is approximately 36% higher than the $1.1B cost for mirrorsats and hub/payload based on Price-H estimates. The Price-H estimates were used in the total mission

Table 2. Cost Comparisons for Mirrorsats and Hub/Payload ('09$)

Component	Price-H/IDC	JSC Estimator (avg. complexity)	Assumptions
30 Mirrorsats	$656M	$759M	65 kg each
One Payload Hub	$466M	$767M	1455 kg
Component Total	$1.1B	$1.5B	

cost derived below, since they were generated using estimates of specific hardware component costs and we believe them to be more reliable.

Project management, mission systems engineering, mission assurance, and system integration and testing are estimated as percentages of hardware costs based on cost averages from the extensive GSFC mission flight experience. Science costs in Phases C/D are based on $500K per instrument per year for a flagship mission class, while Phase E is scaled from Hubble Space Telescope (HST), accounting for fewer discrete targets and programs/year. Four instruments are included: UV and visible cameras, a wavefront sensing and control system, and a "light bucket" spectrograph, which are costed based on HST experience. In Phase E, $ 15 M/year (of line 4.0) goes directly to the astronomical community in the form of grants. Reserves are computed as 30% of Phase B-E cost before adding the launch vehicle cost. EPO costs are based on appropriate levels for a Flagship mission. We selected a Delta-IV Heavy as an example launch vehicle, and very conservatively estimated its cost from information on the Astrophysics Strategic Mission Concept Studies workshop webpage, i.e., 4.2 times the cost of a small launch vehicle ($160M in 2015$), and deflated to 2009$. The SI hardware can be accommodated in other launch vehicles with similar characteristics.

The process above yields a total mission cost in 2009$ of $2.9B. ***We note that there is significant international interest (Carpenter et al., 2009 Ap & Sp. Sci., 320, 217) in the SI mission as well as from the Astrophysics community, which is interested both in the stellar astrophysics and other topics addressable by a high resolution UV/Optical observatory (e.g., central engines of AGNs, mass accretion processes, dynamical structures in supernovae and planetary nebulae) and we would thus expect this cost to be shared by international (esp. European) partners and by the Astrophysics Division.*** The breakout and methodologies used to estimate this mission cost are shown in Table 3. We have left our formal estimates in FY09$, as produced by our last major costing exercise, since the current extremely low level of inflation would not materially change the result relative to other uncertainties in the costing.

Table 3: SI Mission Cost Estimate ($ M, 2009) and Methodology - updated 3/23/09

Cost Element	Phase A	Phase B	Phase C/D	Dev Total	Phase E	Mission Total	Cost Methodology
Project Elements:							
1.0 Project Management	2	9	91	102	9	**111**	9% hardware costs B-D;9% of Mission Ops
2.0 Mission Sys Engr	2	8	81	92	8	**100**	90% of Project Mgmt
3.0 Mission Assurance	2	3	33	38	3	**41**	40% of Systems Engineering
4.0 Science	2	3	6	11	129	**140**	Based on Flagship mission data; analogous ops costs
5.0 Payload & SC	5	52	409	466	0	**466**	MDL and IDL estimates, Price-H
6.0 Flight Systems (incl. 30 mirrorsats)	6	52	598	656	0	**656**	MDL and IDL estimates, Price-H
7.0 Mission Ops	1	4	15	21	103	**124**	MDL estimates; analogous cost, and cost calculators
9.0 Ground System	2	2	2	6	10	**16**	MDL estimates; and analogous cost
10.0 System I&T	1	10	101	112	0	**112**	10% of hardware costs
Sub total	**24**	**143**	**1336**	**1504**	**263**	**1766**	
Reserves	0	43	401	444	64	**508**	
Sub total w/reserves	**24**	**186**	**1737**	**1948**	**327**	**2273**	
Elements w/o cont:							
8.0 Launch Vehicle	0	0	571	571	0	**571**	Delta IV H cost from ASMCS webpg.
11.0 E/PO	0	3	12	15	15	**31**	Phase B-E, appropriate for Flagship
Mission Total:	**24**	**190**	**2320**	**2534**	**342**	**2875**	

The Focusing Optics X-ray Solar Imager (FOXSI)

2010 Heliophysics Decadal Survey White Paper

S. Christe[1], S. Krucker[2,3], L. Glesener[2,4], B. Ramsey , T. Takahashi[6,7], R. P. Lin[2,4], B. Dennis[1], A. Y. Shih[1], G. Holman[1], J. Klimchuk[1], L. Fletcher[8], E. Kontar[8], A. O. Benz[3,10], N. Vilmer[9], A. Mackinnon[8], M. Piana[11], G. Mann[12], L. Golub[13], P. Gallagher[14], S. Masuda[15], T. Sakao[16], S. Freeland[17], D. Williams[18], J. Sylwester[19], A. M. Veronig[20], F. Fárník[21], S. Matthews[18], S. White[22]

[1] NASA Goddard Space Flight Center, Greenbelt, MD, 20771-0001, USA
[2] Space Sciences Laboratory, University of California at Berkeley, Berkeley, CA, 94720, USA
[3] University of Applied Sciences, Institute of 4D, Technologies, 5210 Windisch, Switzerland
[4] Physics Department, University of California at Berkeley, Berkeley, CA, 94720-7450, USA
[5] NASA Marshall Space Flight Center, Huntsville, Alabama, USA
[6] Institute of Space and Astronautical Science, Sagamihara, Kanagawa, 229-8510, Japan
[7] Department of Physics, University of Tokyo, Hongou, Bunkyo-ku, Tokyo, 113-0033, Japan
[8] Department of Physics and Astronomy, University of Glasgow, G12 8QQ, UK
[9] LESIA-CNRS, Observatoire de Paris, 5 Place Jules Janssen, MEUDON, France
[10] Institute of Astronomy, ETH Zürich, 8092 Zurich, Switzerland
[11] Dipartimento di Informatica, Universita di Verona, Strada le Grazie,Verona, Italy
[12] Astrophysikalisches Institut Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany
[13]Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA
[14] Astrophysics Research Group, School of Physics, Trinity College Dublin, Dublin 2, Ireland
[15] STEL, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan
[16] ISAS, JAXA, 3-1-1 Yoshinodai, Kanagawa, Japan
[17] Stanford Lockheed Institute of Space Research, Palo Alto, USA
[18] UCL Mullard Space Science Lab, Holmbury St Mary, Surrey, RH5 6NT, U.K.
[19] Space Research Centre, Polish Academy of Sciences, Kopernika 11, Wrocław, Poland
[20] Institute of Physics, University of Graz, Universitätsplatz 5, 8010 Graz, Austria
[21] Astronomical Institute, Academy of Sciences of the Czech Republic, 25165 Ondˇrejov
[22] Space Vehicles Directorate, AFRL, Kirtland AFB, Albuquerque NM

Executive Summary

Particle acceleration on the Sun is one of the main unsolved problems in heliophysics. Accelerated electrons moving through a plasma radiate hard X-ray (HXR) emission through the well-known process of bremsstrahlung. HXR observations therefore provide quantitative measurements of energetic electrons accelerated by the Sun. Present-day HXR instrumentation does not have the sensitivity to see faint HXR emission from electrons accelerated in the corona, nor the dynamic range to see faint emission in the presence of bright HXR flare footpoints. Existing observations therefore show us only where energetic electrons are stopped, but not where they are accelerated, nor along what path they escape from the acceleration site. The next breakthrough in understanding solar particle acceleration requires that electrons be imaged directly in the corona with good spectral resolution. New HXR focusing optics combined with position-sensitive solid-state detectors with ~1 keV spectral resolution can now provide these measurements. This white paper presents a mission concept for the Focusing Optics X-ray Solar Imager (FOXSI). This future instrument will tell us where electrons are accelerated, along which field lines they travel away from the acceleration site, where they are stopped, and how some electrons escape into interplanetary space, with simultaneous spectral information of the accelerated electrons.

Introduction

The Sun is the most energetic particle accelerator in the solar system, producing ions up to tens of GeV and electrons to tens of MeV. Large solar flares are the most powerful explosions in the solar system, releasing up to 10^{32} – 10^{33} ergs within 10^{2} – 10^{3} s. The accelerated 10 – 100 keV electrons (Lin & Hudson 1976; Emslie et al. 2005) and sometimes 1 MeV/nucleon ions contain up to 50% of the total energy released indicating that particle acceleration and the energy release processes are intimately linked. CMEs, another form of solar eruptive event, release a significant fraction of their energy in accelerated particles as well (see, e.g. Mewaldt 2006). While it is known that impulsive solar energy release is powered by magnetic energy, the energy release and particle acceleration processes are not yet understood. Hard X-ray (HXR) spectroscopic imaging observations provide sensitive quantitative measurements such as the energy spectrum, density, location, and energy content of accelerated electrons. HXR observations, therefore, hold the key to understanding the particle acceleration mechanism present in solar impulsive energy release.

So far, propagating electrons are rarely observed in the corona through their coronal HXR bremsstrahlung signatures though this is where they are thought to be accelerated by many types of solar impulsive energy release such as flare, CMEs, radio bursts (type I, type III, etc.) and jets. This is due to limitations in current HXR spectroscopic imagers. The purpose of this white paper is to describe a new instrument concept, called FOXSI. Short for the Focusing Optics X-ray Solar Imager, FOXSI will enable the next breakthrough in understanding particle acceleration by providing up to two orders of magnitude greater sensitivity and dynamic range compared to previous instruments. **FOXSI will observe energetic electrons directly in the corona where they are accelerated.**

FOXSI Science

Flares

According to the standard model of solar flares, particle acceleration occurs at or near the magnetic reconnection side in the corona. While details of the standard model are still open to question, multiple observations support the existence of a coronal acceleration region such as timing studies by Aschwanden et al. (1995) and, more recently, observations of paired coronal thermal sources (Sui and Holman 2003), and statistical studies of partially disc-occulted flares (Krucker & Lin 2008).

One of the most iconic examples of coronal acceleration is the Masuda flare (see Figure 1), which shows a coronal nonthermal HXR source distinct from the flare

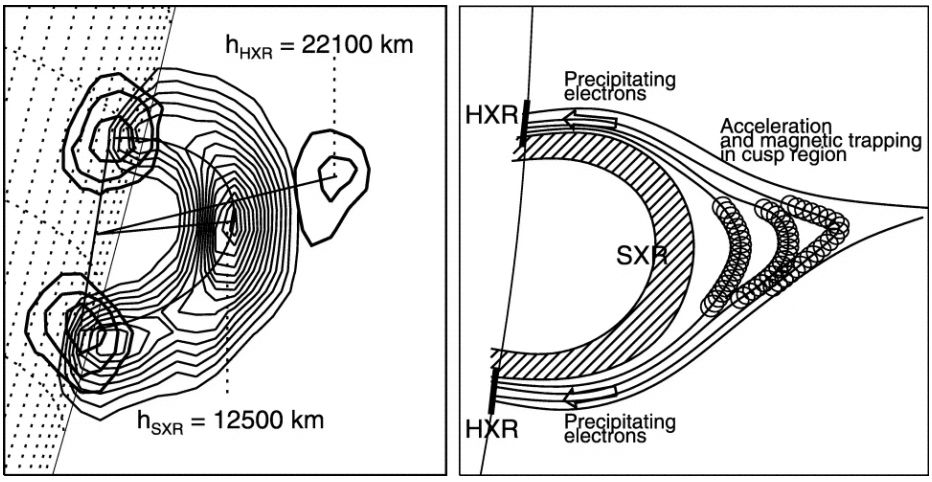


***Figure 1**: **The Masuda flare.** (Left) HXR Observations of the Masuda flare (1992 January 13 1728 UT) which show a Yohkoh HXT 23–33 keV image (thick contours) of two footpoints and an above-the-loop HXR source interpreted as the acceleration region itself. Be119 SXT image (thin contours) outline the thermal loops (Masuda et al. 1994). (Right) The standard flare model interpretation of the Masuda flare. Figure from Aschwanden 2004.*

footpoints (Masuda et al. 1994). Masuda-type observations are currently very rare though coronal sources in partially occulted flares are ubiquitous, found in ~90% of cases studied (Krucker & Lin 2008). This is because coronal HXR sources are generally faint compared to the bright HXR footpoints. Similarly, thermal emission from the footpoints is lost in the limited dynamic range of the much brighter thermal source seen in the corona. The study of the coronal acceleration region is currently limited to cases which display extraordinarily high coronal HXR intensities or have occulted footpoints. A comparison of partially occulted and non-occulted events show that fluxes at 30 keV are 10 to 100 times larger when the footpoints are visible. Current instrumentation has a dynamic range limited to 10 and thus cannot often image coronal sources and footpoints together. **New grazing incidence HXR focusing optics can now provide the necessary dynamic range and sensitivity to always observe coronal HXR sources even in the presence of bright footpoint emission.**

Detailed measurements of coronal bremsstrahlung HXR spectral parameters and their evolution will constrain flare particle acceleration mechanisms. Coronal HXR sources can give direct information on the energetic electrons in the corona such as their spectra, energy content, and location, as they are being accelerated. Observations of simultaneous footpoint and coronal HXR sources will allow us to understand the transport effects at work (e.g. return currents, wave-particle interactions) as electrons travel from the acceleration region to the footpoints. Coronal trapping can be investigated by examining the evolution of coronal HXR spectra, which should harden due to the shorter collisional loss times of lower-

energy electrons. Combined with knowledge of coronal densities provided by radio emission, EUV, or SXR observations can determine the instantaneous number of accelerated energetic electrons and the energetics of the acceleration process as a function of time. Hence **FOXSI will provide quantitative measurements of the population of accelerated electrons within the acceleration region itself.**

Radio-producing Electrons

HXR emission is also expected from electrons that produce coherent radio bursts seen in meter and centimeter wavelengths (Bastian et al. 1998). Radio emission is frequently associated with escaping electrons such as radio type III bursts or type II radio bursts associated with CME shocks (see next section). Type III bursts are thought to be produced by beams of escaping electron traveling through the corona. Observed in plasma emission generated at the local plasma density, they are frequently associated with solar energetic particles (SEPs) detected at Earth. Radio observations provide unparalleled sensitivity to detect these events near the Sun yet they do not provide quantitative information about these electrons. Although these beams generally travel in low densities, they also radiate in HXRs, albeit at a much lower rate than downward moving flare beams (Saint-Hilaire et al. 2009). The HXR observation of a type III beam is therefore difficult for two reasons: relatively small fluxes and the large dynamic range necessary since such bursts are often associated with flares. The HXR flux of an escaping electron beam consisting of 10^{33} electrons above 5 keV, consistent with *in-situ* electron measurements near 1 AU (Lin et al. 1973), propagating away from the Sun can easily be detected by FOXSI (Saint-Hilaire et al. 2009). **FOXSI will, for the first time, provide quantitative information about the energy content and spectra of escaping electrons events (SEPs) directly at the Sun.** This could then be compared with particles observed at Earth in order to understand the transport processes active between here and the Sun and the acceleration mechanism itself.

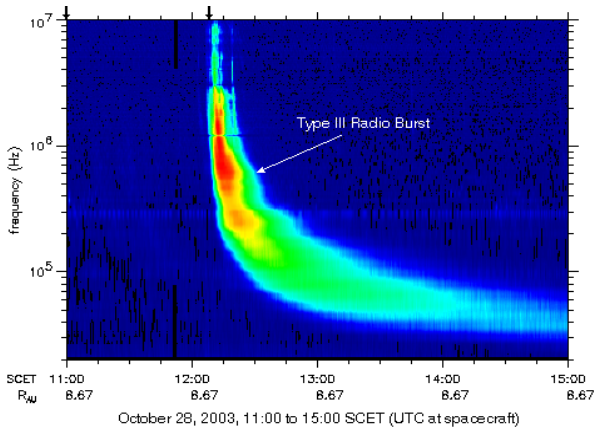


Figure 2: Type III Radio burst*. Frequently associated with flares, radio type III bursts are associated with beams of escaping electrons.*

CMEs and SEPs

Coronal mass ejections are known to drive interplanetary (IP) shocks and are frequently associated with powerful flares that accelerate particles to high energies. These shocks are frequently associated with type II radio bursts produced in the corona and IP medium. Solar energetic particles (SEPs) are frequently associated with type II radio bursts providing strong evidence that these shocks accelerate

particles. Similar to flares, CME-driven shocks are very efficient particle accelerators with up to 10% of the CME's total energy being converted to SEPs (see, e.g. Mewaldt 2006). Though IP shocks are frequently detected *in-situ* at Earth, they have already evolved and do not reflect the properties present at the Sun where most of the particle acceleration is thought to occur. Radio observations are currently the best way of observing these electrons near the Sun yet cannot give detailed information about the energetics of these accelerated electrons. In addition, type II radio bursts are only sensitive to electrons in the energy range of 0.2 to 10 keV (Bale et al. 1999). FOXSI observations of shock-accelerated electrons in the low corona would provide unprecedented information about these electrons before they arrive at Earth such as their spectra and location. **FOXSI observations will allow us understand how shocks accelerate particles and the conditions necessary for shocks to be efficient accelerators.**

Coronal Heating and Quiet Sun network flares

The energy released by flares is one of the proposed mechanisms to explain the million degree solar corona (i.e, the coronal heating problem). Large- and medium-sized flares are well observed but are known to produce insufficient amounts of energy to heat the corona (Crosby, Aschwanden, & Dennis, 1993). Parker (1988) has suggested that a likely candidate to explain coronal heating is continuous flaring at small scales, so-called nanoflares. Therefore, there is great interest in studying the smallest observable events in order to ascertain whether small flares can be responsible for the high temperature of the solar corona.

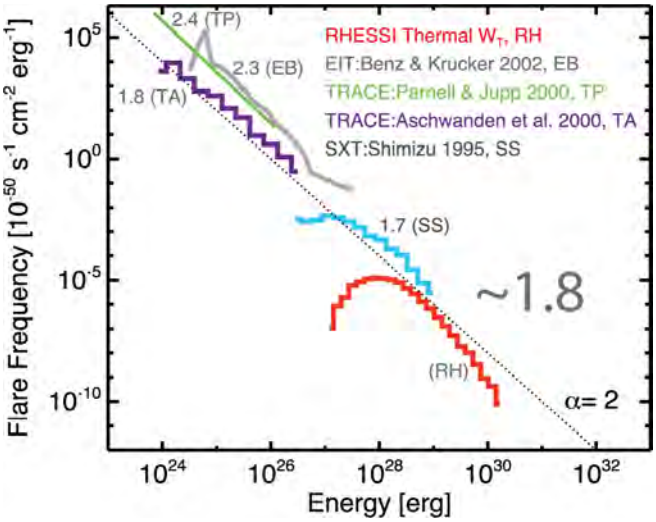


Figure 3: The flare frequency distribution. *The smallest observed flares have only been measured in thermal emission in the EUV by instruments such as TRACE and SOHO/EIT. These events are observed outside of the active regions, in the quiet Sun. Their slope have significant uncertainty. At higher energies, flare energies are more certain because estimated from thermal or nonthermal HXR emission. FOXSI observations of the smallest events would answer whether the quiet Sun flares are the same as larger flares and whether they heat the corona.*

RHESSI, with 100 times the sensitivity of previous mission at 10 keV, has detected nonthermal emission from microflares from active regions in the 3-15 keV range. The nonthermal energy content in these HXR microflares is derived to be down to 10^{26} ergs, showing that previous studies have underestimated the energy content in small flares (Benz and Grigis 2002; Krucker et al. 2002). However, the total

energy of these microflares is still not enough to account for coronal heating (Christe et al. 2008; Hannah et al. 2008).

The smallest observed events have only been seen in thermal soft X-rays and EUV outside of active-regions in the so-called quiet corona (Krucker et al. 1997; Krucker & Benz 1998; Aschwanden et al. 2000; Parnell & Jupp 2000; Benz & Krucker 2002). These events are referred to as network flares, even though it is has not been established that they are similar to normal flares observed in active regions. The current sensitivity of RHESSI is not good enough to detect these events in the quiet corona (Christe, Krucker and Lin 2008, Hannah 2008). **The large sensitivity provided by FOXSI will, for the first time, make it possible to search for nonthermal HXR emission from network flares.**

Radio observations have revealed the existence of nonthermal electrons in some network flares (Krucker et al. 1997), suggesting that network flares are heated in a similar way to regular flares occurring in active regions. If network flares are indeed smaller versions of regular flares, the thermal heating seen in EUV and SXR is due to nonthermal electrons. Since network flares are only heated to temperatures below 2 MK (Krucker et al. 1997), HXR emission above 3 keV is expected to be dominated by nonthermal emission. The absence of nonthermal emission would clearly establish that the brightenings in the quiet corona are not associated with accelerated energetic electrons. **FOXSI will determine whether brightenings in the quiet corona are similar to active region flares and whether they are the cause of the million degree corona. In addition to the search for nonthermal emission, FOXSI will be able to search for faint high temperature (>5 MK) coronal plasma which may be another signature of the coronal heating process.**

FOXSI Instrumentation

RHESSI HXR measurements are obtained with non-focusing rotation modulation collimator (RMC) imaging technique (Hurford et al. 2002). RMCs and other types of non-focusing imaging, however are limited in three ways: (1) limited effective area, (2) large non-solar background, and (3) limited dynamic range (defined as the ratio of strongest to weakest observable source) due to indirect imaging. Newly developed HXR grazing-incidence focusing optics can overcome all of these limitations.

Rotating modulator collimators as used in RHESSI and other types of non-focusing imaging have the disadvantage of a dynamic range limited by the entire field of view. Even for large flares, RHESSI's dynamic range is <30, making it difficult to see faint and bright sources simultaneously. For microflares, the dynamic range is still lower (<5) due to low counting statistics. For focusing optics, the dynamic range (and resolution) is set by the shape of the point spread function (PSF), the more peaked the PSF the better the dynamic range. Currently, the dynamic range of metal

HXR focusing optics for sources separated by 25 arcsec (a typical separations between footpoints and coronal sources), is >50. Furthermore, calibration will allow for the observed images to be deconvolved, further increasing the dynamic range to >100 for events with good statistics. For well-separated sources, the dynamic range is even larger, up to ~1000.

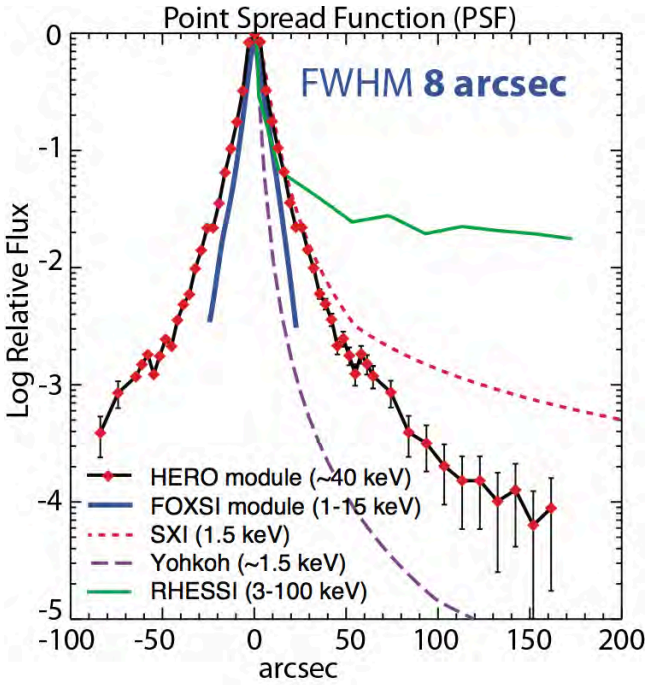


***Figure 4**: **The point spread function** for a single HERO telescope module compared to other similar X-ray instruments. RHESSI is the only non-focusing system and the only solar instrument in the HXR range. The measured point spread function for a FOXSI module is also shown with the expected FWHM of 8 arcsec and an HPD of 25 arcsec.*

Indirect imaging provides effective area that is proportional to physical detector area. High sensitivity can then only be achieved through large detectors yet such detectors are difficult to shield against non-solar background. New solid state detectors (e.g. Si, CdTe, CZT) can now provide high quantum efficiency in the HXR range with pixel sizes perfectly suited for FOXSI (≤500 μm). Compared to RHESSI, the area sensitive to non-solar background for FOXSI is ~10^{-6} times smaller. Combined with active-shielding, negligible non-solar background can be expected. Event detection then requires >10 times fewer counts than RHESSI for the same effective area.

HXR optics typically provide ~3 cm^2 of effective area per kg. Therefore, FOXSI with a mass comparable to RHESSI (130 kg), a Small Explorer class mission, would provide close to 400 cm^2 of effective area nearly 7 times that provided by RHESSI. **FOXSI will be more than 100 times more sensitivity than RHESSI.**

Current and future FOXSI

In preparation for a future satellite-based mission, FOXSI is currently being developed as a sounding rocket payload (Krucker et al. 2009). With a launch date at the end of 2011, the FOXSI sounding rocket has a focal length of only 2 m and an energy range of 1 to 20 keV. Based in large part on technologies developed for the High Energy Replicated Optics (HERO) balloon payload and the Hard X-ray Telescope (HXT), part of the JAXA mission, Astro-H, FOXSI is the first mission which applies HXR focusing optics to solar observations. Making use of silicon strip detectors (to be replaced by CdTe for its second flight), FOXSI will make the first observations of nonthermal electrons in network flares occurring in the quiet Sun.

The Nuclear Spectroscopic Telescope Array (NuSTAR), scheduled for launch in 2012, is a NASA Small Explorer (SMEX) which uses glass HXR focusing optics for astrophysical observations in the 6 to 80 keV energy range (Harrison et al. 2010). NuSTAR consists of two co-aligned grazing incidence optics focusing on two shielded solid state CdZnTe pixel detectors 10 m away. It will provide observations of astrophysical objects with a sensitivity 50-100 times better than the past with an effective area of ~1000 cm^2. NuSTAR, though not optimized for solar observations (see Smith et al 2010 white paper), is an example of what a SMEX-like FOXSI mission with good spatial resolution could achieve.

The FOXSI sounding-rocket and NuSTAR are pathfinders for a future satellite-based FOXSI. Such a mission could be developed as a stand-alone mission in a SMEX-like budget similar to NuSTAR or could be part of a larger mission (e.g. RAM, SEE 2020, SSE white papers) with characteristics such a focal length appropriate to the mission. Table I describes one possible version of FOXSI with a 10 m focal length.

Conclusion

New technological advances in grazing incidence focusing optics and solid-state detectors now make it possible to observe the Sun in HXR with unparalleled sensitivity and dynamic range. **FOXSI will be able to image where electrons are accelerated along which field line they travel away from the acceleration site, where they are stopped, and how some electrons escape into interplanetary space while simultaneously providing detailed information about the energetics of these electrons as a function of time.**

Table I: FOXSI Strawman Mission/Instrument Characteristics. *A larger mission can easily be scaled from this baseline concept mission.*

Focal length	10 m
Energy Range	1 to 60 keV
Energy Resolution	<1 keV
Spatial Resolution	8 arcsec (FWHM) corresponding to pixel size of ~400 μm
Field of View	~12 arcmin
Mass	100 kg (~3 cm^2/kg of optics)
Power	60 Watts
Dynamic Range	Up to 1000 (100 x RHESSI)
Sensitivity	100 x RHESSI
Cost (Mission)	SMEX
Cost (Instrument)	~$40M

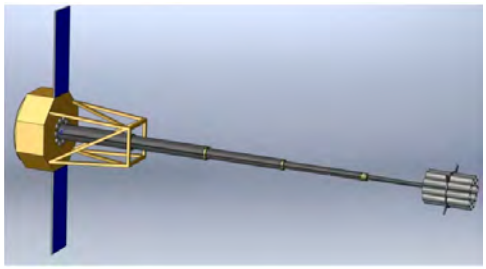


Figure 5: FOXSI Strawman mission concept. *For illustration, a FOXSI strawman concept with an expandable boom and seven telescope modules.*

Works Cited

Aschwanden, M., Tarbell, T., Nightingale, R., Schrijver, C., Title, A., Kankelborg, C., et al. (2000). Time Variability of the ``Quiet'' Sun Observed with TRACE. II. Physical Parameters, Temperature Evolution, and Energetics of Extreme-Ultraviolet Nanoflares. *The Astrophysical Journal , 535*, 1047.

Bastian, T., Benz, A., & Gary, D. (1998). Radio Emission from Solar Flares. *Annual Review of Astronomy and Astrophysics , 36*, 131.

Benz, A., & Grigis, P. (2002). Microflares and hot component in solar active regions. *Solar Physics , 210*, 431.

Benz, A., & Krucker, S. (1998). Heating Events in the Quiet Solar Corona. *Solar Physics , 182*, 349.

Benz, A., & Wentzel, D. (1981). Coronal evolution and solar type I radio bursts - an ion-acoustic wave model. *Astronomy and Astrophysics , 94*, 100.

Christe, S., Hannah, I., Krucker, S., McTiernan, J., & Lin, R. (2008). RHESSI Microflare Statistics. I. Flare-Finding and Frequency Distributions. *The Astrophysical Journal , 677*, 1385.

Christe, S., Krucker, S., & Lin, R. (2008). Hard X-Rays Associated with Type III Radio Bursts. *The Astrophysical Journal , 680*, L149.

Crosby, N., Aschwanden, M., & Dennis, B. (1993). Frequency distributions and correlations of solar X-ray flare parameters. *Solar Physics (ISSN 0038-0938) , 143*, 275.

Emslie, A., Dennis, B., Holman, G., & Hudson, H. (2005). Refinements to flare energy estimates: A followup to ``Energy partition in two solar flare/CME events'' by A. G. Emslie et al. *Journal of Geophysical Research , 110*, 11103.

Hannah, I., Christe, S., Krucker, S., Hurford, G., Hudson, H., & Lin, R. (2008). RHESSI Microflare Statistics. II. X-Ray Imaging, Spectroscopy, and Energy Distributions. *The Astrophysical Journal , 677*, 704.

Hurford, G., Schmahl, E., Schwartz, R., Conway, A., Aschwanden, M., Csillaghy, A., et al. (2002). The RHESSI Imaging Concept. *Solar Physics , 210*, 61.

Klein, K.-L., Schwartz, R., McTiernan, J., Trottet, G., Klassen, A., & Lecacheux, A. (2003). An upper limit of the number and energy of electrons accelerated at an extended coronal shock wave. *Astronomy and Astrophysics , 409*, 317.

Krucker, S., Benz, A., Bastian, T., & Acton, L. (1997). X-Ray Network Flares of the Quiet Sun. *Astrophysical Journal v.488 , 488*, 499.

Krucker, S., Christe, S., Glesener, L., Mcbride, S., Turin, P., Glaser, D., et al. (2009). The Focusing Optics X-ray Solar Imager (FOXSI). *Optics for EUV , 7437*, 4.

Krucker, S., Saint-Hilaire, P., Christe, S., White, S., Chavier, A., Bale, S., et al. (2008). Coronal Hard X-Ray Emission Associated with Radio Type III Bursts. *The Astrophysical Journal , 681*, 644.

Lin, R., & Hudson, H. (1976). Non-thermal processes in large solar flares. *Solar Physics , 50*, 153.

Lin, R., Evans, L., & Fainberg, J. (1973). Simultaneous Observations of Fast Solar Electrons and Type III Radio Burst Emission Near 1 AU. *Astrophysical Letters , 14*, 191.

Parker, E. (1988). Nanoflares and the solar X-ray corona. *Astrophysical Journal , 330*, 474.

Parnell, C., & Jupp, P. (2000). Statistical Analysis of the Energy Distribution of Nanoflares in the Quiet Sun. *The Astrophysical Journal , 529*, 554.

Ramsey, B., Alexander, C., Apple, J., Benson, C., Dietz, K., Elsner, R., et al. (2002). First Images from HERO, a Hard X-Ray Focusing Telescope. *The Astrophysical Journal , 568*, 432.
Saint-Hilaire, P., Krucker, S., Christe, S., & Lin, R. (2009). The X-ray Detectability of Electron Beams Escaping from the Sun. *The Astrophysical Journal , 696*, 941.

White Papers Cited

- Bookbinder et al. 2010 "The Solar Spectroscopy Explorer Mission"
- Golub et al. 2010 "RAM: The Reconnection and Microscale Mission"
- Lin et al. 2010 "Solar Eruptive Events (SEE) 2020 Mission Concept"
- Smith et al. 2010 "The high-energy Sun at high sensitivity: a NuSTAR Solar Guest Investigation Program"

Heliophysics Instrument and Technology Development Program (HITDP)

A white paper for the 2010 Solar and Space Physics Decadal Survey

Principal Author: E. R. Christian (NASA/GSFC)

Co-Authors (alphabetical): E. E. Chollet (Caltech), C. M. S. Cohen (Caltech), J. F. Cooper (NASA/GSFC), J. M. Davila (NASA/GSFC), G. A. de Nolfo (NASA/GSFC), M. I. Desai (SwRI), H. A. Elliott (SwRI), J. Giacalone (U. Arizona), G. C. Ho (JHU/APL), D. Intrilligator (Carmel Res. Ctr.), S. G. Kanekal (NASA/GSFC), L. M. Kistler (UNH), H. Kucharek (UNH), A. W. Labrador (Caltech), L. J. Lanzerotti (NJIT), S. T. Lepri (U. Mich.), R. A. Leske (Caltech), R. J. Macdowall (NASA/GSFC), D. J. McComas (SwRI), E. Möbius (UNH), T. E. Moore (NASA/GSFC), N. A. Murphy (JPL), J. M. Ryan (UNH), E. C. Sittler (NASA/GSFC), T. T. von Rosenvinge (NASA/GSFC), R. F. Wimmer-Schweingruber (U. Kiel)

Introduction

This white paper will put forth the case for a dedicated Heliophysics Instrument and Technology Development Program (HITDP). Given the potential commonality in the required tools and methods across Heliophysics sub-disciplines, this program would be implemented as a single separate program element under the Research Opportunities in Space and Earth Science (ROSES) solicitation. Having a HITDP will re-invigorate hardware development in the Heliophysics community, provide a pathway for new technology to be infused into sub-orbital, Explorer, and larger missions, help develop the next generation of Heliophysics instrument scientists, and help ensure a healthy science mission program for many years to come. HITDP will be invaluable for increasing the Technical Readiness Level (TRL) level of instruments prior to when they are proposed for missions, thus decreasing mission risk, which decreases cost and, in the long run, pays for the program.

This white paper is applicable to all three Decadal Survey study panels.

Discussion

In NASA's Science Mission Directorate (SMD), the Heliophysics Division is the **only** science division that lacks any separate instrument development program as part of the Research Opportunities in Space and Earth Sciences (ROSES). The Astrophysics Division has the Strategic Astrophysics Technology Program (SAT), Planetary Science has the Astrobiology Science and Technology for Instrument Development (ASTID), the Planetary Instrument

Definition and Development Program (PIDDP), the Mars Instrument Development Project (MIDP), and the Mars Technology Program (MTP). The Earth Science Division has an Instrument Incubator Program and an Advanced Component Technology program. Because the Heliophysics Division does not always break out instrument development spending separately it is a little difficult to determine, but it appears that the Heliophysics Division has the lowest ratio of instrument development funding to total budget in SMD.

Heliophysics has a Low Cost Access to Space (LCAS) segment of its R&A program, but this is for higher TRL instruments that are ready for large-scale sub-orbital missions. And there are instruments, especially in heliospheric science, that cannot be tested with sub-orbital because they need very pure vacuum and require outgassing time. Instruments at a low TRL can apply to the Supporting Research and Technology Program in either Geospace (GEO), or Solar and Heliospheric (S&H) science. The Geospace program calls out Instrument Development as a separate funding portion of the SR&T program, and S&H mentions it, but lumps the funding in with all other research. One of the many issues with the current system is that many instruments, such as those that measure particles and fields, are useful in both GEO and S&H mission, but the current selection method forces a bifurcation that can be seen across too much of the community. Also, the Living With a Star Targeted Research and Technology rarely actually targets technology, leaving space weather instrumentation with few opportunities.

Over the last few decades, increased launch and mission costs have reduced the launch rate for SMD missions. This has reduced the number of places in the country that are capable of supporting a flight instrument development. The remaining hardware centers are fighting repeated battles to keep the right skill mix of technicians and engineers employed, as the low launch rate requires effort that is “bursty”, with time periods that require a large amount of effort, but other times when work and funds can be lacking. Scientists are less affected by this problem because data analysis can help to smooth out this rough hardware funding profile. Even the locations that have been able to keep a steady flow of work are often only one-person-deep in some skills, which makes it harder to deal with problems that can occur, increases risk to the instrument delivery schedule, and all but eliminates the organizations’ ability to infuse new talent into instrument and technology development. A dedicated instrument development program can assist by keeping skilled personnel at work, designing the next generation of instruments when flight projects are not at their peak work load. This would help to smooth the funding flow, support highly productive groups with the required skills, and provide needed opportunities for students to get involved in instrument development. In other words, such a program would provide triple benefit, by promoting innovative instrument development, providing crucially needed means to keep essential engineering talent during inevitable valleys between consecutive missions, and promoting the training of new engineering talent.

HITDP will reduce risk on flight missions, by enabling proposals for new instruments at a higher TRL. More early effort in the design, construction, and testing of instruments will clearly reduce risk in the flight development. HTIDP also allows for development of new technology that might

be deemed as high risk and therefore un-proposable without this program. HITDP also decreases the risk and increases the effectiveness of the LCAS program. Delays in LCAS first launches are often due to emerging issue with the instrument development. It would be better to start the effort and incur the cost of a first sub-orbital flight only after the instrument is already designed and tested, thus allowing more sub-orbital opportunities.

Many of the Heliophysics flight instruments recently developed or currently under development are the result of incremental improvements over previous flight instruments. Indeed, there is often a complaint from outside of Heliophysics that a lot of what we propose to do scientifically comes down to just "more of the same". The health and future of Heliophysics research requires that we constantly push the boundaries of what we can observe, and HITDP can help develop new instrumentation that will give us large jumps in measurement capability instead of just incremental improvements.

In addition, HITDP can be used to provide training for future experimentalists. With the few SMD flight programs these days, the number of knowledgeable experimentalists that are involved in hardware development have gradually decreased over time. In the current risk adverse environment, flight programs often cannot provide the right opportunity to teach the next generation experimentalist. HITDP can provide ample opportunity for the community to propose for low TRL instruments, which is a perfect training ground for the next generation experimentalists.

HITDP can also maintain the specialized test infrastructure that is needed for many Heliophysics instruments, such as accelerators, thermal-vacuum test chambers, etc.

Cost and Implementation

At the simplest level, the HITDP could be implemented just by combining the existing funding already set aside for instrument development in the current Heliophysics Geospace and Solar and Heliospheric SR&T programs into a separate ROSES element. The Geospace program has approximately $300K available for new Instrument Development grants in ROSES FY10, and if we assume three year grants and a roughly equivalent amount coming out of the Solar and Heliospheric SR&T program, that gives a $1.8 M program. But this is less than a third of a percent of the Heliophysics Division budget, is smaller than the amount spent on Data Environment improvements, and would only serve to maintain the status quo. Therefore, we recommend that this budget be increased to at least $1M per year of new HITDP grants and preferably more. In addition, co-funding with NASA's new technology development infrastructure should be explored (NASA technologists have often seen Heliophysics as a poor place to invest, due to the lack of potential technology funding within the Heliophysics program). It is understood that this may be a "zero sum" exercise, and that there is probably no significant increase in the Heliophysics Division budget. However, it appears that the increased launch and mission costs have caused the budget to become unbalanced. Rather than decreasing the already

reduced research budget, this new HITDP money needs to come out of the mission budget. Development and delivery of flight instruments is frequently an issue for mission costs. Having a robust HITDP will help reduce risks on the construction of flight instruments and, in the long run, pay for itself.

Conclusion

The Heliophysics Division has clearly been investing very little in technology development, and this puts the future of the field at risk. In order to maintain an active and innovative hardware development community, to reduce mission risk, and to develop the new instruments that will acquire the data that the community needs to further its science goals, Heliophysics should increase the percentage of its budget that goes to new instrument concepts. The easiest and most obvious way to do this is to formally designate a Heliophysics Instrument and Technology Development Program as a ROSES element. This will also help align the division with the other SMD science divisions, all of which have assigned a higher percentage of their budget to technology because they see the utility of instrument development.

Laboratory Solar Physics from Molecular to Highly-Charged Ions: Meeting Future Space Observations of the Solar Plasma and Solar Wind

Dr. Ara Chutjian
Senior Research Scientist and Supervisor, Atomic and Molecular Physics Group
Jet Propulsion Laboratory/Caltech
Visiting Faculty Associate, California Institute of Technology USA
ara.chutjian@jpl.nasa.gov, 818.354.7012 (v), 818.393.1899 (f)

Prof. Nigel Adams
Professor of Chemistry, University of Georgia, Atlanta GA USA

Dr. John A. MacAskill
Atomic and Molecular Physics Group, Jet Propulsion Laboratory/Caltech USA

Dr. Stojan M. Madzunkov
Atomic and Molecular Physics Group, Jet Propulsion Laboratory/Caltech USA

Prof. B. Vincent McKoy
Professor of Chemistry, California Institute of Technology USA

Prof. Horst Schmidt-Böcking
Institut fűr Kernphysik der Johann Wolfgang-Goethe Universität
Frankfurt am Main, Germany

Prof. Reinhold Schuch
Dept. Atomic Physics, Stockholm University, Stockholm, Sweden

Dr. Jurij Simcic
Atomic and Molecular Physics Group, Jet Propulsion Laboratory/Caltech USA

Dr. Kurt Stiebing
Institut fűr Kernphysik der Johann Wolfgang-Goethe Universität
Frankfurt am Main, Germany

Prof. Swaraj Tayal
Dept. Physics, Clark Atlanta Univ., Atlanta GA 30314 USA

ABSTRACT

There has been an astounding return of solar observations of solar-wind velocity and composition, plasma spectroscopy, and plasma imaging from, for example, the ACE, CHIPS, IBEX, RHESSI, SDO, SOHO, STEREO missions. The wide range of planned and studied missions such as IRIS, Solar Probe Plus, and Space Environment Testbeds will almost certainly ensure a continuation of these high-quality measurements. In addition, the solar data will present new challenges for laboratory physics to provide the needed collisional reaction channels and cross sections to help in the plasma simulations. For several decades JPL has been in the forefront of providing such data. Described herein is the use of an electrostatic storage ring (ESR) coupled to a variety of ion sources to generate and store positive and negative molecular ions, and highly-charged positive ions. Studies of collision phenomena with the

relevant projectile and ion target provide insight on plasma properties of the Sun and its interactions – via photons and the solar wind (SW) – with solar-system comets, moons, planetary atmospheres, and the interplanetary medium. A summary is given of the capabilities attained by combining an ESR with a molecular-ion source, and with an electron beam ion trap for highly-charged ions (HCIs). The ESR enables one to measure molecular phenomena such as direct and dissociative ionization, direct and dissociative recombination, and lifetimes of negative ions. One can generate fast H-atoms to simulate solar-atom interactions with planetary atmospheres. For HCIs, one can study collisional excitation, direct and indirect ionization, direct and dielectronic recombination, and excited-state lifetimes in the 10^{-2} to 10^{-9} sec range. Because there is no trapping magnetic field, Zeeman mixing of excited levels is obviated; and since the ions circulate for ≈10-100 sec, the ion and molecular target ions are in their ground states.

INTRODUCTION

The advances in solar physics through ground-based spectrometer observations, and through NASA's, JAXA's and ESA's flight spectrometers and particle detectors, have extended our view of the Sun's plasma to the EUV and X-ray regions of the electromagnetic spectrum; and our understanding of the Sun's SW as to detailed composition and extent . A recent example of the rich atomic HCI emission spectrum from CHIPS is given in Fig. 1 [1]. Here one clearly sees emission lines in highly-charged Fe and Ni ions. Excitation of these lines is almost entirely through electron collisions. And hence absolute collision cross sections – obtained from experiment and benchmarked theory – are required to extract electron density, electron temperature, and ion density from the spacecraft data. In addition understanding the SW, coronal mass ejections, magnetic field reconnection, and the solar corona, chromosphere, and photosphere require a range of instruments, such as found in SOHO. Shown in Fig. 2 are results from the SOHO Mass Time-of-Flight (MTOF) analyzer that detects mass and abundance of the solar-wind species indicated, as well as of the isotopes of ten elements from Ne to Ni [2]. The SW is responsible for a wide range of phenomena. In particular, it gives rise to X-ray emission at sun-approaching comets, and at planetary atmospheres through. Charge exchange occurs in collisions of the solar-wind HCIs with the neutral gases of the target atmosphere. X-rays are produced by relaxation of the hign-*n*, "hollow ion" as it relaxes to its ground state.

Understanding the plasma properties of the Sun is intimately connected to the underlying atomic and molecular collision physics. For HCIs, one requires cross sections for excitation, direct and dielectronic recombination, single- and multiple-charge exchange (in HCI-neutral collisions), HCI level lifetimes (nanosecond to millisecond range), and *f*-values. A review of these processes and needs may be found in Refs. [3, 4]. For the effects of the SW on molecular targets (*e.g.*, in the planetary ionosphere) there is an ongoing need for accurate cross sections and rate constants for direct and dissociative

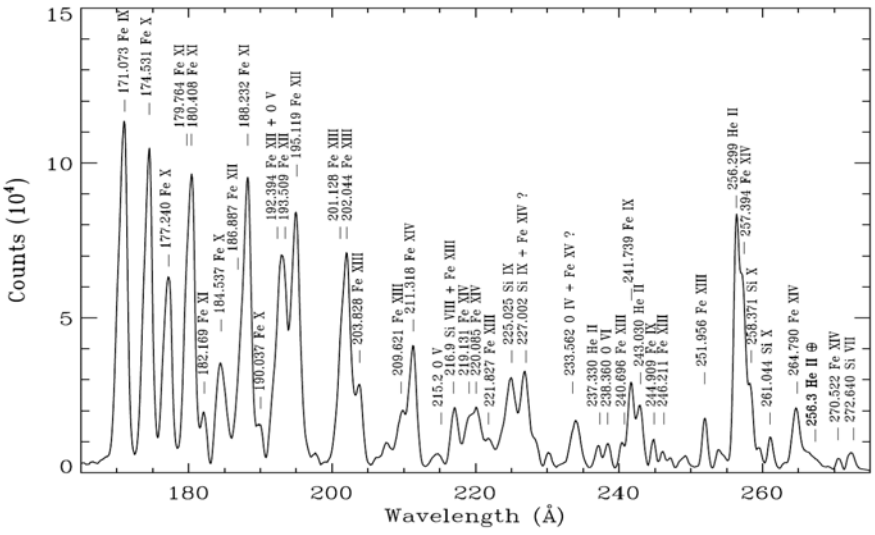


Figure 1. EUV spectrum of the full solar disk from the Cosmic Hot Interstellar Spectrometer (CHIPS) [1].

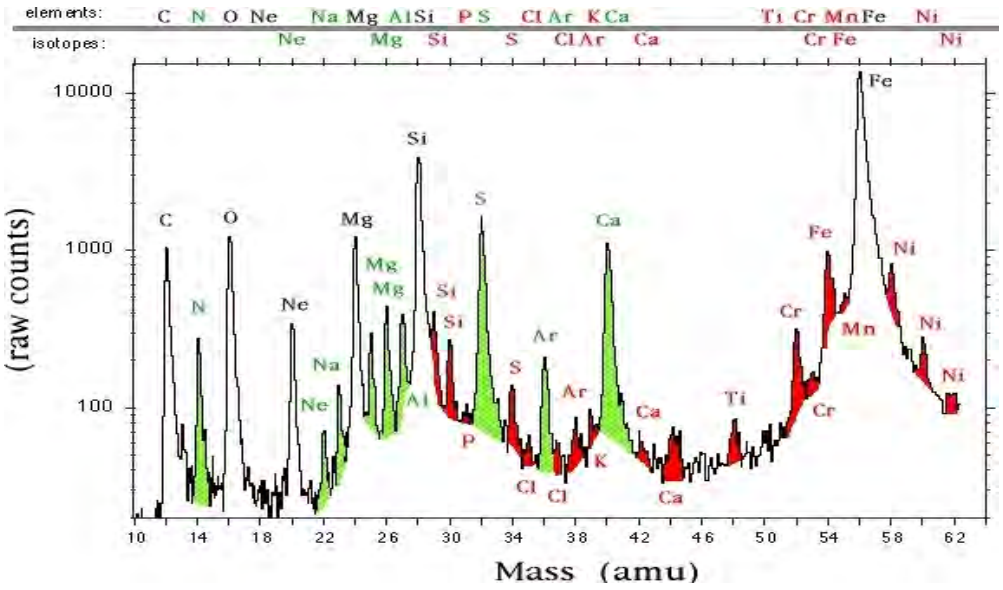


Figure 2. Solar wind elements and isotopes observed by the SOHO/CELIAS/MTOF sensor 16-19 Feb. 1996. [2]

ionization, direct and dissociative recombination, and molecular fragmentation.

Since one cannot measure *every* required cross section or lifetime, benchmarking of theoretical results is an important aspect of this work. As an example, JPL results of absolute e-Fe^{13+} excitation cross sections have helped resolve the so-called "Iron Conundrum " (see Fig 4a and the discussion below).

EXPANDING THE CAPABILITIES OF LABORATORY SOLAR PHYSICS

The extension of space observations into the infrared and X-ray regions of the spectrum has revealed a rich population of new molecules and highly-charged ions. In order to address the current and anticipated return of space data, one requires a concomitant expansion of experimental methods and facilities for measuring, in molecules, ionization and recombination cross sections, and dissociation pathways; and in HCIs collision strengths, lifetimes, charge-exchange cross sections, ionization cross sections, and direct and dielectronic recombination cross sections. Many of these phenomena (cross sections and lifetimes) are already being measured. Use is made of the electron energy-loss method (developed at JPL) with merged beams of electrons and HCIs [5]. Single and multiple charge-exchange cross sections are measured using a collision gas cell and retarding potential difference analysis of the charge-exchanged HCI ions [6]. Metastable lifetimes are measured in a Kingdon ion trap [7]. A photograph of the JPL facility with its three separate beam lines is shown in Fig. 3.

Shown in Figures 4a-4c are representative results from each of the three beam lines. Given in Fig. 4a are absolute experimental and theoretical excitation cross sections for the Fe^{13+} coronal green line, by which the so-called "Iron Conundrum" was resolved [5]: the under-abundance of Fe density relative to Mg, Si, Ca, *etc.* in Seyfert Galaxies compared to its density in the Sun was explained by a too-large collision strength for the Fe^{13+} $^2P^o_{1/2} \rightarrow {}^2P^o_{3/2}$ fine-structure transition calculated in an 18-State R-Matrix theory, relative to that of the more accurate 135-State Breit-Pauli R-Matrix theory [5]. It is clear that the development of accurate (10% level) theoretical approaches to the calculation of solar-physics phenomena is essential to meeting the present and future, high-quality space observations. Shown in Fig. 4b is the lifetime of the coronal green line transition, as measured in the Kingdon ion trap beam line [7]. And shown in Fig. 4c are the modeling results of Ref. [8] for X-ray emission from the comet McNaught-Hartley in which JPL charge-exchange cross sections were used for SW-comet interaction. Recent results of the interaction of solar-wind $Fe^{(5\text{-}13)+}$ ions with cometary and planetary CO are shown in Fig. 5a, and with H_2O in Fig. 5b. Both measurements have been combined with nCTMC calculations and radiative auger and cascade estimates to give good experimental-theoretical agreements [9,10].

To address the future space-mission challenges, several enhancements are needed to the JPL HCI facility: **(1)** the present *Caprice*-type electron cyclotron resonance ion source can produce charge states in

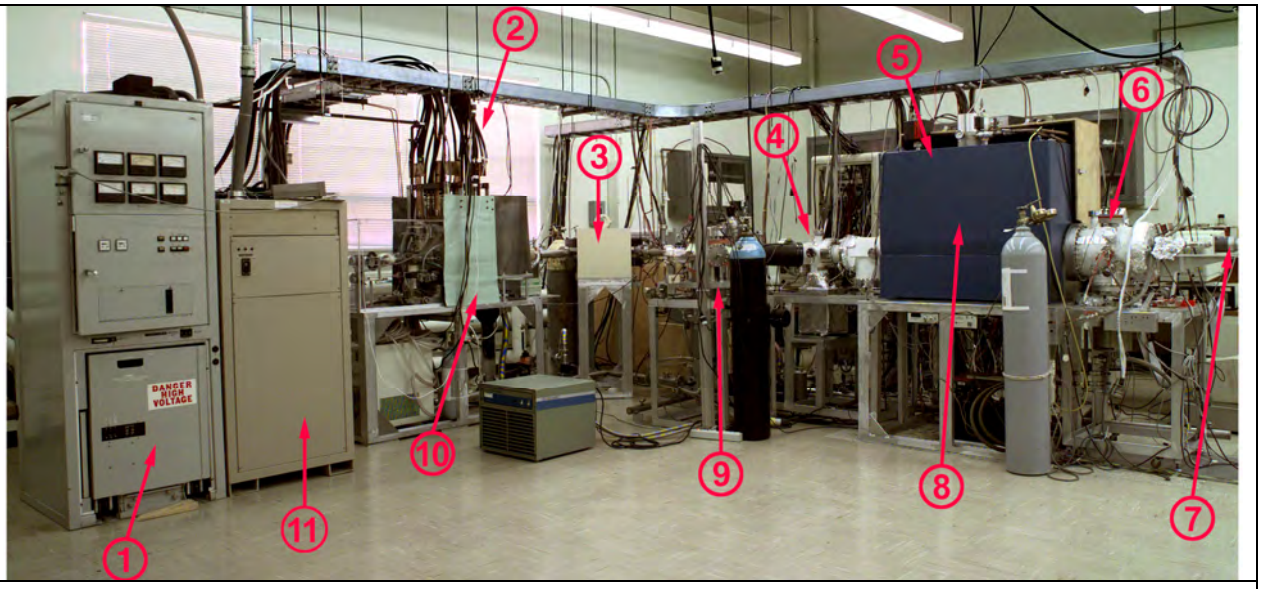


Figure 3: The JPL Highly-Charged Ion Facility. The numbered sections are, ① high-power Ku band amplifier, ② cooling lines for the HCI source magnets, ③ HCI charge/mass selection magnet, ④ ***Y***-switcher for directing the HCI beam into one of three beam lines, ⑤ solenoidal magnet for merging electron and HCI beams, ⑥ vacuum manifold with electrical feedthroughs, ⑦ stepper motor for sweeping beams profiles, ⑧ merged-beams chamber for absolute excitation cross sections, ⑨ Kingdon ion trap for HCI lifetimes, ⑩ *Caprice*-type HCI source with lead shielding, **(11)** 1000-A supply for the *Caprice* solenoidal magnets. The proposed ESR will be accommodated in the open foreground area.

Fe only to about Fe^{15+}, whereas one requires beams of up to Fe^{25+} ions for measuring cross sections applicable to energetic solar-flare emissions and coronal mass ejections. Other relevant HCIs include $Mg^{8\text{-}11+}$, $Si^{5\text{-}13+}$, $Ca^{10\text{-}19+}$, and $Fe^{15\text{-}23+}$. **(2)** Since ions in the SW are in their ground state, one would need

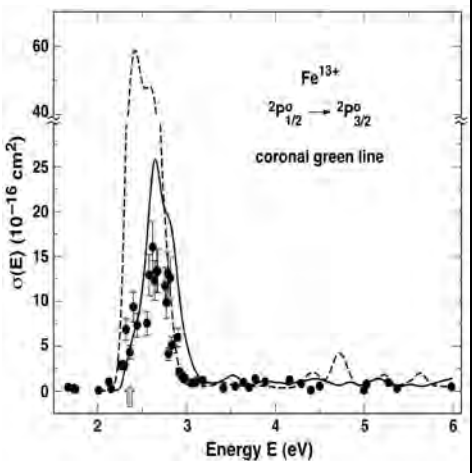


Figure 4a: The "Iron Conundrum" resolved: comparison of JPL experiments (●) and an accurate theory (—) for the Fe^{13+} coronal green line cross section [5]. A less-accurate theory (----) leads to a too-low Fe stellar abundance.

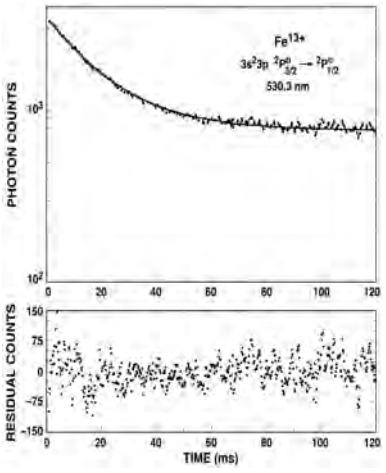


Figure 4b: Photon decay rate for the Fe^{13+} coronal green-line transition $3s^2 3p^2\ ^2P^o_{3/2} \rightarrow\ ^2P^o_{1/2}$ at 530.3 nm [7]. The measured lifetime is 17.0±0.2 ms.

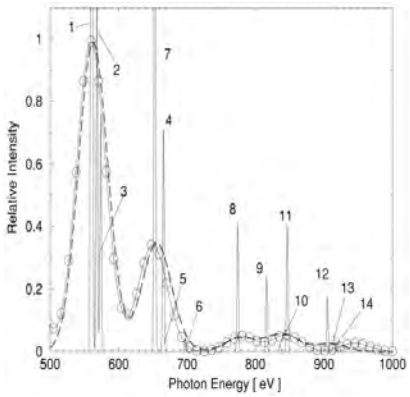


Figure 4c: Comparison of observation (o) and fit (----) to the X-ray emission spectrum from Comet McNaught-Hartley using JPL charge-exchange cross sections [8].

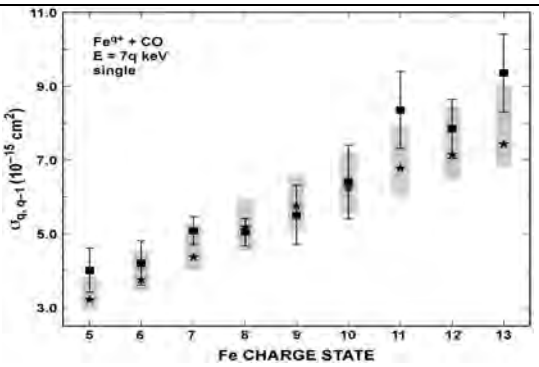


Figure 5a. Results of the nCTMC calculation, with radiative/Auger cascades, for *CE* in Fe^{5-13+} collisions with CO (⋆) compared with the present experiments (■) [9]. Shown are the single *CE* results, with the shaded error limits representing the maximum estimated uncertainty from cascades.

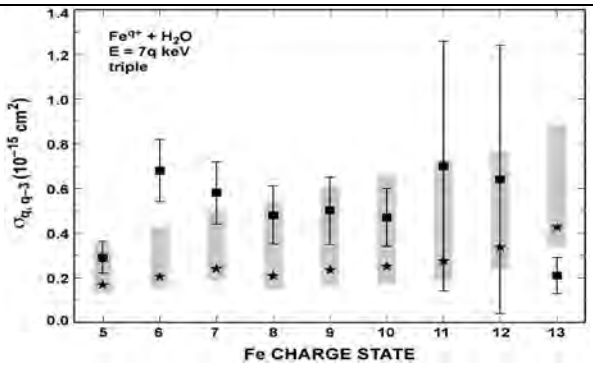


Figure 5b. Results of the *n*CTMC calculation, with radiative/Auger cascade, for *CE* in Fe^{5-13+} collisions with H_2O (⋆) compared with the present experiments (■) [10]. Shown are the triple *CE* results with the shaded error limits representing the maximum estimated uncertainty from cascades.

an experimental beam in its ground state, such as in the ESR where all metastable levels have decayed during their long circulation time. **(3)** One would like to have a means of swapping ion sources to generate negative ions (use a hollow-cathode glow discharge), positive ions (use a Nielsen-type source), or large clusters and biomolecules (use electrospray ionization). Coupling the ring to positive- and negative-ion molecular sources allows one to access a broad range of molecular targets. Ions of planetary and comet interest include CH^+, H_3^+, H_3O^+, CN^-, CO^-, and C_4H^-. **(4)** One should have a means of injecting large molecules, clusters, and biomolecules into the ESR. using electrospray (ES) ionization methods [11], or a simpler, micromachined ES source operating in JPL's MicroDevices Laboratory (P. A. Willis, unpublished results). **(5)** The system should have a straightforward means of studying direct and indirect electron ionization and dissociation phenomena for both molecular and HCI targets. A holistic approach to the issues in **(1)-(5)** is to design and build the ESR coupled to an electron beam ion trap, with utility port to accommodate sources for the low-charged molecular-ions. One example of a tabletop ring is ELISA (**E**lectrostatic **I**on **S**torage **R**ing, **A**arhus) built at Aarhus University in Denmark [11-13]. It has served as the model for rings at Tokyo Metropolitan University, and KEK, Tsukuba. Another design of the ESR stems from the group with Schmidt-Böcking and Stiebing at Goethe Universität [14]. We have selected the latter design because of its size, cost of construction, and provisions for a variety of beam diagnostics and experimental ports. A photograph of the Goethe Universität ESR ("***FIRE***") is shown in Fig. 6, and a brief discussion of its design given in the caption. JPL and Goethe Universität will collaborate on the construction of ***FIRE*** at JPL. This will result in considerable savings of time and labor at JPL, and provide excellent collaboration and researcher-exchange opportunities for both groups.

Generation of HCIs is considered a critical part of understanding the solar plasma, CME's, flares, and their spectroscopy; as well as understanding the SW interaction with planetary atmospheres and comets to generate X-rays (Refs. [6, 8-10] and references therein). As such, an integral part of the infrastructure to laboratory astrophysics will be the addition of a state-of-the-art HCI source. The Refrigerated Electron Beam Ion Trap (REBIT) can generate practically any charge state of any ion, and will provide coverage for the vast majority of HCIs encountered in astrophysics. The REBIT HCI source is commercially available, and would be procured separately as part of the total system build [15].

NEEDED LABORATORY DATA AND RECOMMENDATIONS

A vast range of laboratory data is needed to understand the plasma environment of solar, stellar, and interstellar objects. In addition to providing cross sections to the simulations community accurate, measured absolute cross sections are critical to benchmarking results of calculations. Not all cross

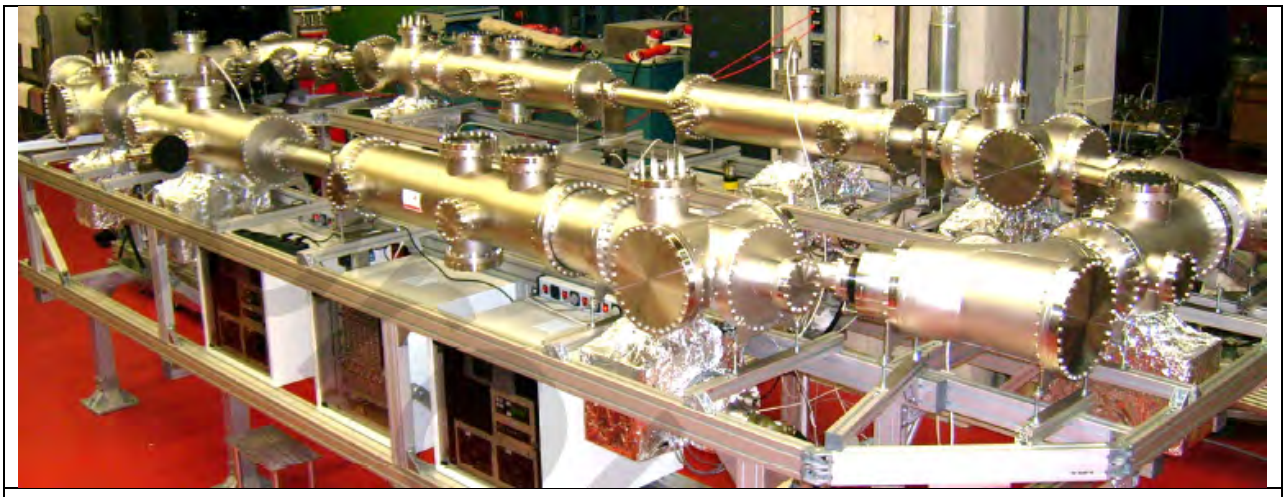


Figure 6. The Frankfurt Electrostatic Storage Ring (*FIRE*) Proposed to be Built at JPL. The ESR has four modular vacuum chambers plus connecting tubes and flanges. The chambers are easily mounted, replaced and interchanged. Two corner regions of the ring are used to carry out 0°-spectroscopy with high photon acceptance angle. Two straight sections are areas for beam diagnostics and manipulation. The experimental areas are separated from the residual ring vacuum by four bakeable UHV valves. Pumping is *via* eight ion pumps and two backing turbopumps. UHV conditions (10^{-12} torr) enable ion storage times of up to 10^{3} sec, with minimal loss due to residual-gas charge exchange.
The Charged-Particle Optical Lattice: The optical lattice utilizes electrostatic quadrupoles, parallel plate deflectors (PPD) and cylinder deflectors (CD). The ring consists of four optical sections of mirror symmetry. Each section includes a PPD, CD, quadrupole doublet (Q1-Q2), singlet (Q3), and one half of a quadrupole triplet (Q4-Q5-Q4). For the vertical orbit correction, parallel deflectors are integrated into the doublets. In order to achieve regions with high ion densities an additional Q3 quadrupole as close as possible to the CD systems is introduced. One full 180° bend is formed by a sequence of seven lens elements: 15°-PPD, Q3, 75°-CD, triplet, 75°-CD, Q3, and 15°-PPD.

sections, lifetimes, branching fractions, collision rates, *etc.* can be measured, and hence the most accurate theories must be used to calculate missing or difficult-to-measure data.

The combination of the ***FIRE*** ring with the REBIT (HCI) and ES (molecular anions and cations) sources will provide a veritable universe of experimental capabilities. Some of these classes of experiments are given in Table 1. The ESR and REBIT combination provides measurement capabilities in which there is no "wavelength barrier" or "ion barrier"! Using the JPL electron energy-loss technique [3-5] transitions from the infrared to the hard X-ray regions are accessed. Since the ESR is completely electrostatic, trapping depends *only* on target energy and *not* mass. Species ranging from hydrogen-like HCIs, to diatomic molecules, to large biomolecules will be generated and trapped!

COST AND SCHEDULE

Design, fabrication, assembly and test of the system are estimated at $1.3M (FY11 dollars) to cover 36 months. An estimated budget for the ESR-REBIT-ES facility in terms of combined, burdened, labor and procurements in each of the three years is as follows: **Year 1**: $365k, **Year 2**: $650k, and **Year 3**: $290k. Included in Year 1 is travel for personnel between JPL and Goethe Universität (***FIRE***), Aarhus (ELISA) and Stockholm (DESIREE) [16] to exchange ideas and broaden collaborations.

The proposed JPL system would be the first ESR in the United States. It would greatly expand JPL's focus on understanding solar, stellar, and planetary physics in the laboratory. The ***FIRE*** ring will add to the *scientific infrastructure* of JPL. Physicists are attracted by challenging research. It is our direct flight experience (from our mass spectrometer presently working aboard the International Space Station) that the ESR physics will attract scientists who will *also* become critical contributors and diagnosticians to JPL's flight instruments and missions!

The proposed research team provides expertise in experimental negative- and positive-ion molecular chemistry (Adams), theoretical and computational quantum chemistry (McKoy), *R*-Matrix theory and calculation of collision cross sections (Tayal), charged-particle optics design (Chutjian, MacAskill, Madzunkov, Schmidt-Böcking, Simcic, Stiebing), ion traps (Chutjian, MacAskill, Madzunkov), use of HCIs in experiments - including ion sources and UHV techniques (Chutjian, MacAskill, Madzunkov, Schmidt-Böcking, Schuch, Simcic); together with experts at the ***FIRE*** ring at Frankfurt (Schmidt-Böcking, Stiebing) and the DESIREE ring in Stockholm (Schuch).

The JPL Atomic and Molecular Group has discussed in detail the intent and ideas herein with its supporting NASA sponsors. As a courtesy, this document has also been provided to the cognizant Program Managers. It is clear that the capabilities listed in Table 1 are congruent with NASA's broad range of space missions. They also resonate with the recent NRC Decadal Survey Committee Report [17] to nurture and expand programs in laboratory collision physics to provide required atomic and molecular physic data for advancing a broad range of space-plasma simulations.

ACKNOWLEDGMENT

The work described in this report was performed at the Jet Propulsion Laboratory/California Institute of Technology, under contract with the National Aeronautics and Space Administration. The cost information contained in this document is of a budgetary and planning nature and is intended for informational purposes only. It does not constitute a commitment on the part of JPL and/or Caltech.
© 2009. All rights reserved.

Table 1. New Experimental Capabilities Enabled by the ESR-REBIT-ES System

• ***measurement of absolute electron-impact cross sections at energies to 5× threshold for up to helium- and hydrogen-like HCIs*** • ***expanded capability, to understand heavy ion ejection in solar flares and energetic CMEs***	• ***measurements of lifetimes in the 10^{-2} to 10^{-9} sec range for excited HCI states*** • ***expanded capability to the present Kingdon trap, for measuring lifetimes in the broader lifetime range***
• ***absolute single and multiple charge-exchange cross sections using the ring-cooled, ground-state HCIs*** • ***present approach uses less accurate methods of measuring metastable populations by gas attenuation and translational energy spectroscopy (in progress)***	• ***measurement of positive- and negative-ion lifetimes in diatomic and polyatomic ions, in the absence of magnetic-field mixing and photodetaching, black-body radiation*** • ***new capability for measuring molecular ion lifetimes dissociation pathways, from simple diatomics to large biomolecules***
• ***absolute direct and indirect ionization cross sections of HCIs*** • ***new capability by merging the ring ions with an electron beam to measure absolute single and multiple ionization phenomena***	• ***absolute direct and dielectronic recombination cross sections in HCIs*** • ***new capability, to understand collisional ionization equilibrium in the Sun by use of a merged or inclined electron beam with the stored HCIs***
• ***accurate measurements of HCI fluorescence lifetimes (10^{-2}-10^{-9} sec) and isotope ratios by tunable laser absorption*** • ***for measuring isotope ratios, this is a new capability that overlaps our high-resolution Paul ion trap program***	• ***laser photodissociation with product distributions of prevalent ISM ions such as H_2^+, H_3O^+, CH_3^+, CN^-, CO^-, and C_4H^-*** • ***new capability for studying molecular chemistry induced by solar ions, electrons, and photons at comets and planetary atmospheres***

REFERENCES

[1] S. M. Sirk, M. Hurwitz, W. Marchant, "EUV Spectra of the Full Solar Disk: Analysis and Results of the Cosmic Hot Interstellar Plasma Spectrometer (CHIPS)," Solar Physics **26**, 287 (2010).

[2] F. Ipavich, P. Bochsler, J. Paquette, and S. Lasley, "Observations of Solar Wind Ni and Fe," in "Ten Years of SOHO and Beyond," ESA SP-617 (Eds. H. LaCoste and L. Ouwehand, 2006).

[3] A. Chutjian, "Ion Collisions in the Highly Charged Universe," Phys. Scripta **T110**, 203 (2004).

[4] J. B. Greenwood, R. J. Mawhorter, I. Čadež, J. A. Lozano, S. J. Smith, and A. Chutjian, "The Contribution of Charge Exchange to Extreme Ultraviolet and X-ray Astronomy," Phys. Scripta **T110**, 358 (2004).

[5] S. Hossain, S. S. Tayal, S. J. Smith, J. C. Raymond, and Chutjian, "Measurement and Calculation of Absolute Cross Sections for Excitation of the $3s^2 3p\ ^2P^o_{1/2}$ - $3s^2 3p\ ^2P^o_{3/2}$ Fine-Structure Transition in Fe^{13+}," Phys. Rev. A **75**, 022709 (2007).

[6] N. Djurič, S. J. Smith, J. Simcic, and A. Chutjian, "Absolute Single and Multiple Charge Exchange Cross Sections For Highly Charged C, N, and O Ions Colliding with CH_4 " Ap. J. **679**, 1661 (2008).

[7] S. J. Smith, A. Chutjian, and J. A. Lozano, "Measurement of Metastable Lifetimes for Transitions in Fe^{9+}, Fe^{10+}, and Fe^{13+}," Phys. Rev. A **72**, 062504 (2005).

[8] V. Kharchenko, M. Rigazio, A. Dalgarno, V. A. Krasnopolsky, "Charge Abundances of the Solar Wind Ions Inferred from Cometary X-Ray Spectra, Ap. J. **585**, L73 (2003).

[9] J. Simcic, D. R. Schultz, R. J. Mawhorter, A. Chutjian, I. Čadež, J. B. Greenwood, C. M. Lisse, and S. J. Smith, "Measurement and Calculation of Absolute Single and Multiple Charge Exchange Cross Sections for Fe^{q+} Ions on CO and CO_2," Phys. Rev. A **81**, 062715 (2010).

[10] J. Simcic, D. R. Schultz, J. B. Greenwood, R. J. Mawhorter, B. V. McKoy, S. J. Smith, C. Winstead, and A. Chutjian, "Measurement and Calculation of Absolute Single and Multiple Charge Exchange Cross Sections for Fe^{q+} Ions Impacting H_2O," Ap. J. **722**, 435 (2010).

[11] J. U. Andersen, *et al.*, "The Combination of an Electrospray Ion Source and an Electrostatic Storage Ring for Lifetime and Spectroscopy Experiments on Biomolecules," Rev. Sci. Instrum.**73**, 1284 (2002).

[12] S. P. Möller, "ELISA, an Electrostatic Storage Ring for Atomic Physics," Nucl. Instrum. Methods A **394**, 281 (1997).

[13] L. H. Andersen, O. Heber, O., and D. Zajfman, "Physics with Electrostatic Rings and Traps," J. Phys B. **37**, R57 (2004).

[14] C. P. Welsch, A. Schempp, and H. Schmidt-Böcking, "Design Studies of an Electrostatic Storage Ring," Hyperfine Interactions **146**, 253 (2003); C. P. Welsch, J. Ullrich, C. Gläβner, K-U. Kűhnel, A. Schempp, H. Schmidt-Böcking, and R. Dörner, Phys. Rev. Lett. – Accelerators and Beams **7**, 080101 (2004).

[15] J. W. McDonald and D. H. G. Schneider, "The Next Generation Refrigerated (Cryogenic) Electron Beam Ion Trap-Source (REBIT-S)," Nucl. Instrum. Methods B **241**, 870 (2005).

[16] H. Y. Schmidt, *et al.*, "DESIREE as a New Tool for Interstellar Ion Chemistry," Int. J. Astrobiology **7**, 205 (2008).

[17] Committee for a Decadal Survey of Astronomy and Astrophysics - National Research Council, "New Worlds, New Horizons in Astronomy and Astrophysics," available at *htttp://www.nap.edu/catalog/12951.html*. See also, "Panel Reports – New Worlds, New Horizons in Astronomy and Astrophysics, available at *www.nap.edu*.

A CubeSat Constellation to Study Magnetospheric Ultra-Low Frequency Pulsations

S. G. Claudepierre and **T. P. O'Brien**
The Aerospace Corporation
S. R. Elkington
LASP, University of Colorado at Boulder

I. Motivation and Background

This white paper addresses the need for a mission to constrain the azimuthal mode number spectrum of magnetospheric ultra-low frequency (ULF) pulsations. These pulsations are a ubiquitous feature of both in-situ and ground-based magnetospheric observations and have been useful in characterizing a wide-range of magnetospheric phenomenon, from substorm onsets to the mass density profile of the inner magnetosphere. In addition, magnetospheric ULF waves are believed to play a role in electron flux enhancements in the Earth's outer radiation belt. The relationship and interplay between the various physical mechanisms responsible for electron energization and loss in the outer zone is complicated and not entirely understood. The proposed mission will help advance the field of radiation belt physics and will lead to a more complete understanding of outer zone electron dynamics. Moreover, the accurate prediction of energetic electron (~ MeV) flux enhancements impacts society as a whole, as several satellite orbits, on which human-kind is increasingly reliant (e.g., geosynchronous, GPS), intersect this highly variable and potentially damaging region of geospace. There is mounting observational evidence that long duration, elevated Pc5 ULF wave power is good proxy for electron flux enhancements in the outer zone [e.g., *O'Brien et al.,* 2001]. This suggests a link between ULF waves and the electron radiation belts.

ULF waves with periods on the order of the electron drift period (e.g. Pc4-5 range) can violate the third adiabatic invariant of the charged particle motion and lead to stochastic diffusion in the *L*-coordinate. This process, known as radial diffusion [*Schulz and Lanzerotti,* 1974], is believed to be responsible for the observed link between magnetospheric ULF wave power and electron radiation belt flux enhancements [*Elkington et al.,* 2003]. The radial diffusion coefficient, which determines the rate of radial transport and energization, is proportional to ULF wave power at frequencies given by the drift-resonance condition, $f = mf_d$. Here, f is the interacting wave frequency, f_d is the electron drift frequency and m is the azimuthal mode number of the interacting wave. Thus, in order to quantify the role that radial diffusion plays in outer zone electron dynamics, a detailed knowledge of both the frequency spectrum and the azimuthal mode number spectrum of the waves is required. However, to determine the azimuthal mode structure up to mode number, m, requires $2m$ equatorial satellite measurements, evenly distributed in azimuth. Thus, constraining the azimuthal mode number spectrum from in-situ observations is especially difficult. Ground-based magnetometer observations can, at times, provide information on the azimuthal mode spectrum of ULF waves, but these calculations suffer from ambiguities related to aliasing and the modification of the ULF signal by the ionosphere. Thus, radiation belt modelers routinely turn to global, three-dimensional magnetohydrodynamic (MHD) simulations of the solar wind-magnetosphere interaction to characterize the azimuthal mode number spectrum of ULF waves.

Figure 1 shows the azimuthal mode number spectrum computed from Lyon-Fedder-Mobarry (LFM; *Lyon et al.,* [2004]) MHD simulations of Pc4-5 ULF pulsations (the figure is reproduced from *Claudepierre et*

al., [2008]). Panel a) shows electric field wave power (azimuthal component) on the color scale, as a function of azimuthal mode number, *m* (horizontal scale), and frequency (vertical scale). In this LFM simulation, there are two ULF wave populations generated: one population is generated near the dawn and dusk flank low-latitude boundary layer, via the Kelvin-Helmholtz (KH) instability; and the other is generated on nightside and is related to periodic fluctuations in the convective electric field. In panel a), the KH wave population is manifest as the enhanced wave power centered near 9-10 mHz; the convective variations are seen at lower frequencies (f < 3 mHz). Panel b) shows integrated wave power as a function of *m*, integrated around three frequency bands chosen to emphasize the different mode number characteristics of the two wave populations. The dashed trace is integrated around 3-15 mHz, to account for the KH wave population, while the dotted trace is integrated around 0.5-3 mHz, to account for the convective variations. The solid trace shows the integrated power for the entire frequency band considered in this study, 0.5-15 mHz. We see that not only are the frequency spectra of these two wave populations distinct (panel a)), so too are the azimuthal mode number spectra. The KH waves have peak wave power centered near $m \approx 15$, whereas the convective variations have peak power centered near $m \approx 8$. This information regarding the frequency and mode number spectra of different ULF wave populations is crucial for conducting radial

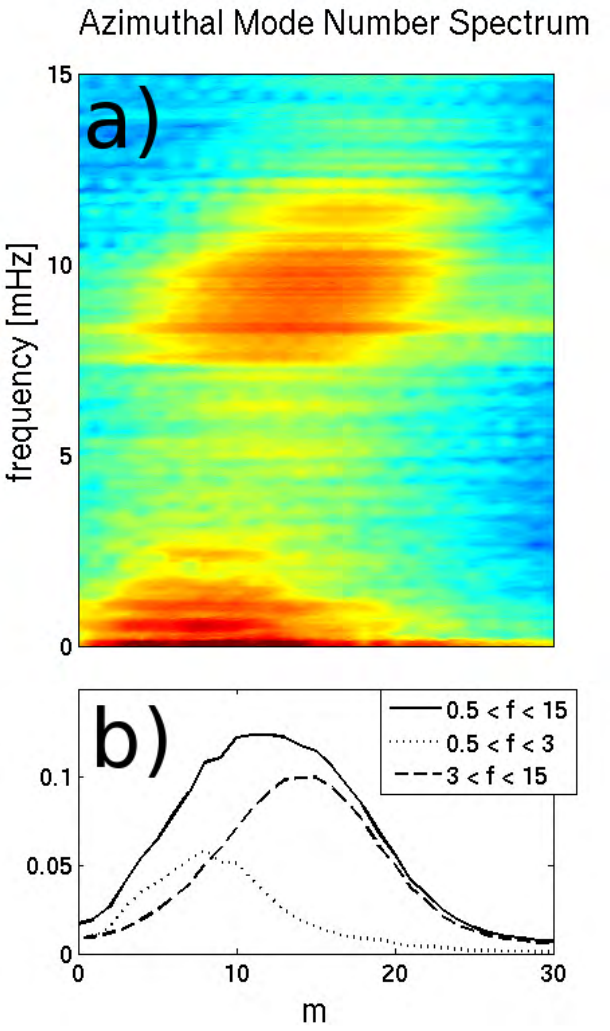


Figure 1. Azimuthal mode structure computed from an LFM simulation of ULF waves.

diffusion-type modeling studies. Without guidance from such global simulations of ULF waves, the *m* number is otherwise arbitrarily specified by radiation belt modelers. We note that *Fei et al.,* [2006] included the azimuthal mode structure derived from global MHD simulations in their radial diffusion calculations and the results showed a substantial improvement over simpler estimates of the mode structure (e.g. all wave power in a single *m* number).

While global simulations such as the LFM provide a reasonable representation of the large scale structure and global dynamics of the magnetosphere, they too suffer from their own shortcomings. For example, the LFM lacks a realistic ring current, plasmasphere and radiation belt and does not realistically account for the microphysics that leads to magnetic reconnection. All of these deficiencies can affect the frequency and mode number characteristics of simulated ULF waves. Thus, the azimuthal mode number spectra deduced from global MHD simulation results, which are frequently used in radiation belt modeling [e.g. *Elkington et al.,* 2003], must be validated. These issues highlight the pressing need

for a mission to accurately and unambiguously determine the mode number spectrum for different classes of observed Pc4-5 ULF pulsations.

II. Mission Concept

A six-to-twelve month database of continuous ULF measurements of the azimuthal mode number spectrum would provide a critical missing piece of information for radiation belt modelers. We envision a constellation of ~30 CubeSats, equipped with fluxgate magnetometers, in a circular orbit near the geomagnetic equatorial plane (e.g. geosynchronous). As noted above, a large number of sensors is required to sample the low m (< 20) portion of the azimuthal mode number spectrum that is crucial for radial diffusion studies. For example, a constellation of 30 CubeSats, evenly distributed in azimuth near the geomagnetic equatorial plane, could determine the mode number spectrum up to m = 15. A moderate time sampling rate would also be required (~30 s) to achieve sufficient sampling of the Pc4-5 frequency range (f < 20 mHz). These two sampling requirements for the temporal and spatial spectra are imposed by the drift-resonance condition given above, which implies that low frequency, low m number ULF waves will be the most effective at energizing and transporting outer zone electrons. We emphasize that both existing (e.g. THEMIS, Cluster, ST4) and forthcoming (e.g. RBSP, MMS) multi-point measurements are not azimuthally distributed, and thus, cannot determine the ULF azimuthal mode number spectrum in a consistent and reliable manner.

A fleet of in-situ magnetometers could also address a secondary science question: dynamics of large-scale magnetospheric current systems. Further, such a data source could be used to seed data assimilative models of the ring current and associated current systems during magnetic storms.

For such a large number of vehicles, a natural fit would be the CubeSat platform. Because of its small size, such a platform likely could not be deployed to geosynchronous orbit and still communicate adequately with the ground. Therefore, a CubeSat mission would likely have to be in MEO, which also resolves the issue of on-orbit disposal. A mission design study would be required to optimize the mission profile to balance the CubeSat bus capability, launch insertion, and likelihood of observing ULF waves. Based on past CubeSat costs and assumed economies of scale, the ballpark cost for the mission is estimated to be less than $10 million.

The azimuthal mode structure of ULF waves has long been a gnawing unknown in our understanding of inner magnetospheric phenomenon. With the advent of CubeSats, it appears that direct observation of that structure is within our grasp.

References

Claudepierre, S. G., S. R. Elkington, and M. Wiltberger (2008), Solar wind driving of magnetospheric ULF waves: Pulsations driven by velocity shear at the magnetopause, *J. Geophys. Res.*, *113*, A05218, doi:10.1029/2007JA012890.

Elkington, S. R., M. K. Hudson, and A. A. Chan (2003), Resonant acceleration and diffusion of outer zone electrons in an asymmetric geomagnetic field, *J. Geophys. Res., 108*(A3), 1116, doi:10.1029/2001JA009202.

Fei, Y., A. A. Chan, S. R. Elkington, and M. J. Wiltberger (2006), Radial diffusion and MHD particle simulations of relativistic electron transport by ULF waves in the September 1998 storm, *J. Geophys. Res., 111*, A12209, doi:10.1029/2005JA011211.

Lyon, J. G., J. A. Fedder, and C. M. Mobarry (2004), The Lyon–Fedder–Mobarry (LFM) global MHD magnetospheric simulation code*, J. Atmos. Sol. Terr. Phys., 66,* 1333.

O'Brien, T. P., R. L. McPherron, D. Sornette, G. D. Reeves, R. Friedel, and H. J. Singer (2001), Which magnetic storms produce relativistic electrons at geosynchronous orbit?, *J. Geophys. Res., 106*, 15,533–15,544, doi:10.1029/2001JA000052.

Schulz, M., and L. J. Lanzerotti (1974), *Physics and Chemistry in Space*, vol. 7, *Particle Diffusion in the Radiation Belts*, Springer-Verlag, New York.

Protecting Science Mission Investment: Balancing the Funding Profile for Data Analysis Programs

C.M.S. Cohen[1], R.A. Leske[1], E.R. Christian[2], G.A. de Nolfo[2],
M.I. Desai[3], D.K. Haggerty[4], L. Jian[5]
[1]California Institute of Technology, [2]NASA/Goddard Space Flight Center,
[3]Southwest Research Institute, [4]Johns Hopkins Applied Physics Laboratory,
[5]University of California, Los Angeles

Our understanding of the Sun and its affects on the inner heliosphere and the Earth has grown exponentially since the dawn of the space age. With a fleet of spacecraft (and ground-based instruments) observing the Sun, the interplanetary medium, and the Earth we are currently making measurements in more detail and with more precision than ever before possible. For the first time we are remotely observing the Sun simultaneously at multiple locations, including ones significantly removed from the Sun-Earth line. Substantial advancements in instrumentation have resulted in measurements being made at cadences and resolutions that are orders of magnitude greater than those of merely 15 years ago. The wealth and quality of data available from current instrumentation provide the opportunity for heliospheric science to move from correlative studies to in-depth examinations of the complex interplay of processes that govern the Sun-Earth system over a range of physical and temporal scales.

Effectively using these data to study complex processes requires research teams with sufficient breadth of knowledge and expertise over a range of sub-disciplines and instrumentation. As a single example, consider the attempt to understand the acceleration of solar energetic particles (SEPs). Our studies have now progressed past the point of noting the correlation with coronal mass ejections (CMEs) and flares, past comparing the characteristics of the SEP events as measured at a single location to the timing and location of the CMEs and flares observed remotely. It is currently recognized that the kinetic energy of the CME plays a large role in generating an interplanetary shock of sufficient strength to accelerate large numbers of suprathermal particles to high energies. It is acknowledged that the accompanying flare may generate SEPs which contribute measurably to the event or possibly provide a seed population to be further accelerated by an interplanetary shock. To understand the generation of an SEP event (much less to predict its occurrence and characteristics, as is the goal for Space Weather applications), studies must be made of solar processes governing the release of magnetic energy generating the flare and CME; the creation and characteristics of the seed particles in the outer corona and solar wind; the development and properties of an interplanetary shock resulting from the interaction of the CME with the solar wind as it propagates through the interplanetary medium; the acceleration process at the shock and how the evolving characteristics of the interplanetary medium, the shock, and the seed particles affect the resulting accelerated particle population. This work requires expert knowledge in remote measurements of the solar environment, the flare, and the CME; *in situ* interplanetary measurements of the solar wind, the interplanetary magnetic field, and interplanetary turbulence; remote and *in situ* particle measurements of the suprathermal and energetic populations; as well as expert theoretical knowledge of solar eruptions and interplanetary shock acceleration. Significant benefits are also to be gained from incorporating expert experience in modeling the processes involved.

The data required to advance our understanding of SEP acceleration must be obtained at as many physically different locations as possible. This and the range of particle energies

involved require careful intercalibration of separate instruments and separate spacecraft. The sophistication of the instrumentation providing the high sensitivity measurements, as well as the large volume of data (e.g., from SDO), make this intercalibration no simple task. Accurately understanding (as well as pushing) the limits of the instruments requires detailed and specific knowledge not widely available outside the instrument teams. It is clear that to make substantial progress on such complex problems a team of researchers must be large enough to provide the necessary expertise and involve the efforts of members of the specific instrument teams in a coherent and effective manner. Unfortunately this is being severely hampered by at least two specific funding issues: the current trend of significantly decreasing funds within a mission for post-launch data analysis (e.g., as has already occurred for STEREO and SDO) and the small size of non-mission-specific grants (e.g., LWS, SHP, HGIP).

Although it is critical for instrument teams to make their data publicly available, in as complete a way as possible, merely funding the generation of these data products is not sufficient to extract the full return value on the investment of building and launching the instruments/spacecraft. All data products and scientific analysis for all studies cannot be foreseen and created by the instrument teams *a priori*. Funding of the instrument teams *to do science* is essential. The expertise of the instrument teams cannot be easily found outside the teams and yet is critical to the success of studies involving multiple instruments/spacecraft. This is not to say that programs such as the Heliophysics Guest Investigator are not also needed. Scientific ideas from outside the instrument teams are key and must be supported. Unfortunately, the current reduction in mission-associated funding of scientific analysis has not been matched by increases in non-mission-specific scientific programs. Thus there is a net loss for funding data analysis.

It is also clear that non-mission-specific funding programs do not fund grants large enough to support the size team required to make significant scientific progress on complicated problems. At ~$150k/year (the size of a typical grant), a team of 5 people at 3-4 different institutions can only be funded for 1-2 months/year. At that level of commitment it is impossible to study a problem in any depth. Smaller teams will yield more time commitment but suffer for a lack of breadth in knowledge which is crucial for success. To make matters worse, the funding for such programs is also decreasing (LWS just received a substantial cut and the HGIP was not competed in 2010). This results in fewer funding opportunities which has an even larger impact on the science accomplished when mission-specific science funding is not available to pick up the slack.

The negative effects of the loss of mission-specific funding are numerous. At a minimum it causes other funding programs to be more oversubscribed, leading to many excellent proposed scientific studies going unfunded. It heavily impairs the ability of scientists to investigate phenomena without preconceived notions of cause and effect, as one must demonstrate a certain likelihood of success for a proposed study in order to be competitive in the oversubscribed funding programs. The small size of the typical non-mission-specific grant does not generally support detailed, time-consuming analysis, but rather encourages automated analysis or limiting scientific studies to utilizing canned data products which may not be best suited, because important results are expected and promised even with only 1-2 months of committed time per year. The size of these grants also forces scientists to study small portions of the problem, e.g., the resulting energetic particle characteristics and their dependence on the characteristics of the seed particle populations. However, in such complex systems, this piecemeal approach is

insufficient to understand the physical processes involved, much less to develop any kind of predictive capability.

Lack of mission-related funding also strongly affects postdocs and early career scientists who typically do not have the experience necessary to successfully compete in funding programs such as LWS, HGIP, and SHP. Traditionally these scientists are funded out of mission MO&DA funds enabling them time to gain knowledge of specific instruments or data analysis techniques. It permits them to be part of an instrument team and learn how to design instrumentation; a process critical to developing the next generation of scientists that will envision and define future missions and instruments. Scientists propose instrumentation because they find the unanswered science questions compelling. By removing the funding for the scientific research associated with the mission, the motivation for designing the instruments is compromised. Such effects on the younger generation of scientists will have lasting negative impacts on the field.

The future of our scientific community and the advancement of Heliospheric science require scientific funding through both mission- and non-mission-specific programs. The mission science funds must be maintained and protected from the cost over-runs that are commonly incurred when building sophisticated, state-of-the-art instruments. The non-mission-specific grants must be substantial enough to support teams of a size and duration sufficient to apply the needed expertise to complex problems in a careful and comprehensive manner. The funding agencies must work together to dovetail these two funding avenues effectively. Without this, our understanding of the complex Sun-Earth system will be subject to merely incremental advances and true predictive capabilities for Space Weather will not be achieved, to the detriment of our field and community.

Space Weathering Impact on Solar System Surfaces and Mission Science

John F. Cooper*, Richard E. Hartle, Edward C. Sittler Jr., Rosemary M. Killen
NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA

Steven J. Sturner
CRESST, University of Maryland College Park, College Park, Maryland, USA

Chris Paranicas, Matthew E. Hill, Abigail M. Rymer
Applied Physics Laboratory, Johns Hopkins University, Laurel, MD, USA

Paul D. Cooper
George Mason University, Fairfax, VA, USA

Dan Pascu
U.S. Naval Observatory, Washington, DC, USA (retired)

Robert E. Johnson, Timothy A. Cassidy
University of Virginia, Charlottesville, VA, USA

Thomas M. Orlando
Georgia Institute of Technology, Atlanta, GA, USA

Kurt D. Retherford
Southwest Research Institute, San Antonio, TX, USA

Nathan A. Schwadron
Boston University, Boston, MA, USA

Ralf I. Kaiser
University of Hawaii at Manoa, Honolulu, HI, USA

François Leblanc
LATMOS, Institut Pierre Simon Laplace, Université Versailles Saint Quentin/CNRS, Verrières-le-Buisson, France

Louis J. Lanzerotti
New Jersey Institute of Technology, Newark, NJ, USA

Claudia J. Alexander, Henry B. Garrett, Amanda R. Hendrix
NASA Jet Propulsion Laboratory, Pasadena, CA, USA

Wing H. Ip
Institutes of Astronomy and Space Sciences, National Central University, Taiwan

***** John F. Cooper, Code 672, NASA GSFC; Phone: 301-286-1193; John.F.Cooper@nasa.gov.

Submitted to Planetary Science Decadal Survey, NAS Space Studies Board, Sept. 11, 2009.
Submitted to Solar and Space Physics Decadal Survey, NAS Space Studies Board, Nov. 12, 2010

Abstract

Space weathering is the collection of physical processes acting to erode and chemically modify planetary surfaces directly exposed to space environments of planetary magnetospheres, the heliosphere, and the local interstellar environment of the solar system. Space weathering affects the physical and optical properties of the surfaces of planetary bodies, so understanding its specifics is critical for interpreting surface data from remote and landed measurements. For full coverage of space environmental measurements, we recommend expanded interdisciplinary cooperation between NASA's Planetary Science and Heliophysics divisions. To grow the field in the next decade and maximize impact on mission studies, we suggest a balanced mixture of laboratory measurements, modeling, and theoretical investigations in support of all missions.

Space Environments

Vast expanse of the space weathering environment interacting with solar system bodies is illustrated in Figure 1 by logarithmic horizontal scale of radial distance from the Sun to α-Centauri. Principal sources of energy for space weathering of planetary surfaces are UV photons, solar wind plasma, and energetic particles from the Sun, within a few hundred AU, and external sources of plasma and energetic particles entering into the solar system from the local interstellar environment. Beyond the realm of terrestrial planets and the asteroid belts, the solar influence significantly wanes with the decline in density of the expanding solar wind plasma and magnitude of the frozen-in magnetic field, while the interstellar influence progressively grows through interaction of interstellar neutral winds with solar ultraviolet radiation and the solar wind plasma. Across the heliospheric boundary region near 100 AU, now being explored by the two Voyager spacecraft, occurs a transition from supersonic (400 – 800 km/s) plasma flows of solar coronal expansion, the solar wind, to 26 km/s inward flow of the interstellar wind. This plasma contains both a bulk flow and thermal components associated with typical ion energies up to a few keV, and energetic components extending to far higher energies, ultimately to the full range

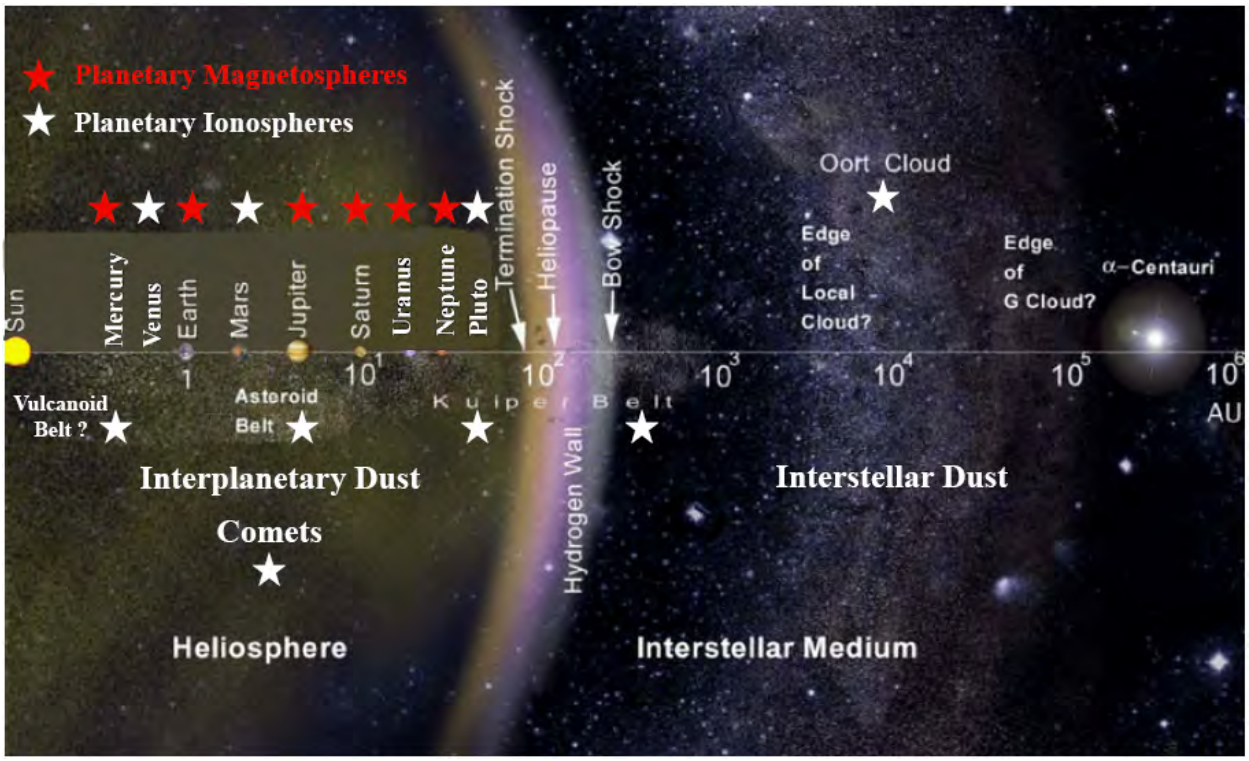


Figure 1. Logarithmic distribution of space weathering environments from the Sun to the nearest star.

of galactic cosmic ray ions easily penetrating into the heliosphere at GeV energies and higher.

The Sun contributes the innermost source of energetic particles in association with solar flare and coronal mass ejection (CME) events, the interstellar environment contributes the outermost sources, and the dynamics of the expanding and variable solar wind provide additional energy to MeV energies within the heliosphere. The solar wind termination shock, the supersonic-subsonic flow boundary crossed by both Voyagers (Stone et al., 2005, 2008), and/or the heliosheath region beyond this shock out to the heliopause, the contact boundary with interstellar plasma, *may* further accelerate plasma particles into the energetic range but this is not yet established by the Voyager measurements. Neither spacecraft detected local particle acceleration at the respective crossings, although the bulk of acceleration may be occurring elsewhere along the shock boundary (McComas and Schwadron, 2006). Other possibilities are that the heliosheath ions are energized by turbulent or reconnecting magnetic fields in the heliosheath, or that these ions originate instead from penetrating interstellar ions (Cooper et al., 2006; Cooper, 2008). As solar activity increases and then again declines within the next decade from the current minimum, the continuing Voyager measurements, supplemented by direct energetic neutral atom measurements of boundary region emissions by the Interstellar Boundary Explorer (IBEX) in earth orbit, are expected to resolve origin of the heliosheath ions and to locate the heliopause. What is already clear is that the termination shock boundary marks the transition from dominance of some space weathering effects, e.g. erosive sputtering, by the supersonic plasma flow to a broader range of effects from plasma and energetic particles at higher energies to the cosmic ray regime.

Within these expanding near-solar to heliospheric to local interstellar space environments we find the objects of primary interest to planetary science: the terrestrial planets, the asteroid belt, the gas and ice giant planets, comets, the Kuiper Belt, and finally the Oort Cloud. Aside from the first known member of the Kuiper Belt, Pluto, now officially designated as an ice dwarf planet, our direct knowledge of Kuiper Belt Objects (KBO) began with the first discovery in 1992, then followed to date by about a thousand other discoveries of such objects, including a few classified as members of the inner Oort Cloud. Presumably there are thousands more of similarly detectable size waiting to be discovered, and far more at smaller sizes. Looking back towards the Sun, there are also thousands of known asteroidal bodies, including Near Earth Objects potentially of concern for future Earth impacts, and as a remote possibility the first members of the fabled Vulcanoid Belt that might be found via increasing sensitivity of near-solar observations. At the smallest scales there are interplanetary dust grains, the source of the zodiacal light, extending down in size to nanometers or less (e.g., molecular clusters) and thought to arise from impact surface weathering of small bodies and from comet outgassing.

The red and white stars of Figure 1 denote the distinctly different space environments of solar system bodies with and without internally generated magnetic fields. Except for the highest energy cosmic rays and their atmospheric interaction products, the direct effects of space weathering do not extend to the solid surfaces of Venus, Earth, and Mars. While the planetary magnetospheres (red stars) substantially deflect interplanetary plasmas and energetic particles away from the atmospheres and underlying surfaces, even an ionospheric (white star) interaction arising from ionization of a thin atmosphere, or surface-bound exosphere, can significantly impede or totally inhibit access of space plasma to otherwise exposed surfaces. The surface of Mars is notably oxidized by solar ultraviolet irradiation and to a lesser extent from high energy (> 100 MeV) cosmic rays and solar energetic particles, while medium-energy (> 1 MeV) energetic ions can easily penetrate Pluto’s microbar-pressure atmosphere to the surface. On the other hand, the acceleration of charged particles within the planetary magnetospheres, and

magnetic pickup of exospheric ions, provides additional and potentially more dominant energy sources for space weathering of surfaces exposed to those environments.

Space Weathering Effects

In remotely viewing and more closely encountering planetary objects of the solar system and beyond, we are often looking "through a glass, darkly" at objects with tenuous atmospheres and direct surface exposure to the local space environment. In each case this space weathering exposure acts via universal plasma-surface interaction processes (Johnson, 1990) to produce a thin patina of outer material covering and sometimes obscuring the endogenic materials of greatest interest for under-standing origins and interior evolution of the affected object. Examples of weathered layers obscuring pristine ones are the radiation crusts (Johnson et al., 1987) on cometary nuclei, also on Kuiper Belt and Oort Cloud Objects (Hudson et al., 2008), and the iogenic components of sulfate hydrate deposits on the trailing hemisphere of Europa. In the latter example, near-infrared spectroscopy is used for mapping of the surface distributions (McCord et al., 1998a,b, 1999, 2001; Carlson et al., 1999, 2005; Dalton et al., 2005; Hibbits et al., 2000), as shown for the icy Galilean moons including Europa in Figure 2. Formation processes of red-orange concentrations of non-H_2O materials are not well understood for any of these moons but appear to correlate to expected magnetospheric irradiation (e.g., Paranicas et al., 2001, 2009) on the trailing hemispheres for Europa and Callisto, and to latitudinal variations expected from deflection of magnetospheric plasma by Ganymede's intrinsic dipolar magnetic field (Cooper et al., 2001; Khurana et al., 2007). Visible and UV spectroscopy are also used to map weathering products, e.g. O_2, $S0_2$, O_3, H_2O_2. Irradiation can make oxidants for life (Chyba 2000) but also obscure and even destroy chemical signatures of prebiotic and potentially biological evolution that would otherwise be high priority targets on potentially habitable worlds such as Europa and Enceladus. *For all these moons, and for others with global variations in relative contributions of different materials, the central one of space weathering remains as to how much of the spectroscopically mapped material is of magnetospheric or meteoritic versus intrinsic origin.*

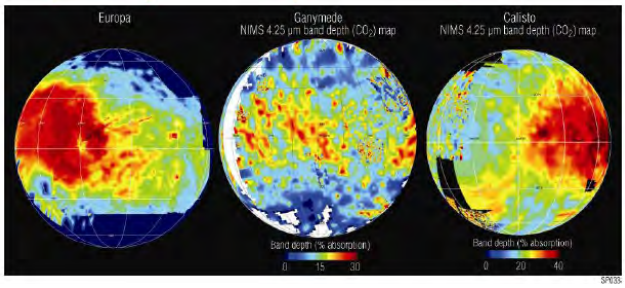


Figure 2. Near-infrared maps from the Galileo mission of non-H_2O surface materials on Europa (left, sulfate hydrates), Ganymede (middle, CO_2), and Callisto (right, CO_2). Red regions for Europa and Callisto are centered on the trailing hemispheres, consistent with expected magneto-spheric effects.

Weathering processes include plasma ion implantation into surfaces, surface sputtering by plasma and energetic ions, photolytic chemistry driven by solar ultraviolet irradiation, and radiolytic chemistry evolving from products of charged particle irradiation. Simultaneous action of multiple processes can substantially change the outcome, e.g. ion sputtering of an impact regolith followed by photon-induced desorption of neutral species such as sodium (Yakshinskiy and Madey, 1999). These microscale interactions are not independent of macroscale structures since regolith structure from impacts, and underlying deeper structures from internal evolution, can affect the efficacy of certain plasma-surface interactions, e.g. sputtering as affected by porosity, and surface irradiation dosage as partly attenuated by local topographic shielding (Cooper and Sturner, 2006). Topographic maps of disrupted icy surfaces on Europa, Enceladus, and other moons (e.g., Schenk, 2009; Schenk and McKinnon, 2009) make it clear that relatively

flat laboratory surfaces cannot really simulate planetary icy surfaces. Chemical products of space weathering could also potentially impact subsurface evolution, e.g. mixed gas clathrate accumulation affecting thermal conductivity within Europa's ice crust (Hand et al., 2006) and gas-driven cryovolcanism from radiolytic chemistry on Enceladus (Cooper et al., 2009).

High porosity of nearly all space-weathered surfaces means that any incident ion, electron, photon or micrometeoroid is likely to enter this porous covering, or regolith, before hitting a surface so that various processes may happen in partially-closed spaces. Grains within the porous material provide enormously greater surface area for chemical reactions than at smooth outer surfaces typically used in laboratory experiments and models as planetary analogues. A prime example of regolith effect on space weathering is V-NIR reddening of surfaces in the inner solar system. Hapke (e.g., 2001) discovered that this was due to the thin patina of reduced material (Fe and Fe/Si compounds, poor in O) on lunar regolith grains. This patina could be produced by a combination of ion and micrometeroid ablation inside of a regolith: Fe and Si ejecta stick to neighboring grains while O-containing compounds desorb and leave the regolith. In icy surface environments like Europa, porosity may play an important role in chemical activity of the irradiation surface for production of the surface-bound exosphere, mostly O_2 due to low efficiency of sticking to the surface (Shematovich et al., 2005; Cassidy et al., 2007), but with minor species produced in the ice and exosphere from sputtering and radiolytic chemistry driven by surface irradiation. Teolis et al. (2006) suggested that regolith porosity may be necessary for trapping of sufficient O_2 for radiolytic production of O_3, as found for simultaneous irradiation and water vapor deposition producing such porosity. More work is needed, Loeffler et al. (2009) found that ion bombardment reddened a flat mineral sample even more than a porous sample.

Processes of surface sputtering, micro-meteoroid vaporization, and photon stimulated desorption, all potentially contributing to space weathering, can be partially identified and characterized by the measurement of the ion and neutral species ejected from sputtered surfaces. Any ion freshly produced at the surface, or from neutrals in the overlying exosphere, is immediately accelerated by local magnetospheric or ionospheric electric fields and moves rapidly to spacecraft altitudes for direct in-situ measurement. The sputtered neutrals can also either be directly measured to determine energy distributions and composition, or the resultant exospheric distributions can be measured with respect to global spatial coordinates and composition. It is usually rather simple to understand how the energy distribution of the neutrals is influenced by the environment of the planetary objects, so a measurement of the atmospheric energy distribution can allow determination of relative sizes of gravitationally bound and escaping components. For the heaviest molecular species, pickup ion detection may be the most sensitive, since the spacecraft orbit (e.g. at Europa) would be far higher than scale heights for these heaviest species. Composition and flux measurements for the ions and neutrals can in turn provide information on the bulk composition of the planetary object and on the resurfacing rate. *These sputtered particle measurements are examples of how space weathering should also be regarded for mission planning both as a science driver and as enabling for instrumentation.*

Extremely important for subtraction of the space weathering contributions to surface and exospheric composition is that the background source composition of the magnetosphere be measured at least as well as the surface and overlying exospheric composition. Plasma composition in the Jupiter and Saturn magnetospheres should largely reflect composition of volcanic outflows respectively from Io and Enceladus. Erosive space weathering, e.g. sputtering, injects other elemental and molecular species into the local magnetospheric environments of the source moons. *But we have little knowledge of the plasma composition in either system or*

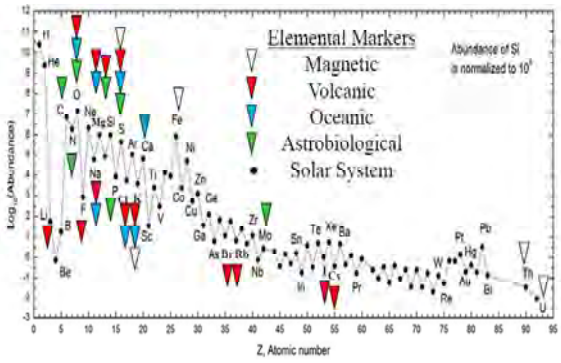


Figure 3. Solar system abundance of elements with markers for elements relating to key processes of high planetary science interest.

elsewhere beyond a few major species, and the enormous differences in the ion to neutral ratios in these magnetospheres are also not understood. Improvement of magnetospheric composition measurements to the full elemental (Figure 3) and isotopic level will inform us on internal chemistry of the sources while also defining the baseline exogenic contributions for the moon surfaces of those systems. If astrobiology and search for biosignatures, and precursors, continue to be major drivers for solar system exploration, then we *must* establish background reference baselines to guide the search for surface and exospheric anomalies indicative of such signatures. A prime example is the mass fractionation from sputtering and exospheric escape that would leaves heavier isotopes on an irradiated icy moon surface. Lighter isotopes, e.g. for carbon, could be used to locate freshly emergent material, potentially even biological.

For the ionic component of direct sputtering, a potentially overlooked but very important mechanism, is electron-stimulated desorption (ESD), e.g. Yakshinskiy and Madey (2001). This very well known process involves the excitation (often ionization) of a surface target followed by charge ejection, bond breaking and ion expulsion due to the resultant Coulomb repulsion. Though the role of electron stimulated desorption processes is not discussed much within in the planetary science community, the impinging energetic electrons that leak through the magnetospheres or reach the surfaces during flux transfer events may induce significant material removal. Given the energetics and the wide band-gap nature of the minerals on planetary bodies, the departing material may also be primarily ionic. *The role of 10 eV – 1 keV electron stimulated desorption and dissociation (ESD), along with keV-MeV electron radiolysis, could be significant.*

Finally, we must also consider impacts of surface irradiation on the electrostatic, magnetic, and electromagnetic field configurations at these surfaces. Static charge buildup can be driven by photonic and charged particle irradiation, and by triboelectic charging of dust particles in Enceladus-like cryovolcanic outflows. Potentials on the Moon have been observed from a few 10's of volts positive to tens of thousands of volts negative. Arc discharging might potentially be detectable via electromagnetic emissions and could transiently affect local chemistry, while also being a potential danger to in-situ surface experiments. On the Moon, dust levitation may be a significant source of surface alteration, while charged dust can rapidly lead to dust deposition and adhesion affecting robotic and manned surface operations. Charging of cometary dust may well be key to understanding plasma-dust interactions in the complex nucleus environment of a comet. Differential charge buildup (e.g. related to topographic variations in surface irradiation) may also lead to current flows (Gudipati et al., 2007) in near-surface micro-ionospheric environments of irradiated ices as a source of joule heating for activation of otherwise-inhibited chemical reactions, detectable emissions of electromagnetic radiation, and contributions to induced magnetic fields. Even asteroid surfaces (and the dusty surfaces of Phobos and Deimos) likely will be affected by charging. A consistent methodology for monitoring and modeling charging is required to understand and mitigate charging effects. *Assurance of reliable spacecraft and instrument operations in plasma charging and energetic particle radiation environments is a major cost driver for missions to planetary magnetospheric environments.*

Laboratory Measurements

Experiments in controlled laboratory configurations are crucial to interpretation and understanding of remote spectroscopic data for weathered planetary surfaces. The experiments must simulate key parameters of the space environments such as temperature, radiation environment, chemical composition (e.g. ices versus minerals), and ultrahigh vacuum conditions. Such experiments must be capable of mimicking target surface interactions of ionizing radiation (charged particles from the solar wind and galactic cosmic rays; solar photons) with low temperature ices and mixtures with silicates and carbonaceous matter. Examples of controlled parameters are: i) the chemical composition of the solid targets (water, ammonia, methanol, methane, silicates, carbonaceous matter), ii) crystalline and amorphous ices, iii) surface temperatures 10-550 K (Mercury to KBOs), iv) kinetic energy and composition of the charged particles, v) photon energy, and vi) cumulative effects of charged particles and photon interactions with ice/mineral surfaces. A key limitation is that sizes and thickness of samples are typically microscopic and not truly representative of real bulk samples, e.g. meters-thick impact regolith. Scattering nature of a thick sample makes optical transmission spectroscopy and other techniques difficult. *New facilities and techniques to acquire data from thick samples are needed.*

For thin film ices complementary detection systems must be interfaced so that newly formed atoms and molecules can be monitored in real time in the gas phase and in ices. By analyzing intermediates and products *on line* and *in situ*, the mechanisms of the formation of new species can be extracted as well. For the solid state, common detection systems are: i) infrared spectroscopy (detection of molecules via an excitation of vibrations which are associated with changes in the permanent dipole moment of the molecule), ii) Raman spectroscopy (monitoring vibrations of a molecule depicting a change in molecular polarizability of, for instance, homonuclear, infrared inactive diatomic species like molecular hydrogen, oxygen, and nitrogen), and iii) ultraviolet-visible spectroscopy (identification of electronic transitions in atoms and molecules, often radicals). Infrared and UV-visible laboratory data can also be useful for comparison with astronomical data, however more near-infrared (overtones of vibrations) laboratory data is needed for comparison to data acquired from orbiting spacecraft that routinely acquire data in this spectral region. Other techniques such as electron-paramagnetic resonance, fluorescence and x-ray spectroscopy could offer unique capabilities to identify radiolysis products in ices. Products released into the gas phase can be sampled via a quadrupole mass spectrometer (QMS) monitoring the complete product spectrum based on the mass-to-charge ratios of the ionized products (residual gas analyzer) or the velocity and hence kinetic energy of the products at a defined mass-to-charge ratio (time-of-flight mode). Products can be ionized via electron impact ionization or by tunable ultraviolet photons (selective ionization of structural isomers). The destruction efficiency of a molecule (the photodissociation cross section) depends strongly on the wavelength of the photon. *At present, no 'universal' detection scheme exists which can probe all products simultaneously in the gas phase and in the solid state.* However, careful combination of various analytical schemes can allow identification of atoms, diatomic, and polyatomic species.

Among the many complexities of space weathering simulation is that affected surfaces are in reality exposed simultaneously to particle and photonic energy sources. In the laboratory the products of these radiation components are usually measured separately, but these components may act in sequence to produce a substantially different result, e.g. higher sputtering yield when ion sputtering is followed by photon- or electron-induced desorption. It may then be difficult to pin down the reaction mechanism and to predict the chemical evolution of surfaces if we cannot distinguish if a product results from the photon and/or charged particle exposure. It is imperative

that all studies must be conducted in an identical experimental setup with exactly the same analytical devices to guarantee that all experimental conditions are well-defined and reproducible from experiment to experiment. A further limitation of restricted sample size is that simultaneous effects of the full range of particle and photon energies in incident spectra are not investigated.

Recommendations

Support comprehensive specification of space weathering environments through expansion of cooperation between NASA heliophysics and planetary science divisions on placement of environmental radiation instrumentation on appropriate missions, compilation of data and semiempirical models from measurements, and on predictive models for each environment.

Rationale: In the heliophysics community there is the concept of the Heliophysics Global Observatory (HGO), the collective fleet of operational heliophysics missions, that should be expanded for interdisciplinary applications to include planetary missions. Heliophysics support for space environment modeling, e.g. the Earth-Moon-Mars Radiation Environment Module (Schwadron et al., 2007) and earlier (e.g., NASA GSFC, JPL) models for solar and cosmic ray energetic particle modeling in the terrestrial planet domain, can be usefully applied to planetary interaction applications. Similarly, missions and data models for planetary interactions can also support investigation and modeling of interplanetary environments. HGO data virtual observatory approaches could be applied to planetary missions.

Encourage community-wide and interdisciplinary investigations of universal space weathering processes through balanced mixture of initiatives on mission instrument data analysis, laboratory measurements, computational modeling, and relevant theoretical investigations.

Rationale: The space environment is universal in the sense of connecting all the planetary environments, and so it most efficient to approach space weathering processes from the universal perspective, e.g. that similar processes act everywhere and the effects differ only in the relative energy deposition rates and compositional impacts of each process in different locations. Process investigations must be well-grounded in measurements for different environments, in broad-spectrum approaches to laboratory simulations, and best available inputs from theory and high performance computing.

Enable development of plasma ion, energetic particle, and neutral composition spectrometers for in-situ analyses to characterize elemental and isotopic range of interconnected planetary surface, atmospheric, ionosopheric, magnetospheric, and heliospheric environments.

Rationale: Our knowledge of composition in these environments beyond the Earth is limited to some major species with little or no information on the full range of elemental and isotopic composition that is critical to determination of origins, evolutionary processes, and astrobiological potential. Sample return is too expensive for general application, advanced in-situ analysis capabilities being required for one-way missions to most non-terrestrial destinations of the solar system. There is also strong coupling of composition for these connected environments and this coupling should be considered in weighting the relative priorities of measurements in each environment.

Provide facilities for more realistic laboratory science and engineering simulations of planetary surface environments under simultaneous influence of extreme limits on pressure, temperature, radiation, composition, physical structure, and endogenic or impact activity.

Rationale: There are no truly flat surfaces, particularly when viewed at the microscopic level of most space weathering processes, and multiple energy sources are typically operating on affected surfaces. The sensible and accessible surfaces have impact regolith layers extending to meters in depth and likely with high porosity under conditions of reduced gravity. Multi-phase interactions of ice, grains, and volatiles in irradiated bulk surface samples need much further investigation with appropriate facilities. Engineering simulation facilities require development to support realistic and extreme environment testing for future orbital and landed missions to irradiated icy bodies such as at Europa, Ganymede, Enceladus, and Triton.

References

Carlson, R. W., R. E. Johnson, and M. S. Anderson, Sulfuric acid on Europa and the radiolytic sulfur cycle, *Science* **286**, 97–99, 1999.

Carlson, R. W., *et al.*, Distribution of hydrate on Europa: Further evidence for sulfuric acid hydrate, *Icarus* **177**, 461-471, 2005.

Cassidy, T. A., R. E. Johnson, M. A. McGrath, M. C. Wong, and J. F. Cooper, The spatial morphology of Europa's near-surface O_2 atmosphere, Icarus, 191(2), 755-764, 2007.

Chyba CF., Energy for microbial life on Europa, *Nature* 403, 381-382, 2000.

Cooper, J. F., Niagara Falls Cascade Model for Interstellar Energetic Ions in the Heliosheath, Paper 386, Proceedings of the 30th International Cosmic Ray Conference, July 3-11, 2007, Merida, Mexico, Rogelio Caballero, Juan Carlos D'Olivo, Gustavo Medina-Tanco, Lukas Nellen, Federico A. Sánchez, José F. Valdés-Galicia (eds.), Universidad Nacional Autónoma de México, Mexico City, Mexico, Vol. 1 (SH), pp. 861–864, 2008.

Cooper, J. F., R. E. Johnson, B. H. Mauk, H. B. Garrett, and N. Gehrels, Energetic Ion and Electron Irradiation of the Icy Galilean Satellites, Icarus, 149, 133-159, 2001.

Cooper, J. F., and S. J. Sturner, Europa Surface Radiation Environment for Lander Assessment, Astrobiology Science Conference (AbSciCon) 2006, Washington, D.C., March 26-30, 2006.

Cooper, J. F., M. E. Hill, J. D. Richardson, and S. J. Sturner, Proton irradiation environment of solar system objects in the heliospheric boundary regions, in *Physics of the Inner Heliosheath, Voyager Observations, Theory, and Future Prospects*, 5th Annual IGPP International Astrophysics Conference. AIP Conference Proceedings, edited by J. Heerikhuisen, V. Florinski, G.P. Zank and N.V. Pogorelov, AIP, Vol. **858**, pp. 372-379, 2006.

Dalton, J. B., *et al.*. Spectral comparison of heavily hydrated salts with disrupted terrains on Europa, *Icarus* **177**, 472-490, 2005.

Gudipati, M, S., L. J. Allamandola, J. F. Cooper, S. J. Sturner, and R. E. Johnson, Consequence of Electron Mobility in Icy Grains on Solar System Objects, *Eos Trans. AGU, 88*(52), Fall Meet. Suppl., Abstract P53B-1248, 2007.

Hapke, B., Space weathering from Mercury to the asteroid belt. Journal of Geophysical Research 106, 10039-10074, 2001.

Hand, K. P., C. F. Chyba, R. W. Carlson, and J. F. Cooper, Clathrate hydrates of oxidants in the ice shell of Europa, Astrobiology, 6(3), 463-482, 2006.

Hibbitts, C. A., T. B. McCord, and G. B. Hansen, Distributions of CO_2 and SO_2 on the surface of Callisto, J. Geophys. Res., **105(E9)**, 22541-22557, 2000.

Hudson, R. L., M. E. Palumbo, G. Strazzulla, M. H. Moore, J. F. Cooper, and S. J. Sturner, Laboratory Studies of the Chemistry of TNO Surface Materials, in *The Solar System Beyond Neptune*, edited by A. Barucci, H. Boehnhardt, D. P. Cruikshank, and A. Morbidelli, pp. 507-523, Univ. of Arizona Press, Tucson, 2008.

Ip, W.-H., D. J. Williams, R. W. McEntire, and B. H. Mauk, Ion sputtering and surface erosion at Europa, Geophys. Res. Lett. **25(6)**, 829-832, 1998.

Johnson, R. E., *Energetic Charged Particle Interactions with Atmospheres and Surfaces.* Springer-Verlag, Heidelberg, 1990.

Johnson, R. E., J. F. Cooper, L. J. Lanzerotti, and G. Strazzulla, Radiation formation of a non-volatile crust, Astron. Astrophys., 187, 889-892, 1987.

Khurana, K. K., R. T. Pappalardo, N. Murphy, and T. Denk, The origin of Ganymede's polar caps, Icarus, 191, 193-202, 2007.

Loeffler, M. J., Dukes, C. A., Baragiola, R. A., Irradiation of olivine by 4 keV He+: Simulation of space weathering by the solar wind. Journal of Geophysical Research (Planets) 114, 3003, 2009.

McComas, D. J., and N. A. Schwadron, An explanation of the Voyager paradox: Particle acceleration at a blunt termination shock, Geophys. Res. Lett. 33, L04102, doi: 10.1029/2005GL025437, 2006.

McCord, T. B., and 11 colleagues, Salts on Europa's surface detected by Galileo's Near Infrared Mapping Spectrometer, Science 280, 1242–1245, 1998a.

McCord, T. B., and 12 colleagues, Non-water-ice constituents in the surface material of the icy Galilean satellites from the Galileo near-infrared mapping spectrometer investigation, J. Geophys. Res. 103, 8603–8626, 1998b.

McCord, T. B., and 11 colleagues, Hydrated salt minerals on Europa's surface from the Galileo near-infrared mapping spectrometer (NIMS) investigation, J. Geophys. Res. 104, 11827–11851, 1999.

McCord, T. B., et al., Thermal and radiation stability of the hydrated salt minerals epsomite, mirabilite, and natron under Europa environmental conditions, J. Geophys. Res. 106, E2, 3311-3319, 2001.

Paranicas, C., R. W. Carlson, and R. E. Johnson, Electron bombardment of Europa, Geophys. Res. Lett., 28, 673-676, 2001.

Paranicas, C., J. F. Cooper, H. B. Garrett, R. E. Johnson, and S. J. Sturner. Europa's Radiation Environment and its Effect on the Surface, in EUROPA, Editors: R. Pappalardo, W. B. McKinnon, and K. Khurana, University of Arizona Press Space Science Series, in press, 2009.

Schenk, P. M., Slope characteristics of Europa: Constraints for landers and radar Sounding, Geophys. Res. Lett., 36, L15204, doi:10.1029/2009GL039062, 2009.

Schenk, P. M., and W. B. McKinnon, One-hundred-km-scale basins on Enceladus: Evidence for an active ice shell, Geophys. Res. Lett., 36, L16202, doi:10.1029/2009GL039916, 2009.

Schwadron, N. A., C. Goodrich, H. Spence, L. Townsend, F. Cucinotta, M. Weyland, M. Golightly, M. Desai, A. Posner, D. Hassler, D.. Krauss-Varban, J. Luhmann, J. Miller, B. Heber, T. Onsager, Earth-Moon-Mars Radiation Environment Module (EMMREM), Aerospace Conference, 2007 IEEE , pp.1-10, doi10.1109/AERO.2007.352781, 3-10 March 2007.

Shematovich, V. I., R. E. Johnson, J. F. Cooper, and M. C. Wong, Surface-bounded atmosphere of Europa, Icarus, 173(2), 480-498, 2005.

Stone, E. C., A. C. Cummings, F. B. McDonald, B. C. Heikkila, N. Lal, and W. R. Webber, Voyager 1 Explores the Termination Shock Region and the Heliosheath Beyond, *Science*, 309, 2017-2020, 2005.

Stone, E. C., A. C. Cummings, F. B. McDonald, B. C. Heikkila, N. Lal, W. R. Webber, An asymmetric solar wind termination shock, Nature, 454, 71-74, 2008.

Yakshinskiy, B. V., and T. E. Madey, Photon-stimulated desorption as a substantial source of sodium in the lunar atmosphere, Nature, 400, 642-644, 1999.

Yakshinskiy, B. V., and T. E. Madey, Electron- and photon-stimulated desorption of K from ice surfaces, J. Geophys. Res. 106, 33303-33307, 2001.

November 10, 2010

Ultraviolet Coronagraph Spectroscopy: A Key Capability for Understanding the Physics of Solar Wind Acceleration

S. R. Cranmer,[1] J. L. Kohl,[1] D. Alexander,[2] A. Bhattacharjee,[3] B. A. Breech,[4] N. S. Brickhouse,[1] B. D. G. Chandran,[3] A. K. Dupree,[1] R. Esser,[5] S. P. Gary,[6] J. V. Hollweg,[3] P. A. Isenberg,[3] S. W. Kahler,[7] Y.-K. Ko,[8] J. M. Laming,[8] E. Landi,[9], W. H. Matthaeus,[10] N. A. Murphy,[1] S. Oughton,[11] J. C. Raymond,[1] D. B. Reisenfeld,[12] S. T. Suess,[13] A. A. van Ballegooijen,[1] and B. E. Wood[8]

[1]*Harvard-Smithsonian CfA,* [2]*Rice University,* [3]*U. New Hampshire,* [4]*U.S. Army Research Laboratory,* [5]*U. Tromsø, Norway,* [6]*Los Alamos National Laboratory,* [7]*Air Force Research Laboratory,* [8]*Naval Research Laboratory,* [9]*U. Michigan,* [10]*Bartol Research Institute, U. Delaware,* [11]*U. Waikato, New Zealand,* [12]*U. Montana,* [13]*National Space Science & Technology Center*

Abstract

Understanding the physical processes responsible for accelerating the solar wind requires detailed measurements of the collisionless plasma in the extended solar corona. Some key clues about these processes have come from instruments that combine the power of an ultraviolet (UV) spectrometer with an occulted telescope. This combination enables measurements of ion emission lines far from the bright solar disk, where most of the solar wind acceleration occurs. Although the UVCS instrument on *SOHO* made several key discoveries, many questions remain unanswered because its capabilities were limited. This white paper summarizes these past achievements and also describes what can be accomplished with next-generation instrumentation of this kind.

1. Background and Motivation

The hot, ionized outer atmosphere of the Sun is a unique testbed for the study of magnetohydrodynamics (MHD) and plasma physics, with ranges of parameters that are inaccessible on Earth. Although considerable progress has been made during the last few decades, we still do not know the basic physical processes responsible for heating the million-degree solar corona and accelerating the solar wind. Identifying these processes is important not only for understanding the origins and impacts of space weather (e.g., Schwenn 2006; Eastwood 2008), but also for establishing a baseline of knowledge about a well-resolved star that is directly relevant to other astrophysical systems.

In order to construct and test theoretical models, a wide range of measurements of relevant plasma parameters must be available. In the low-density, open-field regions that reach into interplanetary space, more parameters need to be measured (in comparison to collision-dominated regions) because the plasma becomes **collisionless.** In other words, individual particle species—e.g., protons, electrons, helium, and minor ions—can exhibit different properties from one another. Such differences in particle velocity distributions are valuable probes of "microscopic" processes of heating and acceleration. In interplanetary space, such kinetic properties have been measured for decades by *in situ* particle instruments (e.g., Marsch 1999, 2006; Kasper et al. 2008). However, measurements in the near-Sun regions that are being actively heated and accelerated have been more limited in scope.

Remote-sensing measurements of plasma properties in the so-called "extended solar corona" (above about 1.5 $R_\odot$ measured from Sun-center) are difficult to make because the photon emission at large heights is fainter by many orders of magnitude than the emission from the solar disk. Standard telescopes, that do not explicitly block out the solar disk, typically contain enough scattered light from the disk to totally mask the faint off-limb emission. Because the corona is highly ionized, the dominant spectral features are at wavelengths accessible only from space. For these reasons, a series of Ultraviolet Coronagraph Spectrometer (UVCS) instruments have been built and flown on rockets, on a Shuttle-deployed *Spartan* payload, and as an instrument on the *Solar and Heliospheric Observatory* (*SOHO*) spacecraft; see, e.g., Kohl et al. (1978,

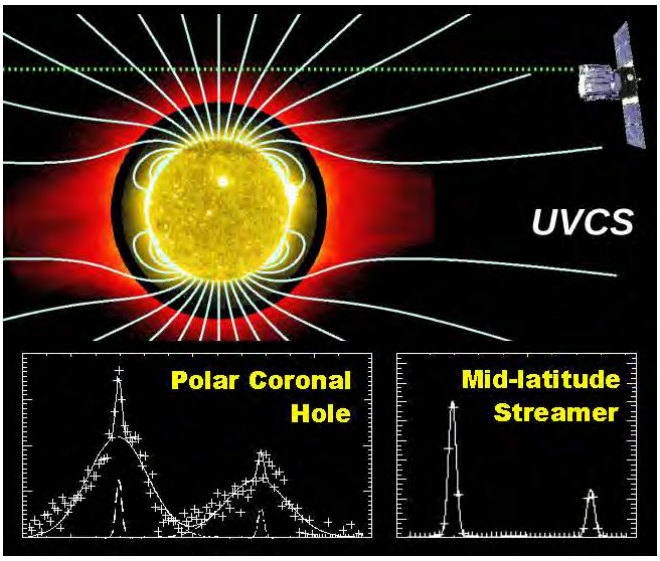


Figure 1: Combined *SOHO* image of the solar corona from 17 August 1996, showing the solar disk in Fe XII 195 Å intensity from EIT (yellow inner image) and the extended corona in O VI 1032 Å intensity from UVCS (red outer image). A typical UVCS observational line of sight is shown in green. Axisymmetric field lines are from the solar-minimum model of Banaszkiewicz et al. (1998), and O VI emission line profiles (bottom) are from various UVCS observations at $r > 2R_\odot$ in 1996–1997 (Kohl et al. 1997).

1994, 1995, 1997, 2006). In these instruments, the coronagraphic blocking of bright light from the solar disk is done efficiently with a pair of occulters: one external to the telescope and one internal. Figure 1 illustrates UVCS/*SOHO* observations of the extended corona.

By measuring off-limb emission lines formed both by collisional excitation and by the scattering of solar-disk photons, UV spectroscopy provides a multi-faceted characterization of the kinetic properties of atoms, ions, and free electrons (e.g., Withbroe et al. 1982; Kohl et al. 2006). The Doppler-broadened shapes of emission lines are direct probes of line-of-sight (LOS) particle velocity distributions (i.e., essentially providing $T_\perp$ when the off-limb magnetic field is ∼radial). Red/blue Doppler shifts reveal bulk flows along the LOS. Integrated intensities of resonantly scattered lines can be used to constrain the solar wind velocity and other details about the velocity distribution in the radial direction (e.g., $u_\parallel$ and $T_\parallel$); this is the so-called "Doppler dimming/pumping" diagnostic (e.g., Noci et al. 1987). Intensities of collisionally dominated lines—especially when combined into an emission measure distribution—can constrain electron temperatures, densities, and elemental abundances in coronal plasma. Even departures from Maxwellian and bi-Maxwellian velocity distributions are detectable with spectroscopic measurements having sufficient sensitivity and spectral resolution.

In the **fast solar wind,** UVCS/*SOHO* measured outflow speeds that become supersonic much closer to the Sun than previously believed. In coronal holes, heavy ions (e.g., O^{+5}) were found to flow faster than the protons, to be heated hundreds of times more strongly than protons and electrons, and to have anisotropic temperatures with $T_\perp > T_\parallel$ (Kohl et al. 1997, 1998, 1999; Cranmer et al. 1999b). Specifically, the anisotropic and super-heated oxygen velocity distributions (with $T_\perp > 10^8$ K) have guided theorists to discard some candidate physical processes and further investigate others. Hollweg & Isenberg (2002) stated in a review paper that "We have seen that the information provided by UVCS has been pivotal in defining how research has proceeded during the past few years." [1]

Figure 2 gives more information about how the temperatures of protons, electrons, and an example minor ion species (O^{+5} in the corona, O^{+6} *in situ*) differ from one another in the high-speed solar wind. Although the measured kinetic temperatures for protons are of order 2–3.5 MK, these are combinations of random thermal motions and "nonthermal" broadening due to Alfvén waves or other unresolved line-

[1] More recently, these results were highlighted in Chapter 3 of the 2010 report of the APS-sponsored Workshop on Opportunities in Plasma Astrophysics (WOPA); see http://www.pppl.gov/conferences/2010/WOPA/

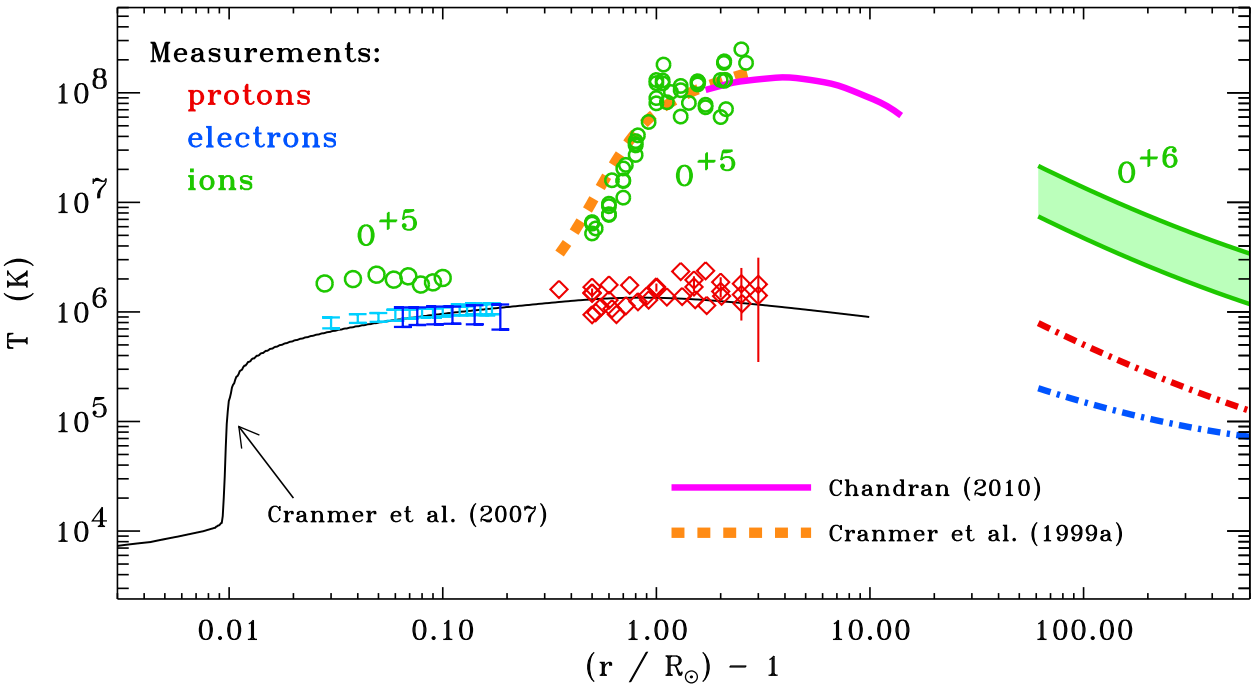


Figure 2: Radial dependence of temperatures in polar coronal holes and fast wind streams. Mean plasma temperature from a turbulence-driven coronal heating model (solid black curve; Cranmer et al. 2007). T_e from off-limb SUMER measurements made by Wilhelm (2006) (dark blue bars) and Landi (2008) (light blue bars). T_p from UVCS data assembled by Cranmer (2004) (red symbols). Perpendicular O^{+5} ion temperatures (green circles) from Landi & Cranmer (2009) ($r < 1.1 R_\odot$) and Cranmer et al. (2008) ($r > 1.5 R_\odot$). *In situ* values of T_p and T_e (red and blue dot-dashed curves) at $r > 60 R_\odot$ are from Cranmer et al. (2009), and the *in situ* range for O^{+6} (light green region) is inferred from, e.g., Collier et al. (1996). Ion heating theories assuming high-$k_\perp$ waves (magenta solid curve) and high-$k_\parallel$ waves (orange dashed curve) both appear to succeed in modeling the O^{+5} data.

of-sight motions. In Figure 2 we attempted to remove the contribution of waves to show the true proton temperatures. The resulting T_p does not show as clear a signal of preferential heating (relative to T_e) as one would have seen from just the kinetic temperature. Although one can still marginally see that $T_p > T_e$, the existing measurements of T_p and T_e do not overlap with one another in radius. Improved measurements are needed in order to better constrain the proton and electron heating rates in the corona. The green points in Figure 2 indicate the key measurements of preferential O^{+5} heating in polar coronal holes. The magenta and orange curves show a subset of the many theoretical models that have been constructed to explain these data.

At solar minimum, UVCS/*SOHO* found that the **slow solar wind** flows mostly along the outer edges of bright streamers, near locations with measured abundance patterns matching those of the *in situ* slow wind (Raymond et al. 1997; Strachan et al. 2002; Abbo et al. 2010). The closed-field "core" regions of streamers, however, exhibit heavy element abundances only 3% to 30% of those seen at 1 AU, indicating gravitational settling (e.g., Raymond 1999; Vásquez & Raymond 2005) or complex flux-tube geometries (Noci et al. 1997). UVCS observed the transition from a high-density collision-dominated plasma at low heights in streamers to a low-density collisionless plasma at large heights, the latter exhibiting high ion temperatures and anisotropies that suggest similar physics as in the fast wind (Frazin et al. 2003). UVCS has also been used to measure plasma properties of coronal mass ejections (CMEs) and put useful constraints on their reconnection rates, 3D flow velocities, energy budgets, magnetic helicities, and shock compression ratios (see, e.g., Raymond 2002; Kohl et al. 2006).

2. Next-Generation Capabilities Are Needed

Despite the advances outlined above, the diagnostic capabilities of UVCS were limited to what was foreseen before the *SOHO* era (when only H I Lyα had been observed). Thus, we are still severely limited in our ability to answer the fundamental questions by our lack of knowledge of the plasma properties in the solar wind's acceleration region. The following enhancements in measurement capability are needed to be able to conclusively identify and characterize the processes that energize the solar wind.

a. The key UVCS discovery of preferential ion heating and anisotropic ion velocity distributions near the Sun was essentially limited to just one ion (O^{+5}). If the kinetic properties of **additional minor ions** were to be measured in the extended corona (i.e., a wider sampling of charge/mass combinations) we could much better determine the nature of the dominant collisionless heating process. Specifically, these measurements would provide an improved empirical description of kinetic waves and turbulence in the corona, with which we could conclusively identify the type of fluctuations (such as ion cyclotron waves, kinetic Alfvén waves, or other nonlinear turbulent modes) as well as their means of dissipation (Hollweg & Isenberg 2002; Cranmer 2002, 2009). Sampling as broad a range as possible in the ion charge-to-mass ratio (Z/A) provides a sensitive probe of the heating; e.g., S^{+5} (933 Å, $Z/A = 0.16$). Ca^{+9} (557 Å, $Z/A = 0.22$), Si^{+8} (296 Å, $Z/A = 0.28$), O^{+5} (1032 Å, $Z/A = 0.31$), Mg^{+9} (610 Å, $Z/A = 0.37$), and Si^{+11} (499 Å, $Z/A = 0.39$). With these ions, the two models shown in Figure 2 (which use either high-$k_\perp$ or high-$k_\parallel$ waves) would be clearly distinguished from one another; see also Leamon et al. (2000). Our confidence in the uniqueness of a successful model increases as the number of ions observed increases.

b. UVCS provided new constraints on the heating of minor ions, but not so much on the heating of the primary proton–electron plasma. Observationally, the biggest missing piece is the **electron temperature** above $\sim$1.5 $R_\odot$. Direct measurements of T_e from, e.g., the broad Thomson-scattered component of H I Lyα, would allow us to determine the bulk-plasma heating rate in different solar wind regions, as well as the partitioning of energy between protons and electrons. This partitioning is a key diagnostic of turbulence models (e.g., Matthaeus et al. 2003; Howes 2010) as well as a driver of the stability of helmet streamers (e.g., Endeve et al. 2004). The downward conduction of electron thermal energy—from the corona to the transition region—sets the base pressure and thus the mass flux of the solar wind (Withbroe 1988; Lie-Svendsen et al. 2002), thus making measurements of $T_e(r)$ especially important. Improved observations of proton temperatures are also important, since the measured rates of proton heating can be compared directly with the energy available to protons in ion cyclotron waves (as constrained by future measurements of additional minor ions; see item [a] above) to reveal the relative importance of resonant wave heating to the primary plasma.

c. In addition to protons and electrons, **helium** also plays a major role in the "primary" plasma of the solar wind. Helium may regulate the wind's mass flux (Hansteen et al. 1994) and set a lower limit on its outflow speed (Kasper et al. 2007). Preferential heating of alpha particles may even dominate the coronal heating close to the Sun and control how much heat is received by the protons (Liewer et al. 2001; Li 2003). Observations of the He II 304 Å and He I 584 Å lines can be used to measure the coronal helium abundance, departures from ionization equilibrium, and the amount of preferential alpha-particle heating.

d. Space-based observations by *Hinode* and *SDO*, as well as ground-based eclipse observations, have highlighted the importance of **greater spatial and temporal resolution** to the direct identification of MHD waves with periods between about 0.1 and 10 minutes (De Pontieu et al. 2007; Pasachoff et al. 2002; Singh et al. 2009). The predicted amplitudes of most MHD waves grow significantly as they propagate up to the extended corona, thus making a UVCS-type instrument with greater sensitivity and resolution ideal for detecting both compressive waves (via intensity oscillations) and Alfvén waves (via Doppler shifts). Measuring the frequency and wavenumber dependence of these waves would open new windows on our understanding of coronal turbulence (e.g., Matthaeus et al. 1999). Greater spatial and temporal resolution also allows improved characterization of small-scale inhomogeneities between neighboring

flux tubes in coronal holes and streamers. Precise measurements of cross-field density gradients can put firm constraints on models of coronal heating via drift instabilities (e.g., Mecheri & Marsch 2008) and MHD discontinuities in the solar wind (Feldman et al. 1997; Borovsky 2008).

e. Recent analysis of *in situ* data shows that the combined **elemental abundances and charge states** of solar wind ions can be used to effectively probe the origin of heliospheric wind streams in closed or open flux tubes (Zurbuchen 2007). There have been some isolated comparisons of ion composition between UVCS and *in situ* measurements (e.g., Ko et al. 2006), but higher resolution and greater sensitivity to more ions is needed to produce corona-heliosphere "mappings" of abundances and charge states that are detailed enough to test models of coronal gravitational settling and ion drag (Lenz 2004; Li et al. 2006) and the first ionization potential (FIP) effect (e.g., Laming 2004, 2009). It is also important to measure a broader range of low-FIP and high-FIP ions in order to better map out the way the solar atmosphere becomes fractionated.

f. Measuring **non-Maxwellian velocity distributions** of electrons and positive ions would provide even more stringent tests of specific models of MHD turbulence, cyclotron resonance, and velocity filtration. An instrument with greater sensitivity than UVCS/*SOHO* could detect subtle departures from Gaussian line shapes that signal the presence of specific non-Maxwellian distributions (e.g., Cranmer 1998, 2001).

New capabilities such as these would be enabled by greater photon sensitivity, an expanded wavelength range, and the use of measurements that heretofore have been utilized only in a testing capacity (e.g., Thomson-scattered H I Lyα to obtain T_e). These would allow the relative contributions of different physical processes to the heating and acceleration of all solar wind plasma components to be determined directly.

3. Synergy with Other Instrumentation and Missions

UV coronagraph spectroscopy has been a key asset to a large number of multi-instrument and multi-mission observation campaigns (e.g., Galvin & Kohl 1999). For example, coordinated observations taken at times when *SOHO* and *Ulysses* could track the same plasma parcels from the Sun to the outer heliosphere led to new methods of mapping field lines over long distances in the solar wind (Suess et al. 2000; Poletto et al. 2002). Next-generation UVCS-type measurements would also provide key context for the regions of the inner heliosphere that *Solar Probe Plus* (*SPP*) and *Solar Orbiter* will fly through, and they would provide direct measurements of coronal regions inaccessible even to *SPP* (i.e., $r < 9 R_\odot$).

Measurements from other instruments also can enhance the ability of UV coronagraph spectroscopy to determine coronal plasma parameters. Although UV emission lines can be used to measure the coronal electron density (e.g., Noci et al. 1997; Antonucci et al. 2004), the most straightforward method to obtain n_e has been the analysis of polarized brightness measured by a white-light coronagraph. The combination of a UVCS-type instrument and a white-light imaging coronagraph provides useful diagnostics that extend beyond what is possible with either instrument in isolation.

There are several possible mission concepts for next-generation UV coronagraph spectrometers. Sub-orbital rockets are useful for demonstrating the capabilities of this kind of instrument, as well as being able to obtain useful scientific results (see § 4.1 of Kohl et al. 2006). Studies have also shown that a standard Low-Earth orbit (350–600 km altitude at 28° latitude) can provide adequate observation time, low UV optical depth from the Earth's atmosphere, and negligible contamination to enable the full range of required UV coronagraph spectroscopy. A polar, Sun-synchronous orbit would of course be an improvement over a low-latitude orbit because of the lack of diurnal solar occultation by the Earth. Another possibility is to position the spacecraft at the **Earth-Sun L5 point** (see also Vourlidas et al. 2010) which eliminates issues of atmospheric absorption and orbital night. This location also positions the prime off-limb field of view (i.e., the "plane of the sky") close to the Earth-Sun line, so that geoeffective solar wind streams and CMEs can be measured directly.

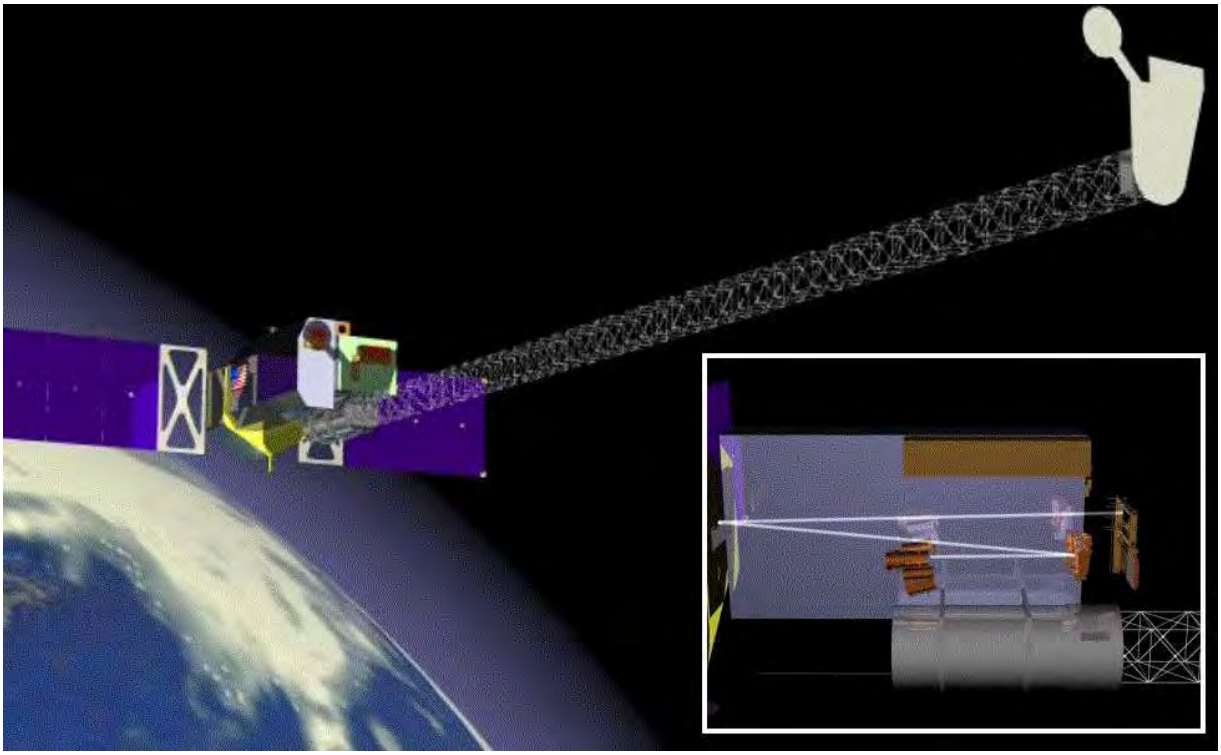


Figure 3: Spacecraft concept for a large-aperture UV coronagraph with an external occulter at the end of a 13 m boom. The inset shows a diagram of an advanced LAUVCS instrument (see Kohl et al. 2006).

4. Instrumentation and Costs

The UV spectroscopic observations required to produce detailed empirical descriptions of solar wind source regions can be accomplished with a large-aperture ultraviolet coronagraph spectrometer (LAUVCS). The large aperture is needed to provide the sensitivity to observe relatively weak spectral lines with the required range of charge-to-mass ratios. The large aperture can be realized with a remote external occulter supported by a deployable mast (see Figure 3). Such an instrument would accommodate a much larger telescope area than UVCS/*SOHO* that is both shielded from direct solar disk rays and also has an unobstructed view of the coronal heights of interest. When equipped with state-of-the-art reflection coatings, high ruling efficiency diffraction gratings, and array detectors, such an instrument exceeds the sensitivity and stray light suppression needed to make all required coronal observations in both streamers and coronal holes up to at least 5 $R_\odot$ and often higher. It is also possible to achieve the large aperture with an internally occulted coronagraph where the solar-disk radiation impinges directly on the telescope primary. The sensitivity is just as good as that of the externally occulted instrument, but the stray light characteristics only allow observations up to $\sim 2R_\odot$ in coronal holes and $\sim 3R_\odot$ in streamers. Compared to UVCS, either instrument would make a tremendous step forward in describing solar wind source regions and identifying the physical processes that heat and accelerate the solar wind (and CMEs[2]).

Both externally occulted and internally occulted LAUVCS type instruments can be built with existing technology. Deployable masts to support the remote external occulter, such as the one baselined for the *NuSTAR* mission, are commercially available. There is also "in space" performance data that demonstrates the ATK ADAM mast exceeds the specifications required for the externally occulted LAUVCS. Intensified

[2]The ability of next-generation UV coronagraph spectroscopy to answer fundamental questions about how CMEs and solar energetic particles (SEPs) are produced will be described in a separate white paper.

CCD (ICCD) detectors similar to those used on *XMM* and *Swift* can be used for a LAUVCS, and they are capable of meeting all requirements for the solar wind science while maintaining the full instrument sensitivity. The KBr-coated microchannel plates that are needed to meet the UV sensitivity requirements are also available commercially, and they can be incorporated into an ICCD with an acceptable air exposure. To achieve the LAUVCS science objectives for CMEs—without the need to attenuate the associated high light levels—it is highly desirable to use an intensified active pixel sensor (IAPS), which is the next step beyond the ICCD. A laboratory model of such a device exists at the Mullard Space Science Laboratory (MSSL), and it has been clocked at 66 MHz to achieve the desired maximum detection rate of 32 Hz per pixel with less than a 10% dead time loss.

Both the solar wind and the baseline CME science objectives can be achieved with an ICCD detector similar to those flown on *XMM* and *Swift*. The advanced technology IAPS detectors currently being developed at MSSL are showing great promise, and they could provide a dynamic range and maximum count rate capability that would accommodate the bright CME emissions without the need for attenuators.

The cost (in FY 2011 dollars) of an externally occulted LAUVCS for a Class C mission with limited redundancy, including design, development, calibration, environmental testing, and support for integration, is approximately $45M including the cost of the deployable mast. The cost of the internally occulted instrument would be about $15M less. These estimated costs do not include any reserve/contingency or margin except for costed schedule reserve, and they do not include any Phase E mission operation and data analysis costs.

References

Abbo, L., et al. 2010, Adv. Sp. Res., 46, 1400
Antonucci, E., et al. 2004, A&A, 416, 749
Banaszkiewicz, M., et al. A&A, 337, 940
Borovsky, J. E. 2008, JGR, 113, A08110
Chandran, B. D. G. 2010, ApJ, 720, 548
Collier, M. R., et al. 1996, GRL, 23, 1191
Cranmer, S. R. 1998, ApJ, 508, 925
Cranmer, S. R. 2001, JGR, 106, 24937
Cranmer, S. R. 2002, Proc. SOHO-11, 361, astro-ph/0209301
Cranmer, S. R. 2004, Proc. SOHO-15, 154, astro-ph/0409724
Cranmer, S. R. 2009, Living Rev. Sol. Phys., 6, 3
Cranmer, S. R., et al. 1999a, ApJ, 518, 937
Cranmer, S. R., et al. 1999b, ApJ, 511, 481
Cranmer, S. R., et al. 2007, ApJS, 171, 520
Cranmer, S. R., et al. 2008, ApJ, 678, 1480
Cranmer, S. R., et al. 2009, ApJ, 702, 1604
De Pontieu, B., et al. 2007, Science, 318, 1574
Eastwood, J. P. 2008, Phil. Trans. Roy. Soc. A, 366, 4489
Endeve, E., et al. 2004, ApJ, 603, 307
Feldman, W. C., et al. 1997, JGR, 102, 26905
Frazin, R. A., et al. 2003, ApJ, 597, 1145
Galvin, A. B., & Kohl, J. L. 1999, JGR, 104, 9673
Hansteen, V. H., et al. 1994, ApJ, 428, 843
Hollweg, J. V., & Isenberg, P. A. 2002, JGR, 107 (A7), 1147
Howes, G. G. 2010, MNRAS, in press, arXiv:1009.4212
Kasper, J. C., et al. 2007, ApJ, 660, 901
Kasper, J. C., et al. 2008, PRL, 101, 261103
Ko, Y.-K., et al. 2006, ApJ, 646, 1275
Kohl, J. L., et al. 1978, in New Instrumentation for Space Astronomy (Oxford: Pergamon), 91
Kohl, J. L., et al. 1994, Space Sci. Rev., 70, 253
Kohl, J. L., et al. 1995, Solar Phys., 162, 313
Kohl, J. L., et al. 1997, Solar Phys., 175, 613
Kohl, J. L., et al. 1998, ApJ, 501, L127
Kohl, J. L., et al. 1999, ApJ, 510, L59
Kohl, J. L., et al. 2006, A&A Review, 13, 31
Laming, J. M. 2004, ApJ, 614, 1063
Laming, J. M. 2009, ApJ, 695, 954
Laming, J. M., & Lepri, S. T. 2007, ApJ, 660, 1642
Landi, E. 2008, ApJ, 685, 1270
Landi, E., & Cranmer, S. R. 2009, ApJ, 691, 794
Leamon, R. J., et al. 2000, ApJ, 537, 1054
Lenz, D. D. 2004, ApJ, 604, 433
Li, B., et al. 2006, JGR, 111, A08106
Li, X. 2003, A&A, 406, 345
Lie-Svendsen, Ø., et al. 2002, ApJ, 566, 562
Liewer, P., et al. 2001, JGR, 106, 29261
Marsch, E. 1999, Space Sci. Rev., 87, 1
Marsch, E. 2006, Liv. Rev. Sol. Phys., 3, 1
Matthaeus, W. H., et al. 1999, ApJ, 523, L93
Matthaeus, W. H., et al. 2003, Nonlin. Proc. Geophys., 10, 93
Mecheri, R., & Marsch, E. 2008, A&A, 481, 853
Noci, G., et al. 1987, ApJ, 315, 706
Noci, G., et al. 1997, Proc. SOHO-5, ESA SP-404, 75
Pasachoff, J. M., et al. 2002, Solar Phys., 207, 241
Poletto, G., et al. 2002, JGR, 107, 1300
Raymond, J. C. 1999, Space Sci. Rev., 87, 55
Raymond, J. C. 2002, Proc. SOHO-11, ESA SP-508, 421
Raymond, J. C., et al. 1997, Solar Phys., 175, 645
Schwenn, R. 2006, Liv. Rev. Sol. Phys., 3, 2
Singh, J., et al. 2009, Solar Phys., 260, 125
Strachan, L., et al. 2002, ApJ, 571, 1008
Suess, S. T., et al. 2002, JGR, 105, 25033
Vásquez, A. M., & Raymond, J. C. 2005, ApJ, 619, 1132
Vourlidas, A., et al. 2010, NRC Decadal Survey White Paper
Wilhelm, K. 2006, A&A, 455, 697
Withbroe, G. L. 1988, ApJ, 325, 442
Withbroe, G. L., et al. 1982, Space Sci. Rev., 33, 17
Zurbuchen, T. H. 2007, Ann. Rev. Astron. Astrophys., 45, 297

Particle Acceleration and Transport in the Heliosphere (PATH)
Objective: Determine the Mechanisms Responsible for Energetic Particle Acceleration and Transport at the Sun and in the Interplanetary Medium

A white paper for the 2010 Solar and Space Physics Decadal Survey

Principal Author: M. I. Desai (SwRI)
Co-Authors (alphabetical): E. R. Christian (GSFC), C. M. S. Cohen (Caltech), J. F. Cooper (NASA/GSFC), G. A. de Nolfo (GSFC), H. A. Elliott (SwRI), J. Giacalone (U. Arizona), S. Geier (JPL), G. C. Ho (JHU/APL), B. Klecker (MPI/Garching), H. Kucharek (UNH), L. J. Lanzerotti (NJIT), M. A. Lee (UNH), S. T. Lepri (U. Mich.), R. A. Leske (Caltech), R. Lin (UCBerkeley), J. E. Mazur (Aerospace), D. J. McComas (SwRI), E. Möbius (UNH), E. C. Roelof (JHU/APL), J. M. Ryan (UNH), N. A. Schwadron (UNH), T. T. von Rosenvinge (NASA/GSFC), R. F. Wimmer-Schweingruber (U. Kiel), P. Wurz (U. Bern)

Abstract

Solar Energetic Particles (SEP) above 10 MeV in energy with proton intensities nearly 5 orders of magnitude more than the galactic cosmic rays pose a serious radiation hazard to both instrumentation and humans in space. Identifying the sources and underlying physics that produce these SEPs is critical, if we are to reliably forecast such events at 1 AU. However, the origin of this acceleration, whether it is in flares or at CME-driven shocks, and the complex roles of the variations in the seed populations in such shocks, is still not well understood. The Particle Acceleration and Transport in the Heliosphere (PATH) mission consists of a suite of high sensitivity and high resolution in-situ and remote sensing instruments designed to determine the mechanisms responsible for the acceleration and propagation of SEPs through the inner heliosphere. PATH is a small-class mission consisting of a spinning spacecraft pointed near the Sun at the L-1 Lagrange point. This configuration and instrument suite also enables PATH to provide the critical real-time solar wind, magnetic field, and energetic particle parameters that are required for Space Weather monitoring and forecasting. In terms of science objectives, estimated cost, and schedule, the PATH mission is fully compatible with (1) the goals of NASA's Heliophysics Division, (2) the future plans described in the 2009 Heliophysics Roadmap and the 2002 NRC decadal survey, and (3) plans and activities of other space agencies.

1. Introduction

Massive explosions on the Sun routinely accelerate energetic ions up to ~GeV energies and electrons up to ~100s of MeV and generate intense electromagnetic (EM) radiation in the form of x-rays and γ-rays. Understanding the mechanisms that are responsible for accelerating these solar energetic particles (SEPs) and influencing their propagation to Earth are of critical importance for developing physics-based capabilities that predict the hazardous effects of solar radiation on technological systems and humans on Earth and in interplanetary space.

Recent observations from many spacecraft in the Heliophysics System Observatory such as Wind, SoHO, RHESSI, ACE, and STEREO have contributed dramatically to our understanding of SEPs. However, it is now abundantly clear that a number of different physical effects that vary with time and location contribute to the properties of SEP events seen at 1 AU. These effects include: the magnetic topology and connection of solar eruptions such as flares and coronal mass ejections (CMEs), the relationship between interacting and escaping particles, the origin, structure, and obliquity of CME shocks, the nature of wave-particle interactions and the type of turbulence that is present near the shocks, the availability, distribution and composition of the suprathermal seed populations, the type of injection and acceleration processes involved, and the possible modification of SEP properties (e.g., time-profiles of Fe/O, intensities etc.) due to transport-related effects.

By making new measurements in the inner heliosphere from ~10 R_s up to ~0.8 AU, future Heliophysics missions such as ESA's Solar Orbiter (SO) and NASA's Solar Probe Plus (SPP) will provide critical breakthroughs regarding SEP acceleration and transport. However, to fully distinguish between the effects of the above processes and achieve closure with models that seek to predict key SEP properties near Earth and beyond, we require coordinated measurements of CMEs and their shocks, the suprathermal seeds and the accelerated SEPs in conjunction with observations of the EM radiation signatures from the flares and CMEs on a single spacecraft located much closer to Earth.

To make these critical measurements, the Particle Acceleration and Transport in the Heliosphere (PATH) mission will be located at the L1 Lagrange point near 1 AU and designed to address three main science topics: A. the origin of the suprathermal seed population; B. the origin and acceleration of SEPs; and C. the role of transport in modifying SEP signatures. PATH is intended to provide comprehensive, stand-alone measurements that bring closure to these science objectives. However, we note that if PATH is operational during the rise-through-maximum of solar cycle 25 (~2019–2023), then together with SO and SPP, PATH will provide unprecedented, radially distributed measurements of the same physical phenomena and address the critical questions of how propagating structures, suprathermal tails and SEP properties evolve en route to Earth. PATH will focus on:

1. Observations of ion charge states across a wide range of energies;
2. Observations of the suprathermal seed population with improved time cadence and sensitivity over previous missions;
3. Observations of neutral (x-rays, γ-rays, neutrons, and energetic neutral atoms) signatures in flare/CME events with improved sensitivity and energy resolution;
4. SEP observations up to several hundred MeV/n; and
5. Vital contextual data (plasma, electric, and magnetic field observations).

The charge state measurements (1) are absolutely necessary because they provide one of the best remote diagnostics of the plasma conditions at the acceleration site (e.g. Geiss et al., 1995 and Klecker, et al., 2007). The SEPs are presumed to be accelerated from the suprathermal tail, rather than the bulk plasma, but these seed particles (2) have never been measured with a cadence that allows understanding of how changes in the tails correlate with the high energy data. Neutral particles (3), which include very energetic neutral atoms, (as observed by Mewaldt et al. 2009), neutrons, and x-rays and γ-rays, give us a direct channel into the acceleration region. SEP observations (4) provide the end result of the acceleration. Since these observations are being made at 1 AU, whereas most of the acceleration happens much closer to the Sun, the transport of the SEPs is impossible to understand without contextual data (5) on the solar wind plasma and magnetic field.

In addition to the prime science goal, the PATH instrument complement will also provide new data to understand the origin of particles accelerated elsewhere in the heliosphere (CIRs, ESPs), as well as provide Space Weather monitoring that is vital to the nation's infrastructure.

The PATH mission comprises sophisticated, heritage-based instruments that measure (1) the distribution functions and composition of the solar wind plasma and energetic particles with comparable or better sensitivity and resolution compared with existing instruments, (2) charge-state composition of suprathermal and energetic particles with significantly improved sensitivity compared with current instruments, (3) magnetic and electric fields at sufficiently high cadence and sensitivity to characterize the plasma and magnetic field waves and turbulence in the interplanetary medium and near shocks, (4) high cadence, high sensitivity measurements of x-rays, γ-rays, and neutrons to provide the energy spectra and timing of solar flares. Finally, the PATH mission concept includes a real-time downlink capability that provides observations of the interplanetary magnetic field, solar wind, and energetic particles to the Space Weather Community.

A reasonable implementation of PATH is a spinning spacecraft pointed toward the Sun at the L-1 Lagrange point, similar to ACE. This configuration also makes communication easy, utilizing a large

antenna on the earthward deck, and also allows the possibility of a Real Time data stream similar to ACE that can be used as a Space Weather station.

2. Science Measurements and Implementation

Summarized below are strawman instruments for PATH that collectively provide measurements of a) the energy spectra, the elemental, isotopic, & charge state composition, and the anisotropy of ions and electrons covering an extended energy range from ~few eV up to 100s of MeV/nucleon; b) the very energetic neutral atoms with temporal, spectral and spatial resolution; c) the energy spectra and timing of neutrons, X-rays and γ-rays from solar flares, d) the solar wind ions and electrons, and e) the DC and AC electric and magnetic fields. These measurement capabilities also provide real-time observations of the solar wind, interplanetary magnetic field (IMF), and energetic particles to the Space Weather community.

2.1 Charge State and Composition of Solar Wind, Suprathermal, and Energetic Ions

High sensitivity, high time resolution measurements of the elemental, isotopic, and charge-state composition of solar wind, pickup, and suprathermal H-Fe ions from ~10s eV/e up to ~50 keV/e. Instruments that employ *E/q* selection techniques, followed by a time-of-flight versus residual energy analysis in solid-state detectors like Ulysses/SWICS and ACE/SWICS [Gloeckler et al., 1992; 1997] are required to make these measurements.

High sensitivity measurements of the differential fluxes of H – Ni ions covering the energy range from ~few keV/q up to ~100s of keV/q to determine the elemental and isotopic (e.g., ^{3}He/^{4}He ratio) composition and the ionic charge states of suprathermal Fe in small and large individual solar energetic particle events. Instruments like Wind/STICS (Gloeckler et al., 1995) and ACE/SEPICA (Möbius et al., 1997) can provide these measurements.

2.2 Very Energetic Neutral Atoms

The fact that there is only one flare during which very energetic neutral atoms (VENAs) are observed (Mewaldt et al., 2009) indicates that only rarely does a flare produce VENAs at a location from where they can escape without being ionized and/or they are narrowly beamed. Yet they provide a very intriguing and unexpected insight into SEP acceleration processes. An instrument specifically designed to observe VENAs could consist of a large area particle detector with a narrow field of view aimed at the sun. This could possibly be combined with either the γ-ray instrument, neutron detector, and/or the energetic particle instrument.

2.3. Neutrons, X rays, & γ-Rays

High sensitivity, good spectral and imaging resolution of neutrons and γ rays to probe the acceleration processes at play directly at the Sun. Neutrons would be measured on an event-by-event basis between 15-150 MeV (the spectral range appropriate for 1 AU) with angular resolution adequate to identify neutrons of solar origin. We envision an imaging neutron spectrometer based on the SOlar Neutron TRACking telescope (SONTRAC) that double (n,p) scatters in densely packed scintillating plastic fibers arranged in alternating orthogonal planes (Ryan et al. 1999). Because it is compact, it would have several times the sensitive area of COMPTEL, the most sensitive such detector to have flown. The imaging property would greatly reduce the local neutron background effects. For γ rays >10 MeV, we require an effective area of order 50 cm^{2} and 20% energy resolution. To measure γ-rays from 0.2 to 10 MeV one or two ganged $LaBr_3$ scintillators would suffice. By adding a thin window to one of the two detectors, we can incorporate x-ray measurements down to 20 keV (just above most thermal emission).

2.4. Suprathermal-through-Energetic Ions and Electrons

Differential fluxes of H – up to Ultra-Heavy nuclei covering the energy range from ~few 10s keV/nucleon up to ~100s of MeV/nucleon to determine the elemental and isotopic composition (e.g., ^{3}He/^{4}He, ^{22}Ne/^{20}Ne ratios) in small and large individual solar energetic particle events. Such measurements can be provided by instruments that employ two types of techniques, namely, (a) time-of-flight versus residual energy technique as in ACE/ULEIS (Mason et al., 1997) and STEREO/IMPACT/SIT (Mason et al., 2006), and (b) dE/dx versus residual energy technique in

instruments like ACE/SIS (Stone et al., 1997b) and STEREO/IMPACT/LET & HET (Mewaldt et al., 2006).

Differential fluxes and anisotropy of ions and electrons between ~2 keV–10 MeV energy at a time resolution of ~1 minute or less to distinguish between particles arriving from the Sunward and anti-sunward directions. These measurements can be made with instruments like ACE/EPAM [Gold et al., 1998] and STEREO/IMPACT/LET & HET (Mewaldt et al., 2006) and will also be used to provide real-time energetic particle (RTEP) observations.

2.5. Contextual Solar Wind Plasma, Magnetic & Electric Field Measurements

Velocity distribution functions (VDFs) of solar wind protons and alphas in the ~0.1-20 keV/q energy range, and 3D VDFs of solar wind electrons in ~0.005-10 keV energy range at a time resolution of <1 second. Measurement of the interplanetary magnetic field is required for the primary science on PATH, but an ACE/MAG equivalent (dual fluxgate magnetometer with 24 vectors per second) is acceptable. A higher cadence and/or the inclusion of a search-coil magnetometer would improve the return on the secondary science goals. Antennae that give electric field/plasma wave measurements would also greatly enhance all five secondary science goals and the radio measurements are also important for remote sensing of acceleration regions and will improve our understanding of the origin of SEPs.

These measurements allow complete determination of the properties of waves, compression regions, CME shocks, and magnetic topology and are critical for understanding in-situ particle acceleration and wave-particle interactions. These data also provide real-time solar wind and IMF data to the community and can be made with instruments like Ulysses/SWOOPS [Bame et al., 1992], ACE/SWEPAM [McComas et al., 1998], and ACE/MAG [Smith et al., 1998].

3. Advances in Solar and Space Physics Science

The PATH Mission will yield significant advances in the following critical research areas.

Primary 1: Origin, Acceleration, and Transport of Solar Energetic Particles. PATH will provide comprehensive measurements of the differential fluxes, anisotropies, and composition (elemental, isotopic, and charge state) of ions, electrons and EM radiation over a continuous and extended energy range during a large number of CME and flare-related SEP events. Also, considerable numbers of protons, ions and electrons are accelerated in flares and are measurable through the neutrons, γ rays and x rays they radiate. The relationship between these particles and those measured in space is often unclear. This neutral radiation is not affected by transport effects and, it carries with it vital information about particle acceleration at the flare site. For example, γ rays and neutron measurements allow us to study the time evolution of different processes throughout the solar event and provide a bulk measure of the metallicity of the accelerated ions. Understanding the clear differences between particle populations accelerated during solar flares and by CME-driven shocks will provide the necessary ground-truth for SEP acceleration models that seek to specify and reliably predict the radiation environment near Earth and its impact on our space and ground-based assets.

Primary 2: Origin and Acceleration of Suprathermal Tails. Suprathermal ions between ~2 keV and 100s of keV are observed continuously at various locations throughout the heliosphere, but their origin and acceleration processes remain poorly understood. Observing and characterizing exactly how their key properties (e.g., intensities, distribution functions, composition etc.) vary on short-timescales ranging from hours to minutes, will provide much needed breakthrough for understanding the origin of this population and how it feeds into the mechanisms responsible for injecting and accelerating particles to even higher energies.

Secondary 1: Self-excited waves, Energy Partition, and Shock Dissipation Mechanisms. CME-driven interplanetary shocks dissipate their energy by heating the bulk plasma and accelerating suprathermal and energetic particles. The exact fraction of energy that is dumped into these different populations varies with radial distance and depends on several factors such as the availability and injection of suprathermal ions, the local structure of interplanetary shocks, the nature and influence of pre-existing turbulence, and the generation and subsequent dissipation of self-excited Alfvén waves.

Comparing properties of shock-associated particle events with CME and shock properties as a function of radial distance will differentiate between the above processes and ascertain the importance of their roles.

Secondary 2: Kinetic, Microphysical, and Turbulence-related Processes in the Solar Wind. Mechanisms responsible for heating and accelerating the solar wind include wave-particle interactions, turbulence-driven dissipation, shocks and compression regions etc. Studying the spatial evolution of the the solar wind at different radial distances in the inner heliosphere is necessary to pinpoint the responsible processes and distinguish between the influence of local effects such as plasma instabilities and the generation and dissipation of MHD turbulence from the remnants of coronal heating processes.

Secondary 3: Origin and Evolution of fast and slow solar wind streams, CIRs, ICMEs, and other solar wind structures. Determining the locations and properties of the source regions of the different types of solar wind and its embedded structures and understanding how they get modified during transit through the interplanetary medium remains a key objective of most Heliophysics missions. Comparing high-resolution solar wind and magnetic field measurements near Earth with properties of the remotely observed coronal structures will provide the elusive missing links between solar and heliospheric observations of the same phenomena.

Secondary 4: Particle Acceleration and Transport of CIR-associated Particles. Corotating Interaction Regions are quasi-stationary, large-scale heliospheric structures that occur primarily around the declining phase of solar activity cycle when fast solar wind streams overtake slower solar wind streams that were ejected earlier. Beyond Earth-orbit, such CIRs are typically bounded by MHD shocks that accelerate ions up to ~20 MeV/nucleon and electrons up to ~few MeV. Studying the properties of CIRs and their associated particles at 1 AU provides deep insights into the origin of the source populations, their efficient acceleration to ~10s of MeV in energy, the mechanisms that influence their subsequent transport back into the inner heliosphere, and the effects of time-dependent magnetic connection between 1 AU and the CIR shocks beyond Earth orbit.

Secondary 5: Origin of Upstream Particle Events near Earth. The origin and acceleration of energetic ions with energies ranging from a few keV up to 1–2 MeV and electrons up to ~100s of keV that escape from the vicinity of the Earth is highly controversial. Using sophisticated instruments that can measure the elemental abundances and ionic charge state composition along with the energy spectra and anisotropy will help unravel the source material and the physical mechanisms that are responsible for accelerating the upstream ions and electrons.

4. Cost Estimate

PATH can be implemented on a spinning spacecraft pointed toward the Sun at the L-1 Lagrange point, similar to ACE. This configuration also makes communication easy, utilizing a large antenna on the earthward deck, and also allows the possibility of a Real Time data stream similar to ACE that can be used as a Space Weather station. The payload is more extensive than ACE, and is more like STEREO in size and complexity. However, by extension from STEREO which was ~$520M (NASA and European) for two spacecraft (and assuming one spacecraft would cost 70 to 75 % of the two spacecraft cost), even with inflation the PATH spacecraft, payload, launch vehicle and mission operations and data analysis should be achievable within the budget of a small class mission (< $500M).

With the loss of Delta-II launch capability there is an uncertainty in launch vehicle availability for missions of this size. Clearly a Pegasus is too small, but this mission should be achievable without using a large rocket such as the Delta IV or Atlas V. Possibilities include the Minotaur, Falcon, or Taurus.

5. Evaluation Criteria for Heliophysics Decadal Survey

5.1 Identification of PATH measurements as a high priority or requirement in previous NRC studies and Roadmaps.

PATH addresses several fundamental questions of critical importance for Heliophysics. For instance, PATH directly addresses the goals and related science questions of two missions recommended in the Future Priority Science Queue in 2009 Heliophysics Roadmap. As stated, the central objective of Solar

Terrestrial Probes (STP #6) SEPAT – Solar Energetic Particle Acceleration and Transport mission is to "Understand how and where solar eruptions accelerate energetic particles that reach Earth", while the goal of the next Living With a Star mission (LWS #7) – Climate Impacts of Space Radiation – CISR after Solar Probe Plus is to "Understand our atmosphere's response to auroral, radiation belt, and solar energetic particles, and the associated effects on nitric oxide and ozone."

PATH also directly addresses four of the five challenges in the previous NRC decadal survey, entitled *"The Sun to the Earth – and Beyond: A Decadal Research Strategy in Solar and Space Physics (2002)"*

Challenge 2: Understanding heliospheric structure, the distribution of magnetic field and matter throughout the solar system, and the interaction of the solar atmosphere with the local interstellar medium.

Challenge 3: Understanding the space environments of Earth and other solar system bodies and their dynamical response to external and internal influences.

Challenge 4: Understanding the basic physical principles manifest in processes observed in solar and space plasmas.

Challenge 5: Developing a near real-time predictive capability for understanding and quantifying the impact on human activities of dynamical processes at the Sun, in the interplanetary medium, and in Earth's magnetosphere and ionosphere.

5.2 Contributions to more than one of the Panel themes.

Stationed at L1, PATH measures the solar wind, suprathermal populations, energetic particles, x-rays and γ rays, magnetic and electric fields, and cosmic rays to revolutionize our understanding of the dynamic and often violent interactions between the particle & plasma populations and the field environments in the heliosphere. PATH also provides real-time high cadence measurements of the IMF and solar wind data that are required for monitoring and predicting Earth's response to solar and interplanetary disturbances.

5.3 Contributions to important scientific questions facing solar and space physics today.

PATH contributes to two of the three critical science investigations associated with the three Research Focus Areas defined in the 2009 NRC Roadmap:

1. What is the composition of matter fundamental to the formation of habitable planets and life?
3. What is the magnetic structure of the Sun-heliosphere system?

5.4 Contributions to applications and/or policy making

PATH monitors the arrival and properties of all major causes (CMEs, shock, CIRs etc.,) of Space Weather at L1 and helps us understand the evolution and impact of radiation hazards from SEPs and GCRs to astronauts and technological assets. Similar measurements from ACE have proved to be of tremendous benefit, with the real-time ACE data being used in essentially all space weather forecasting models and centers. Further, the real-time energetic particle (RTEP) data are particularly relevant during large SEP and Ground Level Enhancement (GLE) events when solar particle populations not only precipitate in polar regions but can also reach lower latitudes, causing severe delays along commercial airline flights flying polar routes (e.g., New York to Tokyo).

5.5 Complement to other observational systems or programs available.

PATH makes the necessary near-Earth measurements to fully study the spatial origin, acceleration and evolution of solar wind and energetic particles as they propagate through the inner heliosphere and reach 1 AU. If operational during the rise-through-maximum of solar cycle 25 (~2019-2023), PATH observations of solar wind structures and SEPs at 1 AU will clearly augment and complement the corresponding SO and SPP measurements of the same phenomena in the inner heliosphere. PATH also measures the composition of different samples of matter, building on the successes of the ACE, Wind, and Ulysses. In addition, PATH fills in a critical between gap the current low-energy space-based observations and the high-energy neutron monitor observations of GLEs, making it possible to identify

the composition content of GLEs. Finally, PATH provides real-time solar wind monitoring and serves the needs of the Space Weather Community.

5.6 Cost-to-Benefit Ratio

PATH is a small-class (< $500M) mission stationed at L1 and has a highly favorable cost-to-benefit ratio since the new measurements will provide major breakthroughs in our understanding of how charged particles are accelerated to energies that produce harmful radiation, and will simultaneously provide real-time Space Weather monitoring capability.

5.7 Degree of technical, resource, and expertise readiness.

All the technology and expertise needed for PATH, including spacecraft systems, science payload, and mission operations at the L1 orbit either already exists or requires simple modifications to heritage instruments and existing sub-systems, thereby resulting in a low-risk, high return mission concept.

5.8 Fitting with other national and international plans and activities.

The major goal of the PATH mission is to determine the mechanisms responsible for particle acceleration and transport in the heliosphere, which concurs with a major objective of two of the most exciting Heliophysics Missions currently in development – SO and SPP – as well as that of the SEPAT mission (STP #6) described in the 2009 Heliophysics Roadmap. The 2009 Roadmap also recommended a launch date of 2021 for SEPAT, which is compatible with the amount of time (<10 years) required for the design, development, and launch of the PATH mission. Thus in terms of science objectives, estimated cost, and schedule, the PATH mission easily fits in with the future plans and ongoing activities of national and international space agencies.

6. References

Bame, S.J., D.J. McComas, B.L. Barraclough, J.L. Phillips, K.J. Sofaly, J.C. Chavez, B.E. Goldstein, and R.K. Sakurai, The Ulysses Solar Wind Plasma Experiment, *Astron. Astrophys. Suppl. Ser. 92*, 237-265, 1992.

J. Geiss, G. Gloeckler, R. von Steiger, H. Balsiger, L. A. Fisk, A. B. Galvin, F. M. Ipavich, S. Livi, J. F. McKenzie, K. W. Ogilvie and B. Wilken, The Southern High-Speed Stream: Results from the SWICS Instrument on Ulysses, *Science*, 268 (1995), 1033–1036

Gloeckler, G., Geiss, J. Balsiger, H.; Bedini, P.; Cain, J. C.; Fisher, J.; Fisk, L. A.; Galvin, A. B.; Gliem, F.; Hamilton, D. C., "The Solar Wind Ion Composition Spectrometer", Astronomy and Astrophysics Supplement Series (ISSN 0365-0138), vol. 92, no. 2, 267-289, 1992

Gloeckler, G., Balsiger, H.; Bürgi, A.; Bochsler, P.; Fisk, L. A.; Galvin, A. B.; Geiss, J.; Gliem, F.; Hamilton, D. C.; Holzer, T. E.; Hovestadt, D.; Ipavich, F. M.; Kirsch, E.; Lundgren, R. A.; Ogilvie, K. W.; Sheldon, R. B.; Wilken, B., "The Solar Wind and Suprathermal Ion Composition Investigation on the Wind Spacecraft", Space Science Reviews, Volume 71, Issue 1-4, pp. 79-124, 1995

Gloeckler, G. Cain, J.; Ipavich, F. M.; Tums, E. O.; Bedini, P.; Fisk, L. A.; Zurbuchen, T. H.; Bochsler, P.; Fischer, J.; Wimmer-Schweingruber, R. F.; Geiss, J.; Kallenbach, R., "Investigation of the composition of solar and interstellar matter using solar wind and pickup ion measurements with SWICS and SWIMS on the ACE spacecraft", Space Science Reviews, v. 86, Issue 1/4, p. 497-539 (1998)

Gold, R. E.; Krimigis, S. M.; Hawkins, S. E., III; Haggerty, D. K.; Lohr, D. A.; Fiore, E.; Armstrong, T. P.; Holland, G.; Lanzerotti, L. J., Electron, Proton, and Alpha Monitor on the Advanced Composition Explorer spacecraft, Space Science Reviews, v. 86, Issue 1/4, p. 541-562, 1998.

Klecker, B., E. Möbius, M. A. Popecki, Ionic charge states of solar energetic particles: a clue to the source, Space Science Rev., 130, 273-282, 2007.

Lanzerotti, L. J.; Gold, R. E.; Anderson, K. A.; Armstrong, T. P.; Lin, R. P.; Krimigis, S. M.; Pick, M.; Roelof, E. C.; Sarris, E. T.; Simnett, G. M., Heliosphere Instrument for Spectra, Composition and Anisotropy at Low Energies, Astronomy and Astrophysics Supplement Series (ISSN 0365-0138), vol. 92, no. 2, Jan. 1992, p. 349-363, 1992.

Mason, G.M., Gold, R.E., Krimigis, S.M., Mazur, J.E., Andrews, G.B., Daley, K.A., Dwyer, J.R., Heuerman, K.F., James, T.L., Kennedy, M.J., Lefevere, T., Malcolm, H., Tossman, B., Walpole, P.H., "The ultra-Low-Energy Isotope Spectrometer (ULEIS) For the ACE Spacecraft", Space Science Reviews, 86, 409-448,1998

Mason, G.M., Korth, A., Walpole, P.H., Desai, M.I., von Rosenvinge, T.T., Shuman, S.A., "The Spurathermal Ion Telescope (SIT) for the IMAPCT/SEP Investigation", Space Science Reviews, 136, no 1-4, 257-284, 2008

McComas, D.J., S.J. Bame, P. Barker, W.C. Feldman, J.L. Phillips, P. Riley and J.W. Griffee, Solar wind electron proton alpha monitor (SWEPAM) for the Advanced Composition Explorer, *Space Sci. Rev., 86*, 563-612, 1998.

Mewaldt, R.A., Cohen, C.M.S., Cook, W.R., Cummings, A.C., Davis, A.J., Geier, S., Kecman, B., Klemic, J., Labrador, A.W., Leske, R.A., Miyasaka, H., Nguyen, V., Ogliore, R.C., Stone, E.C., Radocinski, R.G., Wiedenbeck, M.E., Hawk, J., Shuman, S., von Rosenvinge, T.T., Wortman, K., "The Low-Energy Telescope (LET) and SEP Central Electronics for the STEREO Mission", Space Science Reviews, 136, no 1-4, 285-362, 2008

Mewaldt, R. A., Leske, R.A., Stone, E.C., Barghouty, A.F., Labrador, A.W., Cohen, C.M.S., Cummings, A.C., Davis, A.J., von Rosenvinge, T.T., Wiedenbeck, M.E., "STEREO Observations of Energetic

Neutral Hydrogen atoms during the 2006 December 5 Solar Flare", 2009, ApJ, 693, L11

Mewaldt, R. A., Leske, R.A., Shih, A.Y., Stone, E.C., Barghouty, A.F., Cohen, C.M.S., Cummings, A.C., Labrador, A.W., von Rosenvinge, T.T., Wiedenbeck, M.E., "Observations and Interpretations of Energetic Neutral Hydrogen Atoms from the Decmeber 5, 2006 Solar Event", 2010, in AIP Conf. Proc. 1216, Twelfth International Solar Wind Conference, ed. M. Maksimovic, et al., 592

Möbius, E., Hovestadt, D., Klecker, B., Kistler, L.M., Popecki, M.A., Crocker, K.N., Gliem F., Granoff, M., Turco, S., Anderson, A., Arbinger, H., Battell, S., Cravens, J., Demain, P., Distelbrink, J., Dors, I., Dunphy, P., Ellis, S., Gaidos, J., Googins, J., Harasim, A., Hayes, R., Humphrey, G., Kastle, H., Kunneth, E., Lavasseur, J., Lund, E.J., Miller, R., Murphy, G., Pfeffermann, E., Reiche K.-U., Sartori, E., Schimpfle, J., E. Seidenschwang, E., Shappirio, M., Stockner, K., Taylro, S.C., Vachon, P., Vosbury, M., Wiewesiek, W., Ye, V., "The Solar Energetic Paticle Ionic Charge analyzer (SEPICA) and the Data Processing Unit (S3DPU) for SWICS, SWIMS, and SEPICA", Space Science Reviews, 86, no 1-4, 449-495, 1998

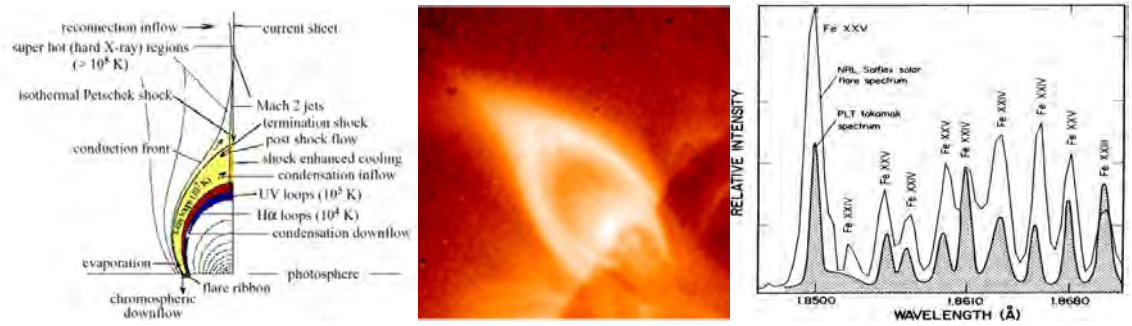
Ryan, J.M., Castaneda, C.M., Holslin, D., Macri, J.R., McConnell, M.L., Romero, J.L., Wunderer, C.B., "A Scintillating plastic fiber tracking detector for neutron and proton imaging and spectroscopy", NIM A, 422, issues 1-3, 49-53, 1999

Stone, E.C., Cohen, C.M.S., Cook, W.R., Cummings, A.C., Gauld, B., Kecman, B., Leske, R.A., Mewaldt, R.A., Thayer, M.R., Dougherty, B.L. Grumm, R.L., Milliken, B.D., Radocinski, R.G., Wiedenbeck, M.E., Christian, E.R., Shuman, S., von Rosenvinge, T.T., "The Solar Isotope Spectrometer for the Advanced Composition Explorer", Space Science Reviews, 86, 357-408,1998

A Concept White Paper for a New Solar Flare Instrument Designed to Determine the Plasma Parameters in the Reconnection Region of Solar Flares at Flare Onset

By George A. Doschek, Space Science Division, Naval Research Laboratory

Rationale: Solar flares are believed to be caused by magnetic reconnection that occurs in a current sheet located above the flares loops that subsequently are observed in soft X-rays. In order to test reconnection theories it is essential to determine the plasma properties within the reconnection region. This region is typically "dark" in soft X-rays and the conditions in it at flare onset are almost completely unknown at this time.



Furthermore, indications are from RHESSI observations that the plasma may be quite hot because the superhot component is seen in X-class flares before there is any sign of activity in the transition region and chromosphere (Robert Lin, Amir Caspi, private communication). I feel that the unique characteristics of the energy release process, via reconnection or some other process, can best be uncovered by observing the region as close to flare onset as possible. For example, polarization and turbulence measurements could be made of magnetic islands in the reconnection region, if they exist as discussed by Drake et al. 2006, Nature, 443, 553. The Fe XXV He-like lines and Fe XXVI H-like X-ray lines near 1.8 Å along with their associated satellite lines are the obvious spectroscopic targets for obtaining the observations. The presence of electron impact excitation lines along with dielectronic recombination lines offers unique plasma diagnostics not available in other spectral regions. The plasma parameters that can be measured are (Doschek et al. 1981, ApJ, 245, 315; Doschek & Tanaka 1987, ApJ, 323, 799; Henderson et al. 1990, Phys. Rev. Letters, 65, 705; Doschek 1990, ApJS, 73, 117):

Electron temperature
- From dielectronic recombination to resonance line ratios
- Independent of ionization equilibrium

Electron density
- From observed ionization times
- From ionization equilibrium conditions

Flows and turbulence
- From line profiles and wavelength shifts

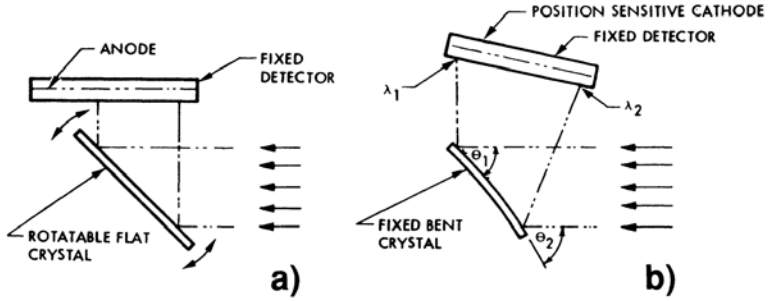
Polarization
- From intersystem and forbidden line to resonance line ratios in Fe XXV

Non-Maxwellian velocity distributions

From dielectronic and resonance lines

It is difficult to investigate the very onset of a flare. The rise phase of a flare is relatively short lived for most flares compared to their overall lifetimes, and instruments with limited fields of view, such as an imaging spectroscopy micro-calorimeter instrument, must get to the flare region at the onset. This is difficult because flare flags respond to an intensity enhancement in some spectral line or waveband, which cannot be set too low or otherwise any transient event will trip the flare flag. When a flare flag is tripped the instrument must be commanded to go to the flare site and enter a flare mode (high data rate) of operation. Even with a full-Sun imaging telescope like *Yohkoh*/SXT, automatic exposure control and flare mode activation took time and the true onsets of flares were not recorded unless the region was already being observed in flare mode by coincidence. Frequently the first few SXT images were overexposed. Furthermore, in flare mode the data rate of an imaging instrument is very high, which tends to saturate spacecraft computer storage. An instrument that requires a flare flag simply cannot make these initial observations, which may contain the strongest signatures of the initial energy release process.

The instrument: The experiment I have in mind is modeled after the Bragg Crystal Spectrometer (BCS) experiment on *Yohkoh* (Culhane et al. 1991, Sol. Phys., 136, 89).
The instrument described below improves on BCS by a factor of 10 in sensitivity. An order of magnitude improvement in any instrument parameter generally results in new discoveries. BCS is a bent crystal spectrometer that records its entire designed wavelength range at once all the time. Because flares evolve rapidly, it is desirable to observe all the diagnostic lines at the same time. This cannot be done with a flat crystal as the crystal rotates and scans wavelengths sequentially. The price for recording everything at once is a loss of crystal area for any given wavelength. On the other hand, a flat crystal may have a large crystal area, but the crystal can only spend a short time recording any given wavelength. Technically, the BCS is attractive because it contains no moving parts. However, a flat crystal spectrometer is technically attractive because flat crystals are much simpler than bent crystals, and also feed non-position sensitive proportional counters, which are simpler than the position sensitive ones required for a bent crystal spectrometer. This might outweigh a small lose in simultaneity with a flat crystal. Both flat and bent crystal spectrometer designs have essentially a NASA TRL level of 6-7, since they have been flown several times before on several different spacecraft.



The sensitivity of a bent crystal spectrometer is proportional to R_c, the integrated crystal reflectivity, the projected area of the crystal on the Sun, A_p, and inversely proportional to the range of angles, $\Delta\theta$, that it must cover to observe all wavelengths. The filter transmission, detector sensitivity, support blocking, etc., are not relevant for this discussion.

BCS contained four crystals, each covering a different X-ray multiplet. One of these crystals covered the He-like iron multiplet while another covered the H-like iron multiplet. Two BCS crystals fed into a single position-sensitive proportional counter detector. The entrances of the two spectrometer boxes were covered with aluminized Kapton filters to eliminate solar heating. For the current instrument, the following non-modifications/modifications are made to the BCS experiment:

1. Eight crystals cover the He-like iron complex, and eight crystals cover the H-like iron complex.

2. Ge 220 crystals (2d = 4.000 Å) are used instead of Ge 422 (2d = 2.310 Å) crystals for both multiplets, as in *Yohkoh*/BCS. For this strawman instrument, the BCS crystal dimensions are assumed for the new instrument.

3. For each iron multiplet, two of the crystals are oriented with the dispersion direction in the solar east/west direction, wavelength increasing westward. Two of the crystals are oriented with the dispersion direction in the solar east/west direction, wavelength increasing eastward. Two of the crystals are oriented with their dispersion directions in the solar north/south direction, wavelength increasing northward. Two of the crystals are oriented with their dispersion directions in the solar north/south direction, wavelength increasing southward.

4. Each crystal has its own detector.

5. There is no change in the entrance filter properties.

6. The spectrometer housings are made much more rigid than BCS. They cannot flex due to temperature variations etc. in orbit. The crystals all have to be accurately aligned such that their normals at a given point on the crystals point within 1”-2” of each other.

7. The wavelength range of the He-like Fe spectrometer is 1.835 – 1.894 Å. The wavelength range of the H-like Fe spectrometer is 1.7636 – 1.8044 Å.

8. The entire instrument package is designed to point to the center of a flare-producing active region and track it continuously across the disk, maintaining the pointing to within about ± 2” accuracy. The short and long wavelength limits are increased by about ± 5’ in Bragg angle to account for different positions of flares in the target active region.

9. The proportional counter detectors must have a much larger dynamic range than the BCS detectors.

10. A context *TRACE*-like imaging instrument is desirable but not necessary. It should image in the *TRACE* 171 and 195 Å bands.

The reasons and impacts of these modifications are given below, keyed to the numbers above:

1. A single "new" Ge 220 crystal is about 1.28 times more sensitive than a BCS crystal assuming the BCS physical dimensions for the new crystals (3.98cm x 18.1cm). By using eight crystals, an overall increase in sensitivity by a factor of 10.3 is achieved for the "new" instrument relative to BCS.

2. A comparison of the Ge 220 versus Ge 422 crystal shows that a Ge 220 crystal has about 3.1 times more sensitivity than a Ge 422 crystal. This overrides the downside that a polarization measurement is less complicated if Ge 422 is used because with Ge 422 the Bragg angles are close to 45 degrees. Crystal polarization can still be unfolded from source polarization with the Ge 220 crystal (Brown et al. 1989, Phys. Rev. A, 40, 4089. See also Henderson et al. 1990, Phys. Rev. Letters, 65, 705).

3. The purpose of the orientations is threefold: a) to distinguish Doppler shifts from positional shifts, (b) to provide the x and y coordinates for the flare on the Sun, and, (c) to make polarization measurements. The internal east-west, north-south reversals in dispersion allow Doppler and spatial shifts to be differentiated, and the use of east-west and north-south dispersion directions allows the x and y coordinates of the flare to be determined as well as enabling polarization measurements.

4. BCS used one detector for two crystals in order to save weight. This is complicated and deemed unnecessary for the new instrument. It is also a risk mitigation, since in BCS failure of a single detector kills two crystals.

5. The filter properties need not be changed.

6. Measurement of the flare x and y coordinates requires that all the crystals function as a calibrated unit in order to minimize the uncertainties in x and y. This requires a very rigid structure, as well as a method for aligning the normals of all the crystals. BCS was not too rigid due to weight concerns.

7. The wavelength range for the He-like iron lines in *Yohkoh*/BCS was considered excessive on the short wavelength side. Shorter wavelength ranges give more sensitivity, a desired goal of the new instrument, so I have modified the He-like iron wavelength range.

8. BCS was designed to observe the entire Sun. This necessitated taking a bigger wavelength range for the crystals than necessary to get a complete spectrum in order to account for different flare positions on the Sun. In the new instrument, the design is set so that this increase in wavelength range is brought to an absolute minimum in order to

increase sensitivity as much as possible. It would be possible to keep the instrument pointed at Sun center by increasing the wavelength ranges somewhat. However, because an east-west orientation for half the dispersion directions is necessary, the increase in wavelength ranges would be larger than for BCS, because BCS's dispersion was north-south and could be confined to the flare production belt on the Sun. This was ± 12'. For the new instrument I would recommend ± 20', in order to see flares above the east and west limbs.

9. It is desirable to have good spectra ranging all the way from flare onset to maximum emission measure in order to have good continuous diagnostic coverage of the reconnecting plasma. Thus the detectors cannot saturate when the count rates are large and/or suffer from excessive deadtime effects. BCS detectors saturated at large count rates.

10. The crystal spectrometer can locate flares and furthermore can provide tomographic images if the entire instrument package is rotated through small angles between spectrometer readouts. When the flare becomes bright enough, high resolution imaging could be provided by an AIA-like imager. The 192 Å Fe XXIV line in the TRACE 195 Å waveband could provide this. An image in Fe XXIV can be directly related to the Fe XXIV line called q in the X-ray spectra. This image would undoubtedly be obtained after flare onset, but this is OK since it is the plasma diagnostics at flare onset that I feel are most important and are supplied uniquely by the Bragg spectrometer.

Instrument Issues (Item 2 applies to the bent crystal design given here, but would not apply to the flat crystal design.)

1. Crystal alignment. It is not necessary to align all crystals to the precise same pointing direction, but it is necessary to know their offsets from one another to within 1-2". Otherwise, the flare location calculation will have too much ambiguity. All spectrometers housings should perhaps be mounted on a common stiff plate. This plate and the spectrometer housings should ensure no changes in relative crystal orientations due to launch and no excessive changes in relative crystal orientations due to temperature effects in orbit.

2. Crystal properties. Each crystal will differ from one another in bend radius, end effects, etc. This is OK provided we know the exact differences between the crystals. If these differences are not known, summed spectra will lose spectral resolution. John Seely at NRL has demonstrated in summing spectra from P78-1 that this problem can be solved.

3. Crystal fluorescence: Ge fluoresces at 11.1 keV, producing an unwanted false continuum. This can be removed by pulse height analysis in the detector electronics.

4. Detector deadtime and saturation. It is highly desirable to have a very large dynamic range.

I remind the reader that one could also design a flat crystal spectrometer package that could be as effective as the bent crystal spectrometer package described above. And as mentioned, the instrument just described has only very crude imaging capability. However, it provides the location of a flare on the Sun, and thus, even without plasma diagnostics, is a very effective flare flag for plasma with temperatures greater than 10 MK. I have not calculated the spatial resolution that might be achieved by rotating the instrument package through small angles between detector readout.

The High Resolution Solar-C International Collaboration: Probing the Coupled Dynamics of the Solar Atmosphere

George Doschek[1], Ted Tarbell[2], Ed DeLuca[3]
Saku Tsuneta[4], Toshifumi Shimizu[5], Yukio Katsukawa[4]
Sami Solanki[6], and Louise Harra[7]

[1]Naval Research Laboratory, [2]Lockheed Martin Solar and Astrophysics Laboratory, [3]Harvard-Smithsonian Center for Astrophysics, [4]National Astronomical Observatory of Japan [5]Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, [6]Max Planck Institute for Solar System Research, [7]University College London

1 Introduction

Many problems in solar physics, such as the heating of the solar chromosphere and corona, the onset of coronal mass ejections, and the origin of the solar wind, remain unresolved. This is because current and planned instruments do not provide comprehensive temperature coverage at the spatial, temporal, and spectral resolution needed to resolve and trace the elementary processes that dominate the energization of the solar atmosphere, such as magnetic reconnection, wave dissipation, and particle acceleration. During the past few years, space-based and ground-based observations and numerical models have provided tantalizing glimpses of the fundamental heating processes, while exposing crucial gaps in our arsenal of observational tools that have hampered progress for decades and that need to be filled to gain a full understanding of how the solar atmosphere is energized.

For the past year solar physicists in Japan, the United States, and Europe, many of whom were involved in the highly successful Solar-A (*Yohkoh*) and Solar-B (*Hinode*) missions, have been studying the possibility of next generation instrumentation that would make major steps forward in all of the necessary observational dimensions. Participation in such a large-scale, international collaboration led by the Japanese would leverage a relatively small investment from NASA to produce an extraordinary scientific return. The Japanese solar physics community is currently considering two mission concepts for Solar-C: an out of the ecliptic mission (plan A) and a high resolution mission (plan B). This white paper gives a brief overview of the science objectives for the high resolution (plan B) Solar-C mission and the instruments required to achieve them.

The unifying goal behind the high resolution Solar-C mission is to understand how the Sun's magnetic field drives mass and energy flows from the thermally dominated lower solar atmosphere into the magnetically dominated corona and solar wind acceleration region. This can only be accomplished with an instrument suite that *simultaneously* observes emission from the photosphere, chromosphere, transition region, and corona at *similar and high spatial, temporal, and spectral resolution.* Solar-C will be the first mission to reveal the spatial and temporal scales on which the energy release processes appear to aggregate according to recent observations and numerical simulations. Unlike previous instrumentation, Solar-C will be able to measure magnetic fields in and above the region where the field is dominated by plasma pressure, providing the first proper boundary conditions for coronal and heliospheric magnetic field extrapolations.

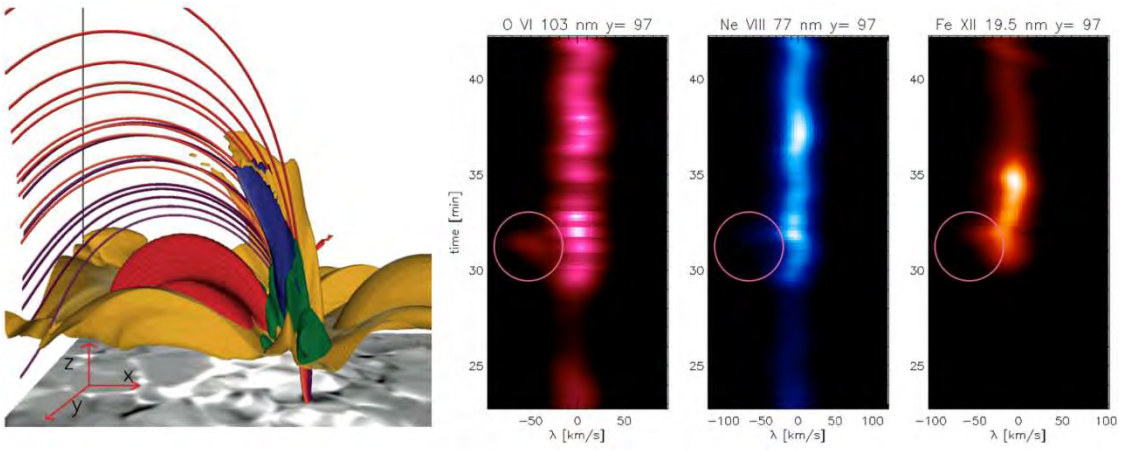


Figure 1: A 3D MHD simulation of a jet produced by emerging magnetic flux. This simulation illustrates the need for simultaneous imaging and spectral data covering the photosphere, chromosphere, transition region, and corona at high temporal, spatial, and spectral resolution. The left panel shows the magnetic field in the low chromosphere (grayscale image), magnetic field lines (red and purple lines), the location of the magnetic heating (green isosurface), the location of a chromospheric jet (blue isosurface), and the transition region and coronal emission (yellow and red surfaces). The right panels show the subtle velocity signature of the jet at three different temperatures as a function of time.

2 Science Goals

2.1 How does the magnetic field structure the solar atmosphere?

Low in the solar atmosphere gas pressure dominates the magnetic field. In the corona the situation is reversed, and plasma is generally frozen to the magnetic field. The $\beta = 1$ transition typically occurs in the chromosphere. Most previous instrumentation has focused on measuring magnetic fields in the photosphere, i.e., below the $\beta = 1$ surface, where the magnetic field properties can be determined relatively easily. This has led to enormous advances in our understanding of the photospheric fields, with *Hinode* providing unparallelled views of the fine-scale structure of flux tubes on spatial scales of 100 km ($0.2''$) and the processes leading to their formation, as well discovering ubiquitous weak fields on granular scales ($< 1''$), with potentially profound consequences for the energetics of the entire solar atmosphere (Lites, 2006).

To understand how the magnetic field drives coronal topology and large-scale field reorganizations in flares and CMEs and to establish the connectivity to the heliosphere, photospheric field measurements are used as a boundary condition for magnetic field extrapolations. But the photospheric field is not force-free, so the extrapolations fail to reproduce the observed topology of solar active regions (De Rosa et al., 2009). The forcing of the field in the chromosphere implies that much of the magnetic energy release occurs well above the photosphere, so tracing the release of free energy of the magnetic field and its effects on the momentum and energy balance of the chromosphere and corona using only photospheric fields is almost impossible.

Solar-C will tackle this problem by measuring the magnetic fields in the chromosphere, both around and above the plasma $\beta = 1$ surface (using, for example, the He I line at

1083 nm and the Ca II infrared triplet lines around 855 nm), with enough sensitivity and spatial resolution to provide meaningful boundary conditions for extrapolation methods. These measurements will reveal where and how magnetic free energy is released in the solar atmosphere. By simultaneously measuring photospheric and chromospheric fields, Solar-C will be able to measure the vertical gradients in the magnetic field, which will provide information on the Poynting flux into the solar atmosphere. This will yield important information on where and how the field is forced, and what role the magnetic fields play in heating and accelerating plasma in the corona and the solar wind.

Recent advances in numerically modeling the solar atmosphere, the impact of the Zeeman and Hanle effects on spectropolarimetric measurements, and the inversion of chromospheric lines to establish magnetic field properties are opening a new window into the magnetized chromosphere. This makes Solar-C particularly timely. Forward modeling with advanced numerical codes indicates that useful chromospheric measurements require a spatial resolution of 0.5″ or better and a signal-to-noise of order 10^3. For the optimal spectral lines, a meter-class optical telescope would be able to provide such measurements at a cadence of 20 s at one spectral position, which is adequate to capture the dynamics of the chromospheric fields, as evidenced by the chromospheric dynamics observed with *Hinode* (which occur on timescales of order 30 s).

2.2 What energizes and maintains the plasma in the complex magnetic environment of the solar atmosphere? — Measuring the signatures of magnetic reconnection, wave dissipation, and particle acceleration

Theoretical work tells us that the smallest spatial scales over which important physical processes such as magnetic reconnection, wave dissipation, and particle acceleration occur may be only meters in size — far smaller than will be resolvable with solar telescopes for the foreseeable future. Solar observations, however, suggest that these physical processes are coherent over much larger spatial scales. Chromospheric jets observed with *Hinode* and ground-based telescopes, for example, have spatial scales of about 0.1–0.3″ (about 70–200 km at the Sun). Similarly, observations of transition region and coronal structures taken with 1″-resolution instruments indicate that about 10% of the observed area is filled with emitting plasma, again suggesting spatial scales that can be observed with next generation solar instrumentation.

Even at the spatial and temporal resolution of *Hinode* there are clues to how the flow of mass and energy through the solar chromosphere can drive coronal activity. De Pontieu et al. (2009), for example, have found evidence of chromospheric jets (type-II spicules) that are heated to high temperatures and appear to supply mass to the corona. It is difficult to follow this chromospheric material into the corona, however, because of the mismatch between the optical, EUV, and X-ray instruments on *Hinode*. Observational challenges such as this drive the design of the spectroscopic and imaging instruments for Solar-C. The optical telescope will obtain intensity measurements of full spectral line profiles of chromospheric lines within 10 s to fully capture the dynamics of the chromospheric plasma. The spectrometer will combine UV observations of the chromosphere and transition region with EUV observations of the corona in a single instrument with very high spatial (0.2″), spectral (∼30,000), and temporal (1–10 s) resolution. This will be complemented by comparable spatial and temporal

resolution observations from narrow EUV passbands using normal incidence, multi-layer telescopes.

It is widely believed that magnetic reconnection plays a central role in many solar phenomena, such as flares and coronal jets. The physics of magnetic reconnection, however, is still poorly understood. For example, it is clear that magnetic reconnection on the Sun occurs much faster than the time scales predicted by the classic Sweet-Parker model, but what drives this fast reconnection? Recent theoretical work has suggested that it is the formation of magnetic islands in the current sheet that leads to fast reconnection (e.g., Bhattacharjee et al. 2009; Edmondson et al. 2010). To reconcile the observations with theory and numerical models, the physical conditions in and around the current sheet must be explored in detail. Of particular importance are the inflow and outflow velocities. Since the density in this region is close to that of the ambient corona and the volume is small, such measurements are very challenging. This motivates the selection of multiple wavelength ranges so that the brightest emission lines at critical temperatures can be observed spectroscopically.

The process of energy release in magnetic reconnection is also not well understood. Acceleration of particles to high energies is the dominant energy transport mechanism during the earliest phase of a flare. The mechanism responsible for this acceleration, however, has not been identified. One theory is that electrons gain energy from interacting with magnetic islands in the current sheet (Drake et al., 2006). Direct detection of energetic non-thermal electrons requires high energy spectroscopic observations (above 10 keV). The impact of these electrons on the chromosphere, however, can be directly measured by high resolution UV/EUV spectroscopy. The new, high-cadence observations from Solar-C will provide insights into the total energy carried by non-thermal particles. The magnetic tension in newly reconnected magnetic flux tubes is another source of energy. The gas dynamic shocks that form as these field lines slingshot away from the reconnection region can heat and compress plasma (Longcope et al., 2009). To understand this energy release process, the dynamics of post-reconnection flare loops must be observed (e.g., Savage et al., 2010). These observations will be possible only with the broad temperature coverage, high spatial resolution, and high cadence of Solar-C.

Solar-C will be able to establish the role of waves in providing the corona with energy and also exploit coronal seismology as a diagnostic of heating processes in the atmosphere. For example, we know that transverse magnetic waves carry enormous amounts of energy upward through the chromosphere (De Pontieu et al., 2007; Cirtain et al., 2007), but Solar-C's high resolution at a wide range of temperatures is required to trace the propagation, mode-coupling, reflection, and dissipation of these waves as they cross the $\beta = 1$ surface into the transition region and corona while resolving their short periods (as low as 60 s) and the displacement of the field lines ($< 1''$).

2.3 What drives large-scale restructuring and energy release in the outer solar atmosphere? — Predicting the onset of space weather

The primary science goals of the high resolution Solar-C mission, understanding the magnetic field throughout the solar atmosphere and the release of energy through magnetic reconnection, wave dissipation, and particle acceleration, lay the foundation for studying a wide range of complex solar phenomena. Of particular interest are phenomena that influence

the near-Earth environment, such as coronal mass ejections (CMEs), flares, and the solar wind. A critical goal for solar physics research is to improve our ability to predict space weather events. Predicting the occurrence of CMEs and flares requires understanding the storage and release of magnetic energy in the solar atmosphere. It is clear that these occur on very different time scales. Magnetic energy is built up slowly, through shearing motions along the neutral line. The release of energy, in contrast, occurs rapidly, but what actually triggers an eruption has not been firmly established.

One class of theories about CME initiation is tied to magnetic topology. In the "breakout model," for example, the overlying, coronal unsheared field is removed through reconnection at a magnetic null, allowing the stressed field along the neutral to erupt (Antiochos, 1998). In contrast, in the "tether-cutting" (e.g., Moore et al., 2001) and "flux rope" models (e.g., Forbes & Isenberg, 1991), the eruption is triggered by reconnection low in the solar atmosphere. In all cases the key to progress is relating signatures of energy release in different atmospheric regions to the magnetic structure of the solar atmosphere. Solar-C is flexibly designed to accomplish this task due to its ability to observe energy release and follow its transport through all regions of the atmosphere.

The solar wind is another important component of space weather. Coronal holes are the source of the fast solar wind (700–800 km s^{-1} at the earth's orbit). The origin of the slow solar wind (300–400 km s^{-1}) is unknown. Recently instruments on *Hinode* have detected persistent high speed outflows from large areas at the periphery of many active regions (e.g., Doschek et al., 2007; Harra et al., 2008; Del Zanna, 2008). These outflows are characterized by bulk shifts in the line profile of up to 50 km s^{-1} and enhancements in the blue wing of up to 200 km s^{-1} (Bryans et al., 2010; McIntosh & De Pontieu, 2009), and may lie on open field lines and contribute to the solar wind. Schrijver & De Rosa (2003), for example, have used magnetic field extrapolations to show that connections of active region fields to the heliosphere are common and suggest that up to 50% of the heliospheric field may originate in and around solar active regions.

Despite the extensive work on the outflows that has been done with *Hinode* a number of important questions remain. Perhaps the most important question that Solar-C can address is how these high speed outflows are heated and accelerated. Temporally resolved observations suggest that the outflows are composed of episodic events that originate low in the solar atmosphere. Because of the high velocities involved, however, the evolution of these events in space and time is difficult to follow. Solar-C will provide the observations necessary to understand how these high speed outflows are formed.

3 Mission Concept

To reach the science objectives outlined above, three state-of-the-art instruments have been designed in considerable detail: A Solar Ultra-violet Visible and IR Telescope (SUVIT), a UV/EUV High-Throughput Spectroscopic Telescope (here called EUVST), and an X-ray Imaging Telescope (here called XIT). These instruments have been well-studied by the Solar-C International Science Working Group and detailed descriptions are available in the Japanese Draft Interim Report; all instruments are at TRL 6. The planned orbit is geo-synchronous to accommodate a high data rate.

Solar-C is envisioned as an International Partnership Mission led by JAXA with substantial contributions from NASA, ESA, the UK, and several other European countries. It is designed to be a Small Strategic Mission from NASA's standpoint; the maximum cost to NASA for such a mission is $250M. This investment will ensure full participation in a large international mission, whose total cost may be of order $750M.

The high resolution Solar-C mission will make fundamental advances in many Research Focus areas listed in the 2009 Heliophysics Roadmap. It was specifically listed as a partnership mission to discover precursors to solar disturbances in the chromosphere and corona and is also key to the outstanding questions of heating and acceleration of the solar wind (Roadmap, P. 50). It is well-timed to provide crucial coordinated observations with Solar Probe Plus, Solar Orbiter, and the Roadmap mission STP #6, Solar Energetic Particle Acceleration and Transport.

SUVIT The primary science requirement of SUVIT is to determine the physical processes responsible for the dynamics of the photosphere/chromosphere and to measure chromospheric magnetic fields. This is done by making intensity and spectro-polarimetric measurements of photospheric and chromospheric spectral lines. From these measurements, physical parameters like temperatures, velocities, and magnetic fields (Zeeman and Hanle effects) will be determined. SUVIT is a meter-class telescope currently under study at 1.5 m.

With this diameter, the polarimetric sensitivity is designed to be 10^{-4} for $0.2''$ pixels and 20 s exposure times. In contrast *Hinode*/SOT and ground-based solar telescopes do not have sensitivities better than $\sim 10^{-3}$. In addition, compared to ground-based telescopes (e.g., ATST), Solar C will have stable, seeing-free imaging quality, which is crucial for high S/N polarimetry, so that the resulting long time series will uniquely allow us to track the evolution of the boundary condition for the coronal and heliospheric magnetic field. The spatial and temporal samplings are dictated by *Hinode* observations of chromospheric dynamics which shows intensity changes on timescales of order 20–30 s. The strawman design parameters are:

Telescope	Aplanatic Gregorian 2 mirror telescope
Primary mirror	1.5 m
Spectral lines (intensities)	Mg II k 2796 Å, Ca II 8542 Å, H I α 6563 Å He I 10830 Å, He I D3 5875 Å,
Spectral lines (polarimetry)	(chromosphere) He I 10830 Å, He I D3 5875 Å, Ca II 8542 & 8498 Å (photosphere) Fe I 8538 Å, Si I 10827 Å
Spectral resolution	100,000 (can observe a 3 km s^{-1} Doppler shift)
Spatial resolution	$0.1''$ at 2796 Å (intensity measurements), $0.2''$ pixels (polarization measurements), with S/N 10^4
Temporal resolution	0.1 s (imager, broadband filters), 1 s (imager, narrowband), 10–20 s (spectropolarimeter)
Field-of-View	$200'' \times 200''$
Focal Plane	Spectrograph Package and Imager Package

EUVST The primary science requirement of EUVST is to observe the flow and dissipation of energy from the top layers of the chromosphere into the transition region and corona, and to observe multi-million degree flare plasma. EUVST can observe all temperatures

from about 20,000 K to about 20 MK with high spatial, spectral, and temporal resolution. EUVST has higher spatial, spectral, and temporal resolution than *Hinode*/EIS as well as *SOHO*/SUMER. Its wavelength coverage in an amalgamation of EIS and SUMER wavebands. The strawman design parameters are:

Telescope	Single parabolic mirror on-axis system feeding an off-axis spectrometer
Grating	4000 lines/mm, elliptical substrate, VLS
Primary Mirror	40 cm
Spatial resolution	0.2″ pixels
Spectral ranges	130–250 Å, 510–630 Å, 690–810 Å, 940–1060 Å
Spectral resolution	40,000 near 1000 Å, 20,000 between 130 and 630 Å
Slit/slot sizes	slits: 0.16″, 0.32″, 0.96″; slots: 10″, 40″, 200″
Temporal resolution	1–10 s
Field-of-view	400″ × 400″
Detectors	5Kx2K CCD/APS

XIT The primary science requirement of XIT is to investigate the physics of magnetic energy storage and release in coronal plasma by observing how the corona reacts to activity that occurred in the transition region, chromosphere, and photosphere. Two X-ray telescopes are under consideration by the International Working Group: (1) a grazing-incidence telescope with photon counting detectors such that imaging spectroscopy can be achieved, and (2) an ultra-high spatial resolution normal incidence X/EUV telescope. In this document we discuss only the normal incidence telescope because it has a higher TRL level. The strawman design parameters are:

Telescope	Ritchey-Chretien form of Cassegrain
Primary Mirror	32 cm
Spatial resolution	0.1″ pixels
Wavebands	Wavebands at 94, 131, 171, 195, 211, 304, 335 Å, and in the UV are being considered
Temporal resolution	10 s
Field-of-view	400″ × 400″

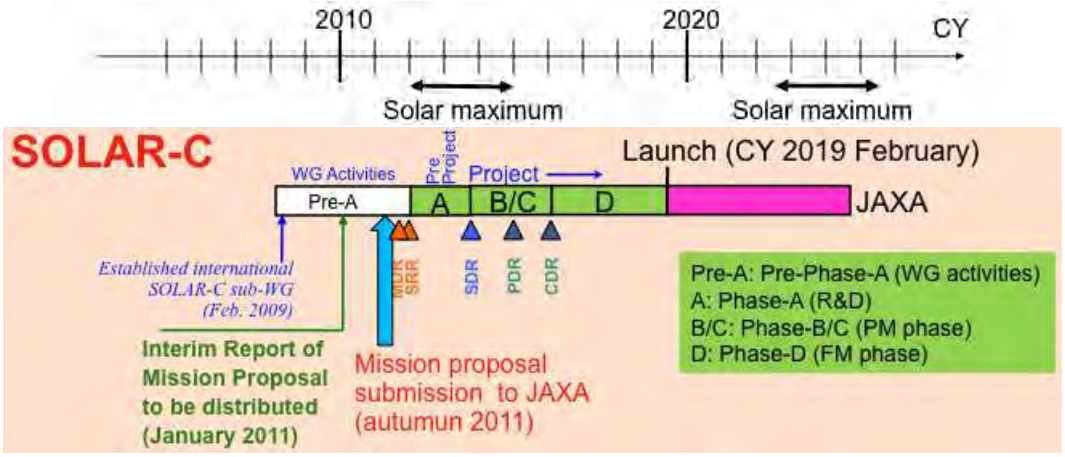


Figure 2: A preliminary schedule for the Solar-C mission.

References

Antiochos, S. K. 1998, ApJ, 502, L181

Bhattacharjee, A., Huang, Y., Yang, H., & Rogers, B. 2009, Physics of Plasmas, 16, 112102

Bryans, P., Young, P. R., & Doschek, G. A. 2010, ApJ, 715, 1012

Cirtain, J. W., et al. 2007, Science, 318, 1580

De Pontieu, B., et al. 2007, Science, 318, 1574

De Pontieu, B., McIntosh, S. W., Hansteen, V. H., & Schrijver, C. J. 2009, ApJ, 701, L1

De Rosa, M. L., et al. 2009, ApJ, 696, 1780

Del Zanna, G. 2008, A&A, 481, L49

Doschek, G. A., et al. 2007, ApJ, 667, L109

Drake, J. F., Swisdak, M., Che, H., & Shay, M. A. 2006, Nature, 443, 553

Edmondson, J. K., Antiochos, S. K., DeVore, C. R., & Zurbuchen, T. H. 2010, ApJ, 718, 72

Forbes, T. G., & Isenberg, P. A. 1991, ApJ, 373, 294

Harra, L. K., Sakao, T., Mandrini, C. H., Hara, H., Imada, S., Young, P. R., van Driel-Gesztelyi, L., & Baker, D. 2008, ApJ, 676, L147

Lites, B. W. 2006, in COSPAR, Plenary Meeting, Vol. 36, 36th COSPAR Scientific Assembly, 3626

Longcope, D. W., Guidoni, S. E., & Linton, M. G. 2009, ApJ, 690, L18

McIntosh, S. W., & De Pontieu, B. 2009, ApJ, 706, L80

Moore, R. L., Sterling, A. C., Hudson, H. S., & Lemen, J. R. 2001, ApJ, 552, 833

Savage, S. L., McKenzie, D. E., Reeves, K. K., Forbes, T. G., & Longcope, D. W. 2010, ApJ, 722, 329

Schrijver, C. J., & De Rosa, M. L. 2003, Sol. Phys., 212, 165

Far Ultraviolet Imaging of the Earth's Thermosphere and Ionosphere

1. Summary

The thermosphere-ionosphere (T-I) system is the interface between the Earth's atmosphere and the space environment. It is here that the Sun's most energetic radiation and the high-energy charged particles from the magnetosphere are absorbed. It is also here that waves propagating from the lower atmosphere finally dissipate. *Determining how the T-I system responds, on a global scale, to variations in these forcings is essential to our physical understanding of how the space environment is coupled to our atmosphere.* To understand the response to external and internal processes that drive the thermosphere-ionosphere, the spatial and temporal variations of the system should be measured simultaneously over the entire globe. Far ultraviolet (FUV) spectral imaging from appropriately spaced geostationary orbits accomplishes this, providing full disk images of atmospheric temperature and composition during the daytime and electron densities in the F2 region of the ionosphere at night. The ideal mission would use three imagers, operating simultaneously in orbits separated by ~120 degrees of longitude. This technique should be validated by initially flying a single imager. This document provides the scientific background for these missions (Section 2), their important practical applications (Section 3), some examples of the science that these missions would provide (Section 4) and an overview (Sections 5-8).

2. Scientific background

Most space missions that have investigated the Earth's T-I system have relied on measurements made along the track of a low Earth orbit (LEO) satellite, but spatial-temporal ambiguity is inherent to such measurements. Full-disk imaging of physical parameters such as temperature and composition, at a cadence of 1-2 hours, will remove this ambiguity. These observations will revolutionize our understanding of the Earth's upper atmosphere, just as global imaging has revolutionized our understanding of the troposphere.

Full-disk, FUV imaging provides an excellent means of obtaining such global measurements of the T-I system. The atmosphere at lower altitudes absorbs FUV light, providing a dark background for light emitted from the thermosphere. During the day, the ratio of emissions from atomic oxygen and molecular nitrogen provides composition information; during the night, imaging of the atomic oxygen emission provides ionospheric density and structure. Through moderate-resolution spectral imaging, daytime temperatures can also be obtained from the rotational distribution of the molecular nitrogen bands. High-altitude imagers such as DE-1 and IMAGE, as well as LEO spacecraft such as TIMED, have demonstrated the value of FUV imaging, and daytime temperature measurements have been successfully demonstrated using FUV data from the ARGOS satellite in LEO. Performing FUV spectral imaging from geostationary orbit would enable continuous global measurement of these key parameters.

Scientific and societal benefits would be maximized using three imagers separated by ~120 degrees of longitude around the equator, providing continuous imaging of the entire Earth. The observations made by these imagers would also provide context for LEO and ground-based observations. However, before a multiple imager mission is attempted, a single imager should be flown to demonstrate the scientific benefits of this concept.

3. Expected scientific benefits

The T-I system consists of a weak plasma subsumed in a more dense neutral gas. This mixture of plasma and neutral gas is unusual, as sometimes the plasma drives the neutral gas, while at other times the neutrals drive the plasma. It has proved difficult to understand the T-I region as a weather-like system because most space-based measurements, being from LEO, have an inherent spatial-temporal ambiguity. High inclination orbits observe at a roughly constant local time, whereas low inclination orbits only provide information at a ~90 minute cadence in a limited latitude range. The result of these samplings is an understanding of the T-I system as essentially a set of climate states rather than as a fully interacting weather system. Ground-based data provide important information about the T-I system, but are not sufficient to move our understanding of the region from a climate-like state to a weather-like state. Comprehensive numerical models describing the T-I system have been developed, but they require global measurements for validation and improvement.

FUV imaging observations of the entire Earth at a cadence of approximately one hour can provide the data necessary to achieve a comprehensive, weather-like understanding of the T-I region. Such observations would provide unprecedented information about coupling of the T-I system to the lower altitudes. FUV imaging will measure atmospheric tides across the entire dayside. A numerical model of thermospheric temperature changes, as viewed from a geostationary orbit, is shown in Figure 1. Such measurements have not been previously obtained for any region of the atmosphere, because the background signals mask tides in the troposphere and there is no method available to image them in the middle atmosphere. A simulation of the temperatures retrieved from observations using a FUV spectral imager is shown in Figure 2. FUV images provide the ability to separate migrating tides from non-migrating tides so that their temporal evolution can be determined directly, rather than being inferred from slowly changing observations at constant local times. This will allow the effects of planetary waves and other variability on the tides to be determined.

FUV imaging from geostationary orbit will also advance our understanding of how the T-I system responds to forcing from above. A key objective is a comprehensive study of the effects of geomagnetic storms on the thermosphere. Tantalizing glimpses of the large-scale changes due to these storms are seen in images from DE-1, IMAGE, and LEO satellites like TIMED. These data show evolving regions where molecular nitrogen densities increase relative to atomic oxygen, but they have not allowed separation of longitudinal and universal time effects during geomagnetic storms on a regular basis. Even less is known about the evolution of temperature during geomagnetic storms. Optical observations of temperatures have been limited to measurements from the AE and DE-2 satellites and some ground-based observations using interferometers. Some information has also been gleaned by observing satellite drag, but this can be difficult to interpret. A limitation of *in situ* observations, another important source of temperature measurements, is the difficulty in distinguishing storm effects from local time or longitudinal variation. The data from full disk FUV imaging will enable comprehensive analysis of the global changes that occur during storms. A numerical model of such thermospheric temperature changes, as viewed from a geostationary orbit, is shown in Figure 3.

A second objective is to better understand the effects of variable solar extreme ultraviolet (EUV) radiation on the T-I system. Solar EUV ionizes neutral atoms and ions or is absorbed as part of the process of dissociating molecules in the Earth's upper atmosphere. The question of whether thermospheric temperature variations are consistent with our knowledge of the ionization and dissociation processes and the variations in input solar radiation can be determined using global observations of neutral temperature. These observations will also enable the effects of solar flares on the thermosphere to be determined. Until now, our understanding of flare effects has been limited by our ability to separate spatial and temporal effects. Full disk observations of the thermosphere will allow these variations to be uniquely determined.

FUV imaging will also increase our understanding of ionospheric bubbles, a phenomenon of both scientific and Space Weather customer interest. These bubbles result from the Rayleigh-Taylor instability mechanism operating on the bottom of the F2 layer. The resulting instabilities expand vertically and along magnetic field lines, becoming bubbles. They occur after sunset within the equatorial ionosphere and are particularly prevalent between dusk and midnight. Our present understanding of bubbles is limited by the available observations, which are from either ground-based measurements or along the track of LEO satellites. These techniques have produced a wealth of information about individual bubbles, but they are not suitable for comparative bubble studies during the same set of geophysical conditions. Imaging bubbles from geosynchronous orbit will permit these comparative studies to be performed, allowing for the separation of longitudinal and universal time effects on bubbles, and providing a better understanding of the effects of geomagnetic activity (and penetration electric fields) on the generation of ionospheric bubbles.

4. Societal benefits

FUV imager observations from geostationary orbit have important practical applications. For example, this technique will enable observations of low-latitude ionosphere bubbles across the whole disk within 15 minutes or less. This will allow, for the first time, the study of their evolution over the entire visible disk. Such observations will not only aid in understanding their spatial and temporal evolution, but will also provide a short term forecast capability for determining communication and GPS outages over the visible disk.

Another practical application of an FUV imager is to provide sufficient data to drive assimilation models of the thermosphere and ionosphere. Successful assimilation models of the ionosphere have been developed and are currently used by the USAF for situational awareness. However, a shortage of thermospheric data has prevented these models from being deployed for the neutral atmosphere. A set of FUV imagers would provide an ideal data set for these assimilation models, enabling improved electron and neutral density specification and forecasts. In turn, better electron density forecasts would allow more reliable estimation of the likelihood of communication and navigation disruption, while better neutral density forecasts would improve estimates of the risks associated with space debris in LEO.

5. Methodology

The ideal FUV spectral imager would have two channels. One channel would perform routine scanning of the whole observable disk in the far ultraviolet, while the other channel would be used to either improve the overall sampling rate or for executing specialized campaigns. An example of the latter would be imaging small regions for bubbles in order to study their temporal evolution at very high cadence.

The FUV imager can also scan the limb to gather information about the height variations of O_2 density using stellar occultations. Bright stars are occulted sufficiently often to permit these observations to be made almost hourly, providing valuable data for understanding how composition evolves during geomagnetic storms and flares.

The imager can be hosted on commercial communications satellites. Such satellites are capable of providing the data stream in real time, which makes the data ideal for either direct application or for inclusion in assimilation models. Approximate costs are given in Section 7.

6. Straw-man mission

A proof-of-concept of the mission can be undertaken with a single imager. One imager can study, over most of one hemisphere, the spatial and temporal variations during geomagnetic storms as well as the effects of changing EUV radiation. A single imager would also observe upwardly propagating tides and enable the separation of their migrating and non-migrating components. In addition, it would allow bubble evolution to be studied on a global-scale, permitting an understanding of their temporal and spatial evolution that has not been possible previously. Flying three imagers would provide global observations and the more complete data set that assimilation models need to drive our theoretical understanding of the upper atmosphere.

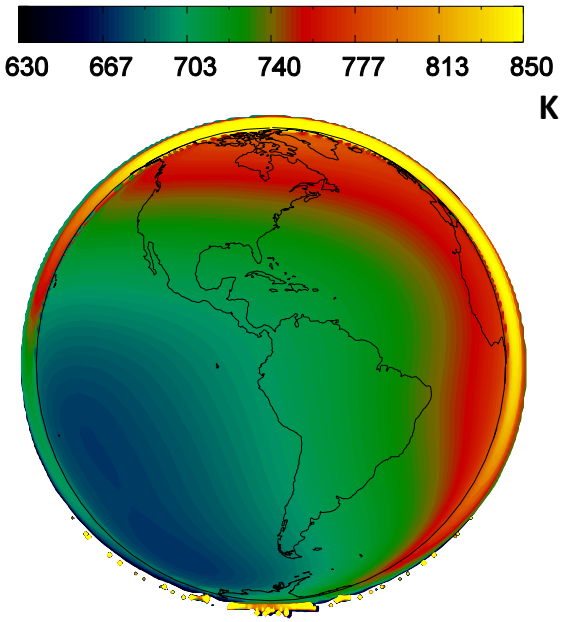
7. Cost estimate

The total mission cost for one imager would be approximately $50 million (FY11 dollars), including accommodation costs, instrument development and the science investigation. This includes the approximately $10 million cost per imager for accommodation on commercial communication satellites. Most of the remaining costs are for the development and fabrication of the imagers. Flying a complete mission, consisting of three identical imagers, would reduce the cost per imager.

8. Why are these studies important?

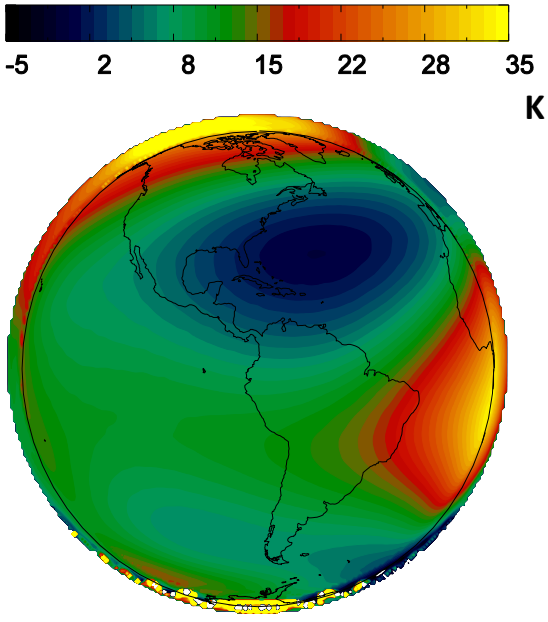
FUV imaging of the Earth's thermosphere-ionosphere from geostatonary orbit will provide breakthrough science. This poorly-understood region has characteristics that are part plasma and part neutral. Properly understanding how the neutral and charged elements interact is only possible if the temporal and spatial variations can be measured simultaneously, allowing the region to be treated as a weather regime rather than a set of climate states. Furthermore, data from dedicated LEO missions, flown in conjunction with the imagers, could be more readily placed into a global context.

Calculated Images of Temperatures & Temperature Changes Due to Tides



Temperatures Calculated using TIEGCM With Tides

(a)



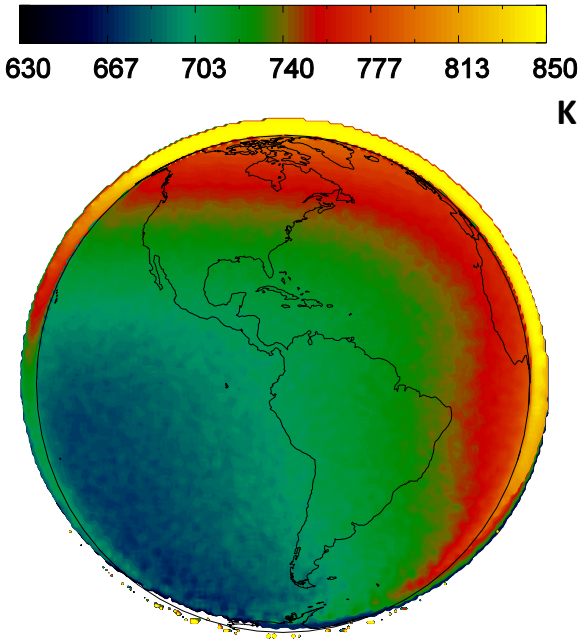
Temperature Difference TIEGCM With-Without Tides

(b)

- Temperatures and densities in the thermosphere are calculated using the TIEGCM model with and without tides from the GSWM model
- For lines-of-sight from geostationary orbit, temperatures and densities from the TIEGCM results are combined with emission profiles derived using AURIC to calculate the N_2 emission temperature that would be seen from geostationary orbit (a, shown on the left).
- Line-of-sight temperatures approximately represent the temperatures at 160 km and tides produce ~30 K perturbations in the calculated temperatures. Temperature change due to tides is shown on the right (b). Storm-time increases can be several times larger.

Figure 1

Retrieved Temperatures from Simulated Observations

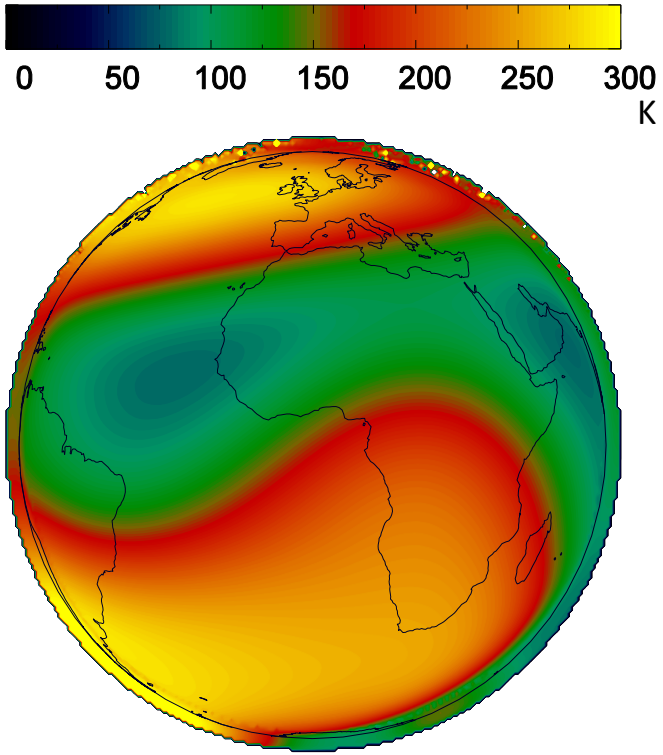


Retrieved Temperatures from Simulated Observations With Tides

- Simulated observation of tidal effects in the thermosphere. Observations are simulated by first calculating synthetic LBH spectra, running the spectra through an instrument simulator to account for instrumental effects (with instrument sensitivity derated by factor of 2) to generate simulated data, and finally retrieving temperatures from the simulated data.

Figure 2

Storm-time Increases in Neutral Temperatures



- Same as Figure 1b, but for storm-time temperature changes. Differences between the neutral temperatures, storm-time – quiet-time, are shown.
- Temperatures and densities in the thermosphere are calculated using the TIEGCM model with and without the storm. Then, for lines-of-sight from geostationary orbit, temperatures and densities from the TIEGCM results are combined with emission profiles derived using AURIC to calculate the N_2 emission temperature that would be observed.

Figure 3

UNCOVERING MECHANISMS OF CORONAL MAGNETISM VIA ADVANCED 3D MODELING OF FLARES AND ACTIVE REGIONS

Gregory Fleishman[1], Dale Gary[1], Gelu Nita[1], David Alexander[2], Markus Aschwanden[3], Tim Bastian[4], Hugh Hudson[5], Gordon Hurford[5], Eduard Kontar[6], Dana Longcope[7], Zoran Mikic[8], Marc DeRosa[3], James Ryan[9], Stephen White[10]

[1]*NJIT,* [2]*Rice University,* [3]*LMSAL,* [4]*NRAO,* [5]*University of California, Berkeley,* [6]*University of Glasgow,* [7]*Montana State University,* [8]*PSI,* [9]*University of New Hampshire,* [10]*AFRL/RVBXS*

ABSTRACT

The coming decade will see the routine use of solar data of unprecedented spatial and spectral resolution, time cadence, and completeness. To capitalize on the new (or soon to be available) facilities such as SDO, ATST and FASR, and the challenges they present in the visualization and synthesis of multi-wavelength datasets, we propose that realistic, sophisticated, 3D active region and flare modeling is timely and critical, and will be a forefront of coronal studies over the coming decade. To make such modeling a reality, a broad, concerted effort is needed to capture the wealth of information resulting from the data, develop a synergistic modeling effort, and generate the necessary visualization, interpretation and model-data comparison tools to accurately extract the key physics.

1. INTRODUCTION

Solar activity, although energetically driven by subphotospheric processes, depends critically on *coronal magnetism*, which, broadly speaking, includes magnetic field generation, morphology/topology, evolution, and transformation into kinetic, thermal, and nonthermal energies in the corona. While reliable direct diagnostic information has been lacking, during the next decade the situation must drastically change if we are to fully take advantage of the large wealth of high resolution data and modeling that is currently, or soon to be, available. New space- and ground-based solar optical telescopes are already capable of precise measurements of the photospheric magnetic field with sub-arcsecond angular resolution and high temporal resolution. Being combined with modern extrapolation algorithms, these data offer important clues on the coronal magnetic field structure and evolution.

Owing to their finite angular resolution, sensitivity, observational errors, and theoretical limitations, those extrapolations are not unique, so the extrapolations require independent verification. An opportunity for quantitative verification through radio coronal magnetography will be available during the coming decade when the new generation of high-resolution solar-dedicated radio instruments, including the expanded Owens Valley Solar Array (OVSA), upgraded Siberian Solar Radio Telescope (SSRT), and Frequency Agile Solar Radiotelescope (FASR), will become operational. There has also been progress made in obtaining the coronal magnetic field of active regions from advanced Stokes Polarimetry of infra red lines from the HAO CoMP instrument (Tomczyk et al. 2008). Microwave radiation from flares is produced by the gyrosynchrotron mechanism as accelerated fast electrons gyrate in the coronal magnetic field. It has recently been demonstrated using ideal (simulated) microwave data, that the coronal magnetic field can, in principle, be reliably recovered from the radio data on the flare dynamical timescale, along with the key parameters of the thermal plasma and accelerated electrons. The ability to detect the magnetic field and its changes on the dynamic time scales is a critically needed element to uncover fundamental physics driving solar flares, eruption, and activity.

2. MODELING COMPONENTS

Comprehensive modeling must include two closely related flows of effort: (i) *direct 3D modeling* based on our theoretical knowledge and constrained by available observations, which adopt some realistic physics and geometry and predict/calculate observables, and (ii) *robust diagnostics tools*, which achieve the complementary goal by starting from observables in order to quantify the geometry and physical processes.

Two major approaches are available to build the diagnostic tools: *true inversions*, which explicitly solve the inverse problem, and *forward fitting*, i.e., finding a number of free parameters of a physically motivated model of the system from fitting the model to observations. These diagnostics tools, being developed, will provide us with an array of the relevant coronal and flare physical parameters through a detailed sophisticated analysis of the data: e.g., radio imaging spectroscopy data for coronal plasma, accelerated electron, and field conditions and X-ray data for flare dynamics and energetics. We concentrate, here, on the solar flare modeling on dynamic time scales (down to $\sim$ 1 s or so); as a byproduct, most of the model components, described below, will also be applicable to the active region modeling, see also `http://lws-trt.gsfc.nasa.gov/trt_sc20063dar.htm`

Direct Modeling.

Very few fully 3-dimensional models of solar flares have been attempted up to the present time. Available models (Preka-Papadema & Alissandrakis 1992; Kucera et al. 1993; Bastian et al. 1998; Tzatzakis et al. 2008; Simões & Costa 2006; Fleishman et al. 2009; Simões & Costa 2010) have used simplified, generic magnetic geometries such as a dipole loop. In order to make contact with observational data, the next generation of models should employ realistic geometries from extrapolations or long-time-evolution MHD simulations based on vector photospheric measurements of the magnetic and velocity fields. As an example of complex concerted efforts leading eventually to the relevant 3D modeling, we itemize below a number of required major (but not all-inclusive) steps needed to develop realistic, interactive, and adjustable 3D models of solar flares.

Elements of direct flare modeling to be covered:

1. **A model of the pre-flare coronal plasma.**

   The magnetic field, density, temperature and elemental abundance of the pre-flare plasma is essential for modeling the subsequent solar flare. The basis of the magnetic field model would typically be a non-linear force free extrapolation from a photospheric or chromospheric vector magnetogram. The expected situation is for a low-beta hydrostatic equilibrium plasma to exist within this magnetic structure. The density and temperature structure would be set by heating determined either by an accepted form (Schrijver et al. 2006; Lundquist et al. 2008) or through comparison to EUV images. Alternatively a time-dependent, slowly-evolving simulation could be used to supply the pre-flare conditions.

2. **The energy release and flaring site.**

   According to current understanding, the rapid magnetic energy release powering a solar flare results form fast magnetic reconnection. Where in the magnetic field this process will occur, how much energy it will release and into which forms the energy will be converted will be the subject of active research over the coming decade. The modeling proposed here will provide critical input and constraints into these investigations. One approach will be to select flare-energized field

lines or flux tubes based on the observed locations of flare signatures, independent of theoretical considerations. Comparing the results of subsequent modeling to observations will then cast light on possible relations between the magnetic environment and reconnection energy release. Alternatively, a model for three-dimensional reconnection could be used to identify the field lines onto which energy is released. Research in recent years has suggested that non-fluid effects may play a critical role in triggering reconnection (see for example, Birn & Priest 2007). It might occur first where the width of the current sheet has decreased to a scale comparable to the ion skin depth, permitting kinetic effects to become effective.

The largest flares tend to be associated with dynamical eruptions called coronal mass ejections (CMEs). The exact relationship between the flare and the CME is still the subject of ongoing research, but it is generally agreed that the magnetic configuration is far from equilibrium at the time of the flare.

3. **Dynamics of the thermal plasma in the flaring loop** (or in the flaring loop system), i.e., prescribing the evolving electron number density, elemental/ion composition, and temperature to each voxel.

   The density and temperature may evolve in time as specified by hydrodynamic response to the flare energy release; ideally the loop will be embedded in a global (pre-flare) coronal model (1). At the level of more advanced modeling the inhomogeneous active region atmosphere is specified self-consistently with the magnetic structure, heating sources (including those driven by the flare), and cooling (see, e.g., Mok et al. 2008). In this case (in place of populating the voxels) the magnetic field model is self-consistently coupled with the thermal plasma distribution, so the data cubes describing the thermal plasma distribution must be imported into the flare modeling tool along with the magnetic field data cube.

4. **Populating the loop by evolving fast accelerated electrons**, which eventually must be determined from the time-dependent solutions of the transport equation (for flare modeling; not needed in case of active regions).

   The full problem is far from its final solution. The problem includes two major ingredients: particle acceleration and particle transport. Currently, there is no consensus about the main acceleration mechanisms operating in flares, although there are numerous models capable of successfully accounting for some of the observed properties of the accelerated particles. The modeling tools must be capable of accommodating arbitrary outcomes of (either analytical or numerical) external models describing the particle acceleration. Once the accelerator has been set up, the consequent transport mechanism is basically known: it is defined by the magnetic field line structure and thermal plasma, which is known within a given model. An additional but typically unknown important ingredient, which can affect the transport, is wave turbulence capable of angular scattering the particles. The presence and amplitude of such turbulence would follow either from the energy release model (2) or be set empirically using other observational input, such as spectral line widths.

5. **Calculation of the thermal and nonthermal** (HXR, gamma-ray, radio etc.) **emission characteristics** in each voxel and solution of the radiation transfer equations along all selected lines of sight, thus forming data cubes of the emission for the preselected viewing angle.

   These tasks, although based on the well understood and well developed theory, are frequently computationally demanding. So specific efforts for minimizing the computation time and fully optimized computing codes are critically needed.

6. **Developing powerful user-friendly visualization tools for the variety of model-derived 3D data cubes.**

   The visualization tools are needed to fully understand the structure and properties of the 3D objects under study. In addition, similar tools are needed to look at the imaging spectroscopy data as the relevant data volumes are often larger than those available from individual context instruments, and require spherical geometry, so that advanced tools are needed for data coalignment, viewing and analysis.

7. **Folding the data cubes through the instrumental response functions for direct comparison with observations.**

Most of the items mentioned represent major sub-projects within the program of the overall 3D modeling effort. The larger effort requires coordination of model standards and compatibility of the model data formats and consistently defined physical assumptions between the component parts, to enable these sub-projects to mesh seamlessly. Ultimately, these simulated models will be analyzed by the forward fitting or inversion tools, early versions of which are already in development (e.g., Fleishman et al. 2009). They are intended for reliably deriving the physical parameters of the emission region. This direct modeling, coupled with the forward fitting or true inversions, is a critically important step for validating the diagnostics tools prior to their application to real observational data.

Forward Fitting Diagnostics.

Robust diagnostics, understood as the determination of physical parameters of a system under study from arrays of observed parameters, is a key outstanding problem in Solar Physics. In some cases, e.g., in the hard X-ray (HXR) range, regularized true inversions can work well (e.g., Kontar et al. 2004). In most cases, however, true inversions fail because of the highly nonlinear nature of the physical systems they are trying to extract information from. In such cases, forward fitting can often be successfully used in place of true inversions. Although we specifically discuss below the forward fitting of the imaging spectroscopy of the microwave data, most of the discussion applies to imaging spectroscopy data in other wavelengths as well.

Assume that we have a sequence of spatially resolved spectra (both intensity and polarization data) from a solar flare (e.g., one spectrum per pixel). Then, we can fit the data to a model spectrum pixel by pixel to derive physical parameters of the source.

This forward fitting includes the following elements.

1. **Identify the model source function based on the radiation mechanism involved.**

   In the case of microwave emission from solar flares the emission mechanism is generally known: it is gyrosynchrotron (GS) emission with a free-free contribution in some cases (e.g., Bastian et al. 1998). Although the corresponding emission and absorption coefficients are known theoretically, the exact GS formulae are very computationally expensive and cannot be used in practice as the forward fitting input. Fortunately, much faster codes giving the same accuracy have recently been developed by Fleishman & Kuznetsov (2010). These codes, being fast, precise, and applicable for a broad range of regimes including anisotropic distributions of fast electrons imply a breakthrough in both 3D direct modeling and the forward fitting, allowing forward fittings of large bodies of data over reasonable time.

2. **Identify the fitting procedure resulting in fast and reliable finding of the true source parameters.**

   The problem is that most of the minimization algorithms often find a *local* minimum of the normalized residual (or of the reduced chi-square), while the ultimate goal of the fitting is to identify the *global* minimum. So far, we have determined that the simplex algorithm is very efficient in finding a local minimum. Then, it needs to be 'shaken' for the simplex solution to overcome any local minima and continue downhill towards the global minimum (a version of the stimulated annealing approach). Further efforts in optimizing the minimization algorithm are still needed especially for more complex cases when the number of the free model parameters is large.

3. **After-fitting inspection of the results.**

   Even when the algorithm performance is very good overall, there is a non-zero probability that the algorithm fails to find the true solution in some pixels. The post-processing must be able to identify and flag/remove those pixels.

4. **Interactive methods (similar to those used for the direct modeling) to deduce changes to model parameters based on observed mismatch.**

This sequential forward fitting must pave the way towards a global fitting in which a global source model is fitted to the whole body of the observational data (a multidimensional data cube containing spectra, light curves, and evolving spatial structure). Ultimately, the above procedures are iterated by quantitative means (to be determined), to adjust the model to match observations. Note that the model is adjusted to simultaneously match all available (multi-wavelength) observations.

3. THEORETICAL INPUT

The outlined modeling efforts require further development of the theory; the following key input is particularly important.

1. **Magnetic field extrapolations, magnetic reconnection, and energy release.**

   The present state of the art, non-linear force-free extrapolation from photospheric vector data, is known to suffer from inadequacies which must be overcome, or at least addressed, if we are to make progress in the coming decade. The photospheric field does not itself satisfy the force-free conditions used to extrapolate it. It must either be somehow modified or replaced entirely with vector field measurements from higher in the atmosphere (Metcalf et al. 1995; De Rosa et al. 2009). There are numerous other sources of information into the coronal magnetic field, such as EUV images (see Malanushenko et al. 2009), radio emission, coronal polarimetry (Lin et al. 2004; Tomczyk et al. 2008) or loop oscillations (Aschwanden et al. 1999; Nakariakov et al. 1999). These data are either indirect (EUV loops or oscillations) or in a form (sparse or integrated) not easily incorporated into the traditional formalism of extrapolation. Nevertheless, they constitute valuable data on the coronal field and should not be discarded entirely. Innovative techniques must be developed for incorporating such data into coronal field models. The modeling effort here proposed will, in the end, provide constraints on the structure of the coronal magnetic field and it would be valuable to be able to incorporate this back into the magnetic model itself.

   The very nature of upward extrapolation is poorly suited to identifying and resolving thin magnetic structure far from the boundary. Such structures, current sheets in particular, are

believed to play an essential role in the rapid energy release powering flares. In order to make genuine progress in understanding flares new coronal models must be developed which are capable of including these structures and their associated free energy.

2. **Turbulence generation, evolution, and parametrization of wave-particle interactions.**

   Turbulence is a highly important (but elusive for direct probing) element of the coronal plasma. It can play a key role in particle acceleration and transport and in the generation of electromagnetic emission. Since it often cannot be measured, the role of the theory here is crucial to provide the necessary relevant input for the modeling.

3. **Particle acceleration and transport including simplified fast solutions.**

   Particle acceleration remains an outstanding problem in solar flares. Any progress in understanding the acceleration mechanisms and how they work in realistic geometries will be exceedingly valuable for the modeling discussed above. The accelerator plays a role of the source function for the particle transport. Although the transport equations are generally known, their numerical solution is computationally demanding; thus, like in case with radiation coefficients, simplified and optimized (but still accurate) solutions are needed.

4. **Radiation processes including emissivity, absorption coefficients, and radiation transfer solutions; including fast codes.**

   Although many radiative processes are well studied, there is a need to optimize the computing codes for speed in many cases (e.g., Fleishman & Kuznetsov 2010). In addition, new emission processes have recently been identified, e.g., diffusive synchrotron radiation (Li & Fleishman 2009), which can give dominant contribution in the very site of the stochastic acceleration. Thus, both new regimes and mechanisms of emission and efficient computing codes must be developed to provide valuable input to the modeling.

4. EXPECTED RESULTS AND OUTSTANDING QUESTIONS

The modeling effort outlined above would have a broad impact on the field of solar astronomy. Its most immediate results would be in progress toward answering the following questions.

- Determination of the location of the energy release, the means by which magnetic energy is rapidly released, magnetic field reconfigurations, and the mechanisms of particle acceleration.
- Quantitative verification of magnetic field extrapolations, and refinement of extrapolation methods.
- Quantitative understanding of the hydrodynamic response of the atmosphere to energy input.
- Quantitative understanding of accelerated particle distributions, including energy, pitch angle, and relative proportions of thermal/nonthermal partition.
- The role of turbulence and wave-particle interactions on transport of particles and energy.

5. BROADER IMPLICATIONS

The modeling efforts we have outlined can only bring fundamental knowledge about flare/active region physics if used in conjunction with high-resolution modern observations. Key observations of the coronal plasma parameters can only be made by radio instruments that combine high sensitivity, temporal, spatial, and spectral resolution, which are unavailable now. A small part of the required science will be possible soon with the upgraded OVSA instrument (the funding of the upgrade just [10/01/2010] started; anticipated start of the upgraded instrument operation is fall, 2013). However, the full required capability has to await the completion of the full FASR, see Concept Papers submitted by Bastian et al., Gary et al., White et al.

6. CONCLUSIONS

As is clear from this Concept Paper, sophisticated development of three elements will be needed over the coming decade to significantly improve our understanding of the coronal magnetism, turbulence generation, and particle acceleration: (i) direct 3D modeling; (ii) theory; and (iii) iterative methods and tools for analysis of new, high-resolution, observations through forward fitting.

REFERENCES

Aschwanden, M. J., Fletcher, L., Schrijver, C. J., & Alexander, D. 1999, ApJ, 520, 880

Bastian, T. S., Benz, A. O., & Gary, D. E. 1998, ARA&A, 36, 131

Birn, J. & Priest, E. R. 2007, Reconnection of magnetic fields : magnetohydrodynamics and collisionless theory and observations, ed. Birn, J. & Priest, E. R.

De Rosa, M. L., Schrijver, C. J., Barnes, G., Leka, K. D., Lites, B. W., Aschwanden, M. J., Amari, T., Canou, A., McTiernan, J. M., Régnier, S., Thalmann, J. K., Valori, G., Wheatland, M. S., Wiegelmann, T., Cheung, M. C. M., Conlon, P. A., Fuhrmann, M., Inhester, B., & Tadesse, T. 2009, ApJ, 696, 1780

Fleishman, G. D. & Kuznetsov, A. A. 2010, ApJ, 721, 1127

Fleishman, G. D., Nita, G. M., & Gary, D. E. 2009, ApJ, 698, L183

Kontar, E. P., Piana, M., Massone, A. M., Emslie, A. G., & Brown, J. C. 2004, Sol. Phys., 225, 293

Kucera, T. A., Dulk, G. A., Kiplinger, A. L., Winglee, R. M., Bastian, T. S., & Graeter, M. 1993, ApJ, 412, 853

Li, Y. & Fleishman, G. D. 2009, ApJ, 701, L52

Lin, H., Kuhn, J. R., & Coulter, R. 2004, ApJ, 613, L177

Lundquist, L. L., Fisher, G. H., Metcalf, T. R., Leka, K. D., & McTiernan, J. M. 2008, ApJ, 689, 1388

Malanushenko, A., Longcope, D. W., & McKenzie, D. E. 2009, ApJ, 707, 1044

Metcalf, T. R., Jiao, L., McClymont, A. N., Canfield, R. C., & Uitenbroek, H. 1995, 439, 474

Mok, Y., Mikić, Z., Lionello, R., & Linker, J. A. 2008, ApJ, 679, L161

Nakariakov, V. M., Ofman, L., Deluca, E. E., Roberts, B., & Davila, J. M. 1999, Science, 285, 862

Preka-Papadema, P. & Alissandrakis, C. E. 1992, A&A, 257, 307

Schrijver, C. J., De Rosa, M. L., Metcalf, T. R., Liu, Y., McTiernan, J., Régnier, S., Valori, G., Wheatland, M. S., & Wiegelmann, T. 2006, Sol. Phys., 235, 161

Simões, P. J. A. & Costa, J. E. R. 2006, A&A, 453, 729

—. 2010, Sol. Phys., 266, 109

Tomczyk, S., Card, G. L., Darnell, T., Elmore, D. F., Lull, R., Nelson, P. G., Streander, K. V., Burkepile, J., Casini, R., & Judge, P. G. 2008, Sol. Phys., 247, 411

Tzatzakis, V., Nindos, A., & Alissandrakis, C. E. 2008, Sol. Phys., 253, 79

This preprint was prepared with the AAS LaTeX macros v5.2.

The outer heliosphere – solar system's final frontier

V. Florinski, J. Heerikhuisen, D. Shaikh, G. P. Zank
University of Alabama in Huntsville

Introduction

In the summer of 2002 the twin Voyager 1 and 2 spacecraft, launched in 1977, had long since completed their primary mission to explore the giant outer planets of the solar system and were approaching the very edge of the solar system. With most of theirs instruments turned off and the spacecraft operating on a fraction of the original design power, the expectations for future science results were relatively modest. There was even talk that these two stalwarts of the NASA fleet should be retired. What happened instead was a chain of remarkable new discoveries that redefined the face of space physics for the next decade.

In 2002 Voyager 1 started to observed highly energetic ions arriving from larger heliocentric distances [*Krimigis et al.*, 2003; *McDonald et al.*, 2003]. The spacecraft had arrived in the vicinity of the inner edge of the heliospheric boundary, known as the solar wind *termination shock* (TS), whose existence had long been predicted by theoretical models [*Parker*, 1961], and was sampling the particles accelerated close to this immense plasma structure. The shock itself was crossed at the end of 2004 and in the mid-2007 by Voyager 1 and Voyager 2 respectively. While the shock was approximately where it was predicted to be, many of measurements of solar wind and energetic particles were in stark contrast with the models.

A major disagreement between theory and observations was Voyager's apparent failure to detect the source of *anomalous cosmic rays* (ACRs), which was believed to be the TS as a whole [*Pesses et al.*, 1981]. Energetic particle signatures were not consistent with steady first order Fermi acceleration at a strong collisionless shock [*Stone et al.*, 2005; *Decker et al.*, 2005]. Rather, the observed ions appeared to belong to two different populations, one (low energy) accelerated locally at the TS, and the other (high-energy ACR) accelerated elsewhere. The disagreement led to a major revision in our understanding of particle acceleration mechanisms in space plasmas (Section 3).

The termination shock itself and the region of heated shocked solar wind plasma beyond, known as the *heliosheath*, proved to be full of surprises. Far from a stationary structure, the TS was traveling with a speed of up to 100 km/s toward and away from the Sun, in response to variations of the dynamic pressure of the solar wind. Voyager 2 discovered that the TS was in a rare state known as an energetic particle-mediated shock possessing an extended upstream precursor [*Florinski et al.*, 2009]. The temperature of the bulk plasma increased by only a factor of 10 rather than ~100 that would follow from energy conservation [*Richardson et al.*, 2008]. The missing energy is evidently absorbed by *pickup ions* (PUI), former interstellar neutrals ionized in the solar wind via a process of charge transfer (Section 5).

Whereas the Voyagers are not instrumented to measure PUIs directly, their presence in the distant heliosphere is deduced via remote sensing of *energetic neutral atoms* (ENAs) produced in charge transfer collisions between ions and interstellar neutrals by the Earth-orbiting IBEX spacecraft and by Cassini in orbit around Saturn. During its first year of operation the IBEX team made a startling discovery of a ribbon of enhanced ENAs sources threading its way across the sky [*McComas et al.*, 2009]. One explanation is that the ring is composed of secondary ENAs produced from neutralized solar wind protons because it displays signatures (such as latitude variations) characteristic of the solar wind (Section 2).

Much insight was gained from observations of *galactic cosmic rays* (GCRs) in the outer heliosphere and heliosheath. These highly energetic particles are messengers from distant supernova shocks where they are believed to be accelerated. It is now revealed that GCRs spend much of their heliospheric transit time trapped in magnetic structures in the heliosheath, which plays a prominent role in modulating their intensities [*Florinski and Pogorelov*, 2009]. As the Voyagers approach the outer edge of the solar system, called the *heliopause*, the pristine interstellar spectrum of GCRs is being gradually revealed [*Webber and Higbie*, 2009]. The heliosheath is the last puzzle for a comprehensive theoretical model of GCR modulation in the heliosphere (Section 4).

More discoveries are in store for the upcoming decade. The Voyagers are expected to be operational until 2020, by which time both would have crossed the heliopause. Apart from the great symbolic value of first venturing outside our solar system, the heliopause itself should prove to be no less significant than the TS. The region could be the source of ACRs according to some theories, as well as a generator of turbulence via instabilities driven by charge exchange with interstellar hydrogen (Section 5). It is hardly necessary to emphasize the importance of a diverse theory program to analyze and explain past, present, and future observations. The global structure of the region, including magnetic field topology, needs to be modeled in detail based on 1 AU and in situ observations. The TS and heliosheath comprise a unique space plasma laboratory with conditions found nowhere else in the solar system, providing a broad field for theoretical investigation. An investment in an outer heliospheric theory program over the next decade will benefit multiple space-related fields from solar physics to astrophysics.

Elucidating the global structure of the heliosphere

Past achievements. The two Voyagers crossed the TS at a heliocentric distance of 94 and 84 AU, respectively, indicating a distinct North-South (and possibly East-West) asymmetry of the global heliosphere [Stone et al., 2008]. This effect has been primarily attributed to a tilt of the interstellar magnetic field relative to the velocity vector between of the solar system relative to the surrounding interstellar cloud [Opher et al., 2007, Pogorelov et al., 2007]. A second effect that plays a role is the motion of the shock toward the Sun as a result of a reduction in the solar wind dynamic pressure approaching the 2007-2009 deep solar minimum. Based on the knowledge of Voyager 1 crossing, a successful forecast of Voyager 2 crossing was made [Washimi et al., 2007]. The Voyager TS observations allowed to place theoretical constraint on the direction and magnitude of the interstellar magnetic field, one of the last unknowns of our local galactic environment (Figure 1a). Another unexpected result of the Voyager mission is the significantly faster heliosheath plasma speed measured at Voyager 2. The flow is now turning away from the radial direction as the plasma becomes deflected on approach to the heliopause. There is more deflection in the azimuthal direction than in the polar direction [Richardson et al., 2009].

Much progress has been made in explaining the origin of the IBEX ribbon and the related structure at higher ENA energies ("belt") uncovered by Cassini [Krimigis et al., 2009]. Ribbon ENAs could be produced beyond the boundaries of the heliosphere, in the interstellar space around the heliopause from PUIs of solar wind origin, or closer to the Sun, upstream of the TS from shock-reflected ions. The leading theory [McComas et al., 2009], quantified in Heerikhuisen et al. [2010] currently explains both the intensity

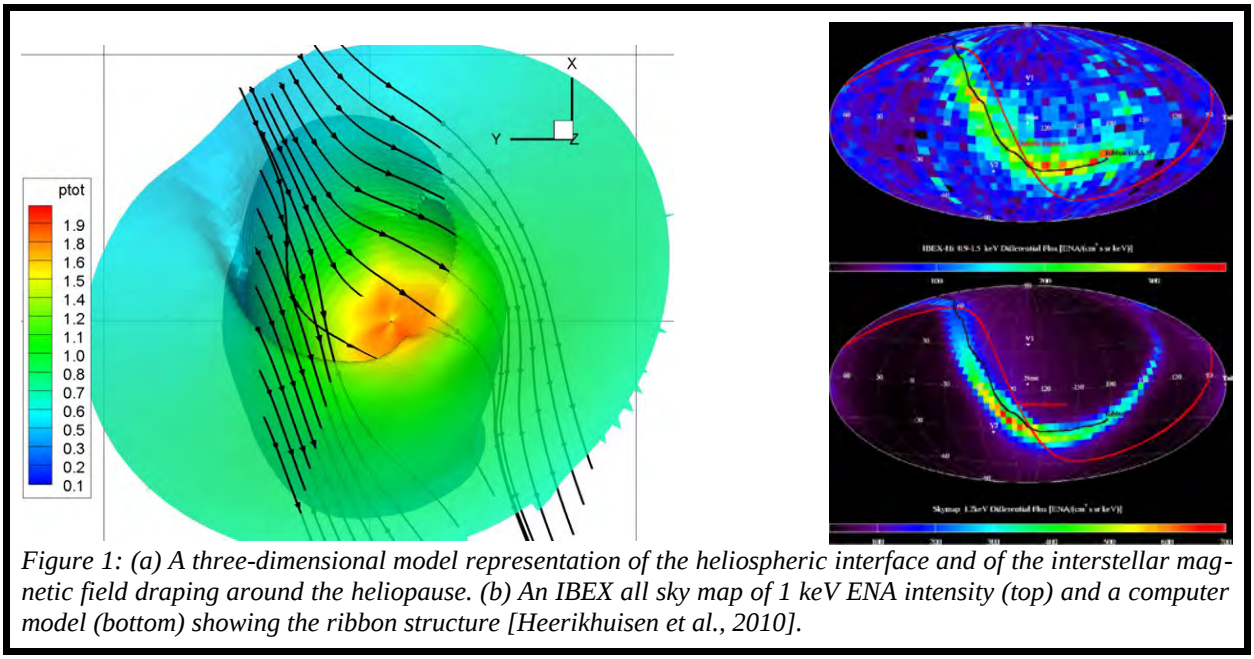


Figure 1: (a) A three-dimensional model representation of the heliospheric interface and of the interstellar magnetic field draping around the heliopause. (b) An IBEX all sky map of 1 keV ENA intensity (top) and a computer model (bottom) showing the ribbon structure [Heerikhuisen et al., 2010].

and the geometry of the ribbon (Figure 1b), although outstanding issues remain. It also provides an important clue to the direction of the magnetic field in the local interstellar space.

Case for future study. A synergy of Voyager, IBEX, and Cassini to study the remotest regions of the solar system cannot be overstated. The New Horizons space probe, currently en route to Pluto, could potentially compliment Voyager data from the direction of the heliospheric tail. We need to continue coordinated efforts to elucidate the global structure of heliospheric boundaries using both in situ and remote observations. A strong and consistently funded theory and simulation component should be an essential part of this work. Efforts to develop a complete self-consistent computer representation of the heliosphere that includes both the solar wind plasma and the interstellar neutrals, should be a priority target for funding in this field.

Much work is still needed to reconcile most of spacecraft observations with a single time-dependent global model of the heliosphere and the surrounding interstellar medium. A crucial component of this effort is the development of more accurate (kinetic) representations for interstellar and heliospheric atoms and for pickup ion populations. A more remote goal is the eventual coupling with the models of the inner heliosphere that would provide a time-dependent input of solar wind data to the models of the outer heliosphere. Computational models of the global heliosphere are in many ways still in their infancy. But in the same way that continued development of space-weather models – and terrestrial weather models in the decades prior – has greatly improved our understanding and predictive power, so too will global heliospheric models continue to improve over the next decade, provided such development continues to be supported. Due to the limited observational data, all the topics discussed in this paper rely on accurate simulations, so that continued support to develop models of the global heliosphere should be a corner stone of any outer heliospheric program.

A particle acceleration laboratory

Past achievements. The TS was traditionally considered the only source of ACRs [Jokipii, 1986]. Yet neither Voyagers collected evidence that the process was operating at the highest energies (up to 100 MeV) at the locations of their shock crossings [Stone et al., 2008]. A plethora of new theories were development in response to these puzzling results. It was recognized that the TS is an immense structure with a large degree of spatial and temporal variability. Unlike most interplanetary shocks, the TS can be both quasi-parallel and very oblique depending on latitude and longitude. Its compression ratio, speed, and magnetic field connection to other parts of the heliosphere all vary on timescales from hours to years.

If the TS is the source of ACR, then their acceleration occurs at sites not accessible to the Voyagers. The most promising locations are the flanks and the tail regions. These sites enjoy a longer magnetic field connection times to the downstream regions which allows particles to return to the shock multiple times to boost their initial energy [McComas and Schwadron, 2006; Kota, 2010]. The process is facilitated by the bluntness of the TS (Figure 2a). If ACRs are accelerated at the flanks, their Voyager-measured intensities would appear modulated, as if they were coming from larger heliocentric distances. It is also possible that time-dependent reformation of the shock caused some of the observed spectral anomalies.

Another intriguing possibility is that ACRs originate close to the heliopause, where turbulent conditions could also favor particle energization. Proposed mechanisms include stochastic acceleration and acceleration at reconnecting current sheets. The former process has been shown to produce power laws in momentum for a range of ion species similar to what is observed at Voyager 1 [Fisk and Gloeckler, 2009]. The latter is a result of a highly complex structure of the magnetic field that includes a rapidly oscillating magnetic field due to a piling up of bends in the heliospheric current sheet [Drake et al., 2010]. Reconnection leads to a formation of magnetic islands that could serve as particle acceleration sites (Figure 2b).

Case for future study. The origin of ACRs remains an outstanding problem in space physics. Observations currently do not permit us to eliminate any of the theories outright. The riddle may be answered when the Voyagers approach the heliopause, which will happen within the next decade. However, we also must acknowledge a possibility that the acceleration site will not be visited by any spacecraft. For this reason it is essential to make full use of available Voyager 1 and 2 data, including spectral, spatial, and

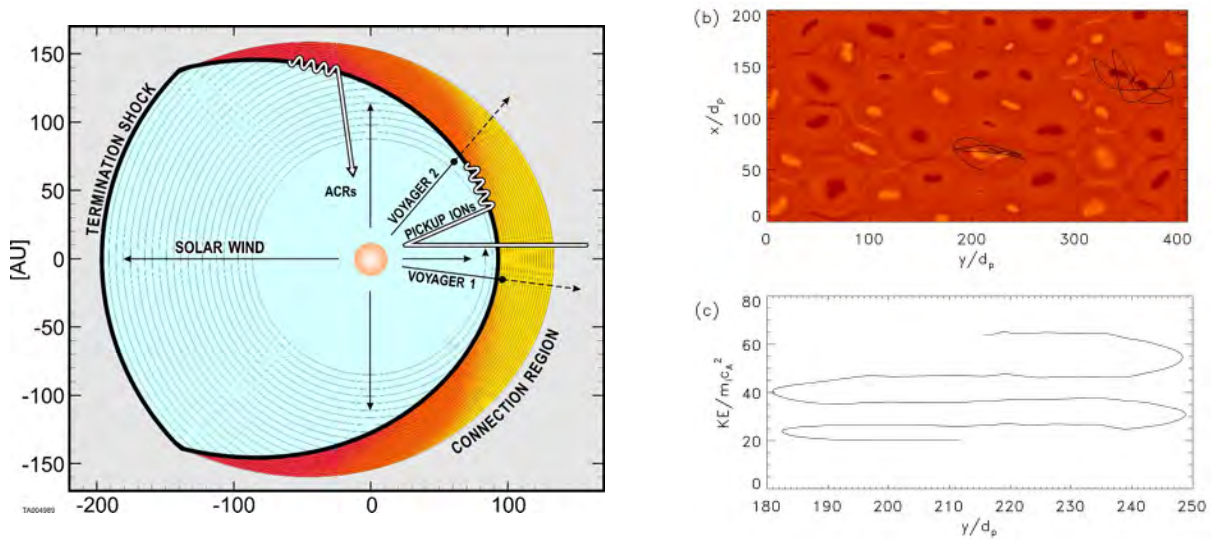


Figure 2: (a) A conceptual drawing of a blunt termination shock showing magnetic field connectivity. Acceleration is favored at the flanks (red) where the field stays connected to the shock for a long time [McComas and Scwadron, 2006]. (b) A simulation of a reconnecting current sheet leading to a formation of magnetic island that could serve as energetic particle accelerating sites [Drake et al., 2010].

temporal features, to reconstruct ACR acceleration history as accurately as possible.

Theoretical work should focus on developing working concepts of ACR acceleration capable of explaining the entire compendium of observational results. We need a realistic three-dimensional representation of the blunt TS and a model of injection of PUIs into the acceleration process at a quasi-perpendicular turbulent shock. A parallel effort should focus on elucidating the plasma and magnetic field conditions near the heliopause and on evaluating their effectiveness in accelerating charged particles to high energies. These modeling objectives tie in closely with those required to improve our understanding of the global structure. The TS and the heliosheath represent a unique particle acceleration laboratory with conditions not found anywhere else in the solar system. It is essential to make use of the opportunity to investigate the alternative acceleration mechanisms while we still have spacecraft taking in situ measurements in the region for the next decade.

Galactic cosmic rays in the heliosheath

Over the past decade Voyager observations have shown just how important the role of the heliosheath is in protecting the inner heliosphere from the ionizing cosmic radiation in the form of GCRs. This barrier effect is very significant at moderate GCR energies (below a few hundred MeV). Intensities in the heliosheath during the 2007-2009 extended solar minimum were about 50% lower than the model interstellar values. At the same time radial cosmic-ray gradients were small indicating that a stronger barrier effect at even larger distances, most likely close to the heliopause [Stone et al., 2008].

Much was learned from computer simulations about the pattern of GCR transport in the turbulent heliosheath. Figure 3a illustrates the very complex topology of magnetic field in the outer heliosphere and heliosheath. The oscillating neutral sheet is folded very tightly in the heliosheath and segments of the opposite polarity, corresponding to different solar cycles, are simultaneously present on the leeward (tail) side. We now have the capability to simulate a range of cosmic-ray trajectories in this very complex field using Monte Carlo simulations. A time history of galactic protons is illustrated in Figure 3b. Clearly these particles spend 3 to 5 times longer in the heliosheath than they do in the solar wind. The heliosheath thus serves as a long term storage for GCRs.

Case for future study. The next decade will finally reveal the unmodulated spectrum of GCRs at the edge of interstellar space. This is crucial for the models of GCR production and transport throughout the

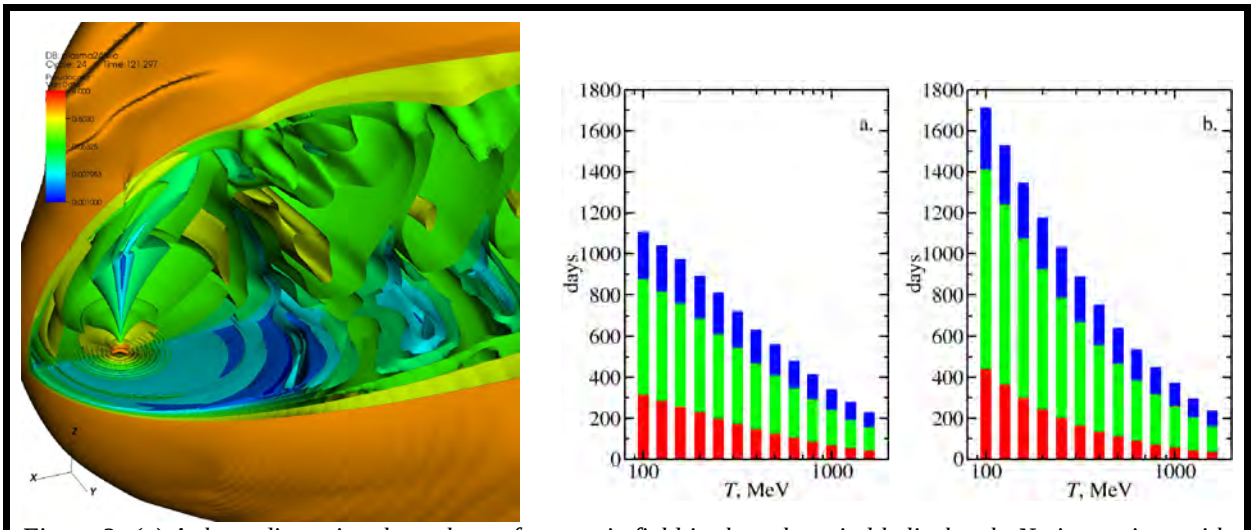


Figure 3: (a) A three-dimensional topology of magnetic field in the solar wind heliosheath. Notice regions with changing polarity of magnetic field corresponding to positive and negative solar cycles. (b) Time spent by cosmic ray protons in the solar wind (red), the heliosheath (green) and the heliotail (blue) regions [Florinski and Pogorelov, 2009].

Galaxy. The unmodulated intensity will provide the absolute upper limit on the intensity of ionizing galactic radiation in near Earth and interplanetary space with important implications for the future of manned space missions to other planets. We will learn, through Voyager measurements and theoretical modeling, just how effective is the heliosheath as a cosmic-ray shield for the inner solar system during different phases of the solar cycle. The role of the TS and the heliotail region in re-accelerating GCRs [Florinski et al., 2003] also needs to be evaluated.

Despite much progress, a number of observations remain puzzling. For example, the intensity of low-energy electrons has recently risen at Voyager 1, hinting that the heliosheath could be more "porous" than previously thought. It is possible that this observation means Voyager 1 is near the heliopause and is detecting a galactic electron component. We still do not have a good understanding of turbulence in the heliosheath, which controls scattering and diffusion of cosmic rays. This needs to be one of the theory focus areas for the next decade.

The termination shock, the heliosheath, and the heliopause

Voyager 2 discovered the TS to be in a highly unusual state known as an "energetic particle mediated shock" [Richardson et al., 2008; Florinski et al., 2009]. The plasma flow is slowed down upstream of the shock,in the region called the shock precursor, by the back-pressure of the shock-accelerated ions. Such shocks were only observed a handful of times in interplanetary space and the fact that the largest shock in the solar system is of this type is a boon to the science of mediated shocks in space and astrophysics. For example, supernova remnant shocks, believed to be sources of GCRs, could be themselves mediated by cosmic ray particles [Berezhko, 2001]. The TS may be regarded as a reference case for these distant astrophysical objects, an important cross-discipline connection.

Another landmark Voyager discovery was a lack of heating of solar wind ions as they crossed the TS [Richardson et al., 2008]. The missing energy is believed to be transferred to PUIs through efficient reflection from the shock potential [Lee et al., 1996; Zank et al., 1996]. An effort is currently underway to reconstruct the spectrum of low-energy PUIs using IBEX measurements of heliosheath ENAs. Figure 4a shows a comparison between IBEX-derived fluxes below 6 keV and higher-energy in situ Voyager proton data [Gloeckler and Fisk, 2010]. This comparison provides important clues as to how PUIs are heated and accelerated to eventually become anomalous cosmic rays.

The character and nature of turbulent fluctuations in the heliosheath has proven to be entirely different

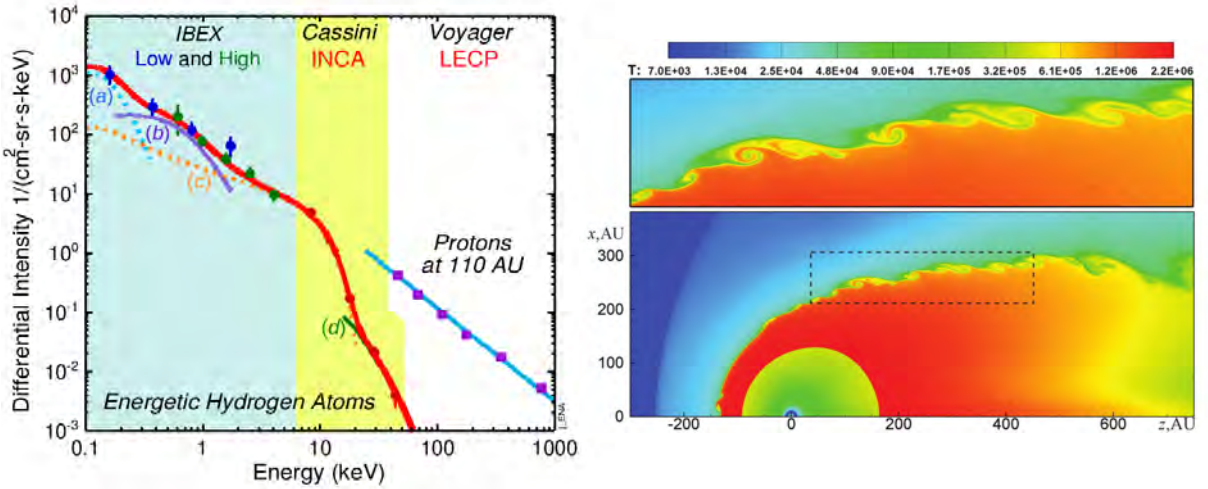


Figure 4: (a) Spectra of energetic protons and ENAs in the heliosheath from a model compared with Voyager, IBEX, and Cassini [Gloeckler and Fisk, 2010]. (b) A computer simulation showing vortices forming on the surface of the heliopause as a result of instability driven by charge exchange [Borovikov et al., 2008].

from the more familiar turbulence in the solar wind sampled by ACE, Wind, and Ulysses at 1-5 AU. Little is known about the physical processes that govern the intricate multi-scale interactions. The heliosheath region features unusual nonlinear structures, such as magnetic hole and humps. Another intriguing observation is that the magnetic field distribution is log-normal in supersonic solar wind, but Gaussian in subsonic heliosheath, with marked differences between the regions sampled by Voyager 1 and by Voyager 2 [Burlaga et al 2009].

The heliopause is now believed to be in a turbulent state as well. This boundary of the solar system may be unstable owing to a different degree of drag from plasma-neutral interaction on the solar and the interstellar sides. This process could lead to a formation of vortices shown in Figure 4b [Zank, 1999; Borovikov et al., 2008]. A mixing between the interstellar and heliosheath plasmas could lead to a production of X-rays that opens an intriguing possibility of observing astrospheres around other stars.

Case for future study. A realistic modeling of the heliosheath plasma, one that includes a self-consistent treatment of the PUIs, is critically important and essential to our understanding of the highly variable heliosheath plasma. This calls for theoretical and modeling development of the magnetic structures and turbulent interactions in the presence of PUIs in mutli-scale subsonic heliosheath plasma. More realistic models of heliosheath plasma taking into account its essentially non-thermal (non-Maxwellian) nature need to be developed over the coming decade.

Work continues on deducing the low-energy heliosheath PUI component, invisible to the Voyagers, from IBEX and Cassini ENA observations. Future IBEX measurements will reveal the differences in the suprathermal components in the fast and slow solar wind and in the upwind vs. heliotail directions. Research should also focus on elucidating the conditions near the heliopause, including the possible plasma depletion layer, magnetic reconnection sites, and instability due to charge exchange.

Summary

The next decade will be an exciting time for the outer heliospheric physicist. It is hard to overestimate the importance of the Voyager interstellar mission in being our only chance to study directly the very edge of the solar system and, for the first time, the interstellar space itself. In this paper we presented a list of discoveries made in the outer heliosphere over the past decade and discussed their significance in other areas of space physics and astrophysics. New discoveries of this magnitude are still out there, and call for a focused and reliably funded supporting theory and modeling program concentrating on the outer heliosphere. At present, the field is a relatively minor component of the Heliophysics program element at

NASA. While in principle, the outer heliosphere is included in this component, funding has traditionally been redirected to disciplines perceived to be more “mainstream” (solar physics and space weather).

The following is of what we believe would be an appropriate range of focus areas in the outer heliospheric research for the next decade.

- Large-scale structure of the heliospheric interface, heliosheath flow geometry,
- Conditions in the very local interstellar medium (magnetic field, galactic cosmic ray specra),
- Interstellar and heliospheric neutral atoms, composition,
- Suprathermal plasma component, pickup ion evolution in the heliosheath,
- Deducing acceleration sites and mechanisms of anomalous cosmic rays,
- Physics of the termination shock, ion reflection and energization, role of turbulence and shock variability,
- Understanding galactic cosmic ray filtration by the “heliospheric shield”,
- Properties of turbulent fluctuations and nonlinear structures in the heliosheath, their effects on energetic particles,
- Physics of the heliopause, magnetic reconnection, instabilities of the heliopause.

To sustain the current outer heliospheric theoretical science community a modest level of funding is required. We estimate support is needed for an equivalent of 20 full time researcher positions per year at $200,000/year, which comes to a total of $40M over ten years. Funding could be delivered via dedicated theory or guest investigator program as 7 or more three-year awards of $500,000-$600,000 each per year.

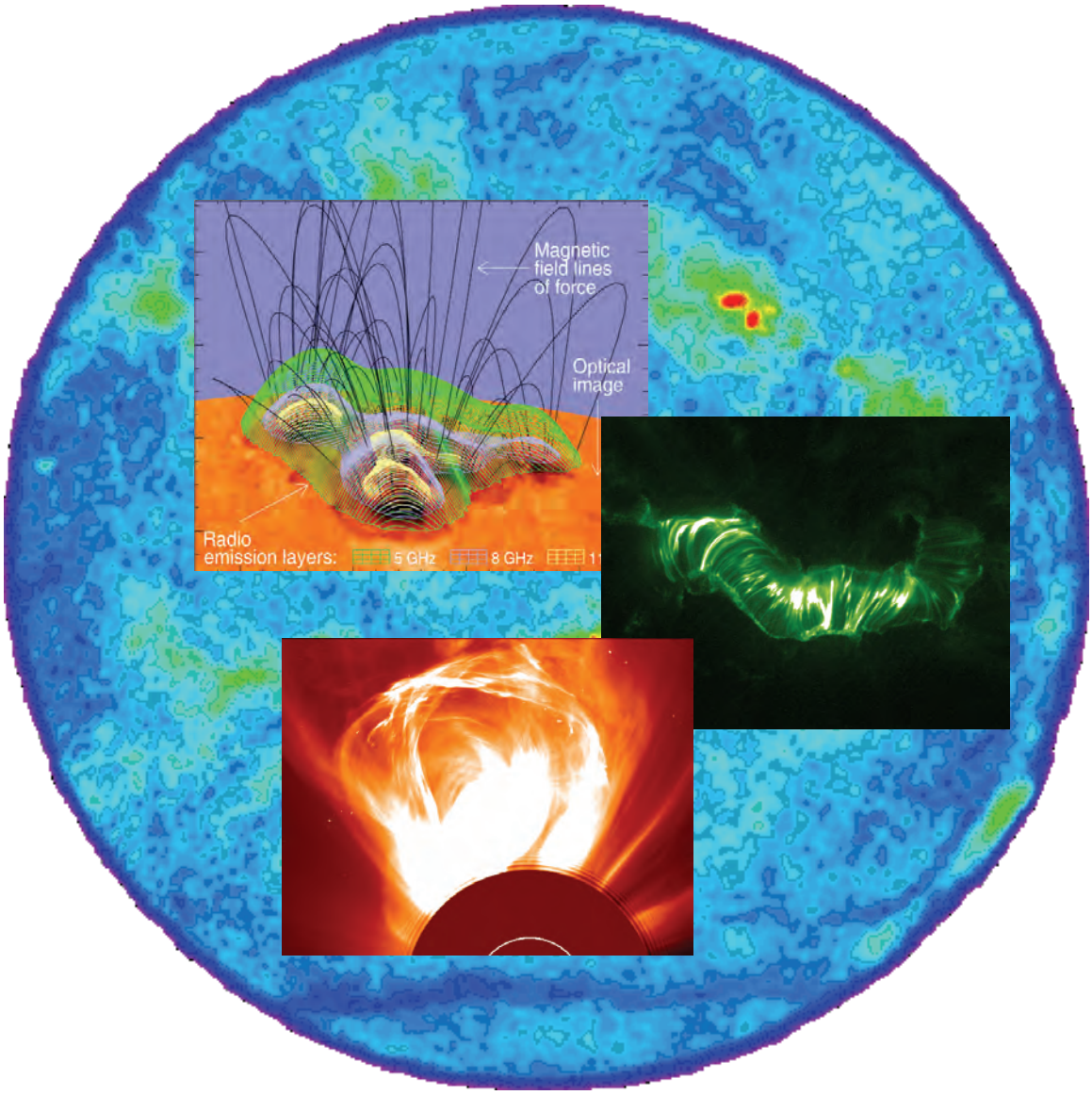
References

E. Berezhko 2001, Space Sci. Rev., 99, 295
S. N. Borovikov et al., Astrophys. J., 682, 1404
L. F. Burlaga et al. 2009, Astrophys. J., 692, 1125
R. B. Decker et al. 2005, Science, 309, 2020
J. F. Drake, et al. 2010, Astrophys. J., 709, 963
L. A. Fisk and G. Gloeckler 2009, Adv. Space Res., 43, 1471
V. Florinski et al. 2003, J. Geophys. Res., 108, 1228
V. Florinski and N. V. Pogorelov 2009, Astrophys. J., 701, 642
V. Florinski et al. 2009, Geophys. Res. Lett., 36, L12101
G. Gloeckler and L. A. Fisk 2010, AIP Conf. Proc., 1302, 110
J. Heerikhuisen et al. 2010, Astrophys. J., 708, L126
J. R. Jokipii 1986, J. Geophys. Res., 2929
J. Kota 2010, Astrophys. J., 723, 393
S. M. Krimigis et al. 2003, Nature, 426, 45
S. M. Krimigis et al. 2009, Science, 326, 971
M. A. Lee et al. 1996, J. Geophys. Res., 101, 4777
D. J. McComas and N. A. Schwadron 2006, Geophys. Res. Lett., 33, L04102
D. J. McComas et al. 2009, Science, 326, 959
F. B. McDonald et al. 2003, Nature, 426, 48
M. Opher et al. 2007, Science, 316, 875
E. N. Parker 1961, Astrophys. J., 134, 20
M. E. Pesses et al. 1981, Astrophys. J., 246, L85
N. V. Pogorelov et al. 2007, Astrophys. J., 668, 611
J. D. Richardson et al. 2008, Nature, 454, 63
J. D. Richardson et al. 2009, Geophys. Res. Lett. 2009, 36, L10102
E. C. Stone et al. 2005, Science, 309, 2017
E. C. Stone et al. 2008, Nature, 454, 71
H. Washimi et al. 2007, Astrophys. J., 670, L139
W. R. Webber and P. R. Higbie 2009, J. Geophys. Res., 114, A02103
G. P. Zank et al. 1996, J. Geophys. Res., 101, 457
G. P. Zank 1999, AIP Conf. Proc. 471, 783

The Frequency Agile Solar Radiotelescope

A White Paper for

A Decadal Strategy in Solar and Space Physics



D. E. Gary[1], T. S. Bastian[2], G. J. Hurford[3] and S. M. White[4],
for the FASR Team

[1]*New Jersey Institute of Technology,* [2]*National Radio Astronomy Observatory,*
[3]*University of California, Berkeley,* [4]*U. S. Air Force Research Lab*

1 Summary & Background

Solar activity represents a uniquely accessible example of the ubiquitous processes of energy release and particle acceleration that occur in a wide range of contexts, from the magnetosphere to the galaxy. Radio observations provide both a unifying perspective on solar activity and a unique sensitivity to solar magnetic fields, the energy source for solar activity. In addition, radio emission is unsurpassed in its sensitivity to accelerated particles. As a result, radio emission provides a powerful source of diagnostic information with the potential for transformational insights into solar activity and its terrestrial impacts.

The *Frequency Agile Solar Radiotelescope* (FASR) is a solar-dedicated radio telescope that provides a unique combination of superior imaging capability and broad instantaneous frequency coverage to address a wide range of science goals. The potential value of such a facility has been recognized by high-priority recommendations for FASR from three previous surveys of future science in the US conducted by the National Research Council. The previous decadal survey of astronomy and astrophysics, recommended FASR a moderate-sized initiative. The *Solar and Space Physics Survey Committee (2003)* made FASR its highest-priority "small project", in recognition of the unique and transformative role it will play in addressing basic research questions, as well as its complementary role to other important instruments being developed and deployed to address the Sun-Earth system. The recently concluded *New Worlds, New Horizons in Astronomy and Astrophysics (2010;* Astro2010) amplified its previous recommendation to endorse FASR as its number-two rated project in all of radio, millimeter, and submillimeter (RMS) astronomy, citing the project as "compelling" and in a high state of readiness.

FASR will be designed and constructed by a university consortium organized under Associated Universities, Inc. (AUI), the not-for-profit science management corporation that manages the National Radio Astronomy Observatory (NRAO). The project was presented in detail to Astro2010. More than 70 pages of documentation were provided to the survey via two "requests for information" (RFI-1 and RFI-2). As part of RFI-2, the project was independently assessed by the *Cost Risk and Technical Evaluation* (CATE) process (Aerospace Corp.). Based on this assessment the Astro2010 RMS panel report concluded "FASR is a mature effort that has been recommended in two previous NRC decadal surveys. Extensive development and design review have been funded and achieved, including the testing of prototype instrumentation and the detailed planning of FASR operations, maintenance, and management. Independent analysis of FASR characterized it as 'doable today'." A project based on FASR science and technology, roughly 10% of FASR in size and cost, is the Owens Valley Solar Array Expansion, funded by a $5.2 M grant from NSF to New Jersey Institute of Technology. This project, to be completed in 2013, will further mature FASR's technology, software and data handling.

The FASR project cost in FY09 dollars, provided to Astro2010, was $68.5M, including contingency and management fees. The CATE analysis estimated the cost to be $109M. Although the team is not privy to the details of the analysis, the RMS report states that the difference is "based primarily on higher estimates for project management, antennas, and reserves." The project team estimated annual operations costs of FASR to be $3M whereas the CATE analysis put the cost at $4M. Astro2010 assumed that FASR construction would be funded equally by the AST and AGS divisions at the NSF.

2 FASR Science

FASR is an interferometric array designed to perform Fourier synthesis imaging at radio wavelengths. Moreover, as a radioheliograph, FASR is designed and optimized to observe the Sun. The major advances offered by FASR over previous generations of solar radio instrumentation are its unique combination of ultra-wide frequency coverage, high spectral resolution, and outstanding image quality. FASR measures the polarized brightness temperature spectrum (from 50 MHz to 21 GHz) along every line of sight to the

Sun as a function of time. Radiation in this wavelength range probes the solar atmosphere from the middle chromosphere to well up into the corona. In essence, FASR images the entire solar atmosphere once every second from the chromosphere through the corona, while retaining the capability to image a restricted frequency range with as little as 20 ms time resolution. In so doing, FASR enables fundamentally new and unique observables. Moreover, FASR's panoramic view allows the solar atmosphere and the physical phenomena therein to be studied as a *coupled system*.

Here we summarize the main science goals of the proposed instrument while at the same time emphasizing the new observables enabled by FASR. The potential of FASR for breakthroughs in a wide range of science topics, resulting from extensive discussions among members of the solar and space physics community, is presented in detail in the book *Solar and Space Weather Radiophysics*[24]. Elements of FASR science are also discussed in various science white papers. With its unique and innovative capabilities FASR also has tremendous potential for new discoveries beyond those presently anticipated.

2.1 The Nature and Evolution of Coronal Magnetic Fields

Quantitative knowledge of coronal magnetic fields is fundamental to understanding essentially all solar phenomena above the photosphere, including the structure and evolution of active regions, prompt energy release, charged particle acceleration, flares, coronal mass ejections (CMEs), filaments, coronal heating, and acceleration of the solar wind (see the white paper entitled *Coronal Magnetic Fields* by White et al.). Useful quantitative measurements of the coronal magnetic field have been largely unavailable to date. Thus, considerable resources are devoted to *indirect* means of constraining the coronal magnetic field. For example, the field topology can be traced by hot, optically thin, thermal plasma emitting at EUV/SXR wavelengths. Measurements of the (vector) magnetic field distribution can be measured at photospheric and numerically extrapolated into the upper chromosphere and corona. Such extrapolations depend sensitively on measurement errors and rely on assumptions whose validity is questionable – that the magnetic field is potential or force free[35]. Hence, new observational techniques for directly measuring the coronal magnetic fields are being vigorously pursued. At optical/IR wavelengths, the Hanle and Zeeman effects are being measured for this purpose, but such measurements are only possible at the solar limb where photospheric fields are not well measured and results values are integrated along the line-of-sight[40].

Radio observations provide the most direct means of measuring coronal magnetic fields. The key radio technique for doing so is the well-established gyroresonance method[30,44,45]. Such radio measurements, effective from ≈120-2000 gauss, allow coronal magnetic fields in active regions to be measured both on the disk and above the solar limb. Gyroresonance emission is optically thick; by imaging at multiple frequencies, the thin resonance location varies in a systematic way, allowing a CAT-scan-like view of nested isogauss layers. The most straightforward application is to exploit the break in the brightness temperature spectrum along each line of sight to infer the (force free) magnetic field at the base of the corona, thereby producing a 2D **coronal magnetogram**. Several additional radio techniques[24] are available for measuring coronal magnetic fields in a wide variety of physical environments, including polarization measurements of free-free emission to measure the longitudinal magnetic field in the quiet corona[25,27] and dense structures such as filaments and prominences[2]; radio bursts at meter and decimeter wavelengths can be used to measure fields in the middle corona[17]; gyrosynchrotron emission can be used to deduce the magnetic field in flaring magnetic loops and in CMEs[10,33,] (see also §2.2-3); mode coupling phenomena can place additional, independent topological constraints on the coronal magnetic field[41].

FASR will be unique in its ability to deliver these transformative means of measuring the coronal magnetic field both on the solar disk and above the limb.

2.2 The Physics of Flares

Outstanding problems in the physics of flares include those of magnetic energy release, particle acceleration, and particle injection and transport. There has been significant theoretical and observational progress in recent years on three-dimensional magnetic reconnection[7,18,22,39,]. FASR brings new tools to bear on the problem of 3D magnetic reconnection. At centimeter wavelengths, gyrosynchrotron emission – radiation from nonthermal electrons with energies of 10s of keV to several MeV gyrating in a magnetic field – illuminates any magnetic coronal loop to which energetic electrons have access, showing when and where accelerated electrons are present. Such loops provide a valuable topological tool, linking the site of particle acceleration to more distant sites to which accelerated particles are transported during all phases of the flare. Moreover, FASR will provide quantitative measurements of magnetic fields in the flaring source through forward fitting of the spectrum, coronal loop oscillations[6,43,21], and timing comparisons between microwave and HXR fine structures[1,9].

At longer wavelengths, decimetric spike bursts and type-III-like bursts accompany impulsive energy release in flares[3]. They may be intimately related to the magnetic energy release process[8,12] associated with reconnection in both the impulsive phase of flares, and in post-flare loops[13]. FASR's dynamic imaging spectroscopy is urgently needed to identify the location of these emissions. The combination of spatial, spectral and temporal resolution gives the fascinating opportunity to directly image the diffusion region where the primary acceleration of the particles takes place.

In addition to probing the site, topology, and evolution of magnetic fields involved in 3D magnetic reconnection, *FASR will be a superb probe of the (anisotropic) distribution function of accelerated electrons and its spatiotemporal evolution*. Nearly every flare accelerates electrons whose gyrosynchrotron spectrum in microwaves can be forward fit and the relevant physical parameters thereby inferred[8,11,33]. The key to further exploitation of this approach is to add high-spatial-resolution imaging, allowing temperature, density, magnetic field strength[23], non-thermal electron energy and angular distribution[1,19,20,32] (see also the white paper entitled *Uncovering Mechanisms of Coronal Magnetism via Advanced 3D Modeling of Flares and Active Regions* by Fleishman et al. and references therein), and even the level of turbulence to be extracted for every resolution element and time. These diagnostics provide powerful means of testing theories of electron acceleration[36] and transport[4,5,34] (see also the white paper entitled *Particle Acceleration and Transport on the Sun* by Gary *et al.*). These observations will be an invaluable complement to hard X-ray emission, which typically originates from the conjugate magnetic foot points of coronal magnetic loops into which energetic electrons are injected, as well as to EUV/SXR emission.

FASR will provide new insights into magnetic energy release and quantitative measurements of the magnetic field and the evolving energetic electron distribution. It will provide access to the flare phenomenon as a coupled physical system.

2.3 The Drivers of Space Weather

The term "space weather" refers to phenomena that disturb the interplanetary medium and/or affect the Earth and near-Earth environment[29]. This includes recurrent structures in the solar wind (fast solar wind streams, co-rotating interaction regions), the ionizing radiation and hard particle radiations from flares, coronal mass ejections, and shock-accelerated particles. These drivers result in interfering radio noise, geomagnetic storms, changes in the ionosphere, and atmospheric heating which can, in turn, result in a large variety of effects that are of practical concern to our technological society: ground-level currents in pipelines and electrical power grids, disruption of civilian and military communication, spacecraft charging, enhanced atmospheric drag on spacecraft, etc. They can also disrupt the work, and endanger the lives, of humans living and working in space. The drivers of space weather are all solar in origin.

FASR will be sensitive to a large number of space weather phenomena, and particularly to coronal mass ejections (CMEs), filament/prominence eruptions, and type II radio bursts. Interest in CMEs is particularly strong because they are associated with the largest geo-effective events and the largest solar energetic particle (SEP) events. With the detection of synchrotron radiation from CMEs in cm to m wavelengths[10,33] *a new tool has become available to detect, image, and diagnose the properties of CMEs.* FASR will exploit this tool extensively. FASR will also make important contributions to the study of shocks and waves in the corona. Fast ejecta and/or a blast wave are sometimes produced by flares; a CME produces a piston-driven shock wave. The relationship between these shocks, their radio-spectroscopic signature, and other phenomena of interest such as Moreton waves and "EIT waves" – many of which FASR will also image[46] – remains a matter of considerable interest and controversy[14,26]. With its unique ability to perform imaging spectroscopy, FASR will be able to simultaneously image the basic shock driver (flare or CME signature), the response of the atmosphere to the driver (chromospheric/coronal waves and coronal dimmings), and the shocks themselves that may form due to the associated flare or the CME. Similarly, FASR will yield important new insights into the relationship between flares, CMEs, and solar energetic particles (SEPs) as it will image the CME precursor environment.

FASR's ability to provide a comprehensive and integrated picture of the flare phenomena applies equally to space weather drivers. It will enable consideration of the entire coupled physical system from the chromosphere out into the corona.

2.4 The Physics of the Quiet Sun

Radio emission from the quiet Sun is dominated by free-free opacity, which is proportional to the square of the electron density, and whose polarization is proportional to the longitudinal *B* field. At the highest FASR frequency, 21 GHz, FASR will sample the middle chromosphere while, by systematically tuning to lower frequencies, the sampled height of the (optically thick) emission moves through the top of the chromosphere and into the corona. The multifrequency imaging capability of FASR again provides a heretofore unexploited CAT-scan-like probe into the temperature, density, and *B* field structure of this important region of the Sun. The importance of this region for determining the global properties of the corona and solar wind is hinted at by the polar cap brightenings seen with Nobeyama[42]. FASR's multifrequency imaging will clarify the connection between the chromospheric structure and the origin of the solar wind.

FASR will also address the dynamics of this region of the solar atmosphere. First, resonance wave heating[16,37] represents one of many models for coronal heating and makes specific predictions of the location and time scales of energy deposition in coronal loops. Using free-free diagnostics, FASR observations of the resulting temperature changes in this region encompassing both the chromosphere and the adjacent corona can address the validity of such predictions. A second area is the role of "nanoflares"[38], and a host of perhaps related microflare phenomena that, collectively, may dominate the structure of the lower solar atmosphere. The energetics of such small events depends critically on their thermal properties and the role, if present, of accelerated electrons[15,28]. Again, the combination of FASR's frequency coverage and imaging can apply diagnostics to quantitatively address both thermal and nonthermal signatures of such events. See also the white paper entitled *The Solar Chromosphere* by McIntosh et al.

A third example, already mentioned in §2.1, is the strength, distribution and evolution of magnetic fields in prominences and filaments, as deduced from the polarization of free-free emission[25,27]. FASR can use imaging spectroscopy to directly measure the longitudinal magnetic field in filaments and prominences. Such measurements are critically important to understanding the formation, structure, and destabilization of such structures and their role in CMEs and interplanetary flux ropes.

The exploitation of free-free emission to determine the spatially-resolved temperature, density and magnetic fields in the quiet chromosphere, corona, and prominences is possible only as a result of FASR's unique attributes.

3 Overview of the Instrument

FASR's design is driven by a number of factors. First, the instrument meets the science specifications derived from the key science objectives outlined §2. Second, it meets the needs of the large and diverse scientific community that will exploit FASR data. Since most users will not be radio interferometrists, they should not bear the burden of data calibration and reduction. As a result, these functions will be performed by pipeline data processing that delivers high-level, fully calibrated data products to a broad user community. Third, it is imperative to control the long-term operating and maintenance costs. Finally, it is important to design an expandable/upgradable system in order to respond flexibly to opportunities presented by new resources, new technologies, and/or new science objectives in the future. *These broad drivers have resulted in a design that emphasizes simplicity, utility, reliability, maintainability, and expandability.*

The science objectives outlined in §2 flow down to the high-level science and technical specifications summarized in the Table. FASR is designed to perform Fourier synthesis imaging using well-established interferometric techniques. For an array of N antennas there are $\sim N^2/2$ independent antenna pairs, each of which measures a single Fourier component, or complex *visibility*, of the Sun's radio brightness distribution at a given frequency, time, and spatial frequency. The ensemble of antenna pairs, or *antenna baselines*, therefore measures many Fourier components. Fourier inversion of the visibility measurements yields an image of the Sun's radio brightness at a given frequency, time, and polarization. Deconvolution techniques are then used to remove the effects of the *point spread function*, the response of the instrument to a point source.

Angular resolution	$20/\nu_{GHz}$ arcsec
Frequency range	50 MHz – 21 GHz
Number data channels	2 (dual polarization)
Frequency bandwidth	500 MHz per channel
Frequency resolution	Instrumental: 4000 channels Scientific: min(1%, 5 MHz)
Time resolution	~1 s (full spectrum sweep) 20 ms (dwell)
Polarization	Full Stokes (IQUV)
Number antennas deployed	**A** (2-21 GHz): ~100 **B** (0.3-2.5 GHz): ~70 **C** (50-350 MHz): ~50
Size antennas	**A** (2-21 GHz): 2 m **B** (0.3-2.5 GHz): 6 m **C** (50-350 MHz): LPDA
Array size	4.25 km EW x 3.75 km NS
Astrometry	1 arcsec
Flux calibration	<10% absolute 1% relative

FASR observations will cover the frequency range of 50 MHz to 21 GHz using three separate arrays of antennas, denoted FASR-A, -B and -C. Each array provides frequency coverage over roughly a decade of bandwidth: 2-21 GHz (FASR-A); 0.3-2.5 GHz (FASR-B), and 50-350 MHz (FASR-C). The number, type, and configuration of the antennas in each array are chosen to address the key science objectives summarized in §2. The antennas will be distributed over an area of approximately 4.25 x 3.75 km, providing an angular resolution of 1" at 20 GHz over an hour angle range of at least ±3 hr. The angular resolution scales linearly with wavelength from this fiducial value and is sufficient to resolve solar radio sources observed as well as being well matched to

the capabilities of missions such as TRACE, SOHO, RHESSI, Hinode, and SDO. The instrument will be calibrated against sidereal sources to enable precision astrometry and an absolute flux calibration of better than 10%. The relative calibration across the more than two decades of bandwidth will be better than 1%.

4 FASR Construction, Operation, and Costs

The FASR construction project will be of five years duration. It will be implemented by a consortium of university partners and a national facility (NRAO) under AUI management. FASR has undergone several costing exercises. The most comprehensive costing exercises were the result of the FASR team's response to the NSF ATM Mid-Sized Infrastructure (MSI) Opportunity in 2008 and the Astro2010 RFI responses in 2009. The cost of FASR construction is estimated by the FASR team to be $68.5M in FY09 dollars. The amount includes contingency and management fees. The Astro2010 *Cost Risk and Technical Evaluation* (CATE), an independent assessment overseen by Aerospace Corp. on behalf of the Astro2010 survey, put the cost at $109M. While the FASR team is not privy to the CATE analysis, the Astro2010 RMS panel states that the higher cost estimated by CATE is "based primarily on higher estimates for project management, antennas, and reserves." The RMS panel adopted $100M as FASR construction cost and commented that FASR is a "mature" effort and characterized the project as "doable today".

As part of the FASR design and development effort, a detailed Operations and Maintenance (O&M) plan was prepared to assess the long-term costs of operating the instrument at the request of NSF. Details are available upon request. The FASR O&M Plan estimated the core long-term annual O&M cost to be approximately $3M in FY09 dollars whereas the Astro2010 CATE analysis estimated O&M costs to be $4M annually. Both estimates are significantly lower than the 10% of capital costs often used as a "rule of thumb" for estimating operating costs. The low operating cost estimate reflects the fact that FASR does not employ high maintenance technologies. Rather, FASR employs simple, non-cryogenic front ends; small, mechanically simple antennas; analog data transmission; signal processing confined to a shielded and environmentally controlled central hub, etc.). Since FASR is a solar dedicated instrument, pipeline data processing is both feasible and necessary to meet the needs of the user community. The software and data management plan describes the data pipeline – including data calibration, RFI excision, and data selection – the interim data archive, the applications data base, the user interface, and data acquisition and analysis tools in some detail. Instrument monitor and control functions are also considered. The user support model includes both quick-look data for operational needs, and a permanent archive of applications databases that can be accessed and mined via a web interface. The volume of data that will be permanently archived is estimated to be ~200 GBytes a day.

The O&M cost estimate does not include support of special observing modes, data products, or mission support, which, if added, would be supported with funds from other agencies (e.g., NASA, USAF, etc). Nor does the estimate include possible contributions, financial or in kind, from university and/or foreign partners.

Acknowledgment

Many individuals have contributed to FASR design, development, and planning: M. Azarian, T. Bastian, B. Block, D. Bock, D. Boyd, R. Bradley, L. D'Addario, N. Dalal, D. Das, C. Edmonson, G. Fleishman, D. Gary, R. Gawande, S. Gross, G. Hurford, Z. Liu, M. Morgan, G. Nita, C. Ruf, K. Saini, A. Thompson, S. Tun, J. Webber, S. White, T. Zurbuchen. Thanks also to C. Allen, G. Clark, P. Jewell, K. Lo, T. Miller, N. Radziwill, T. Readhead, A. Sargent, P. Donahoe, and E. Schreier for input on site, organizational, and administrative matters.

References

1. Altyntsev, A. T., Fleishman, G. D., Huang, G.-L., & Melnikov, V. F. 2008, ApJ 677, 1367
2. Apushkinskij, G. P., Topchilo, N. A., Tsyganov, A. N., & Nesterov, N. S. 1996, Astron. Nach. 317, 417
3. Aschwanden, M. J., & Benz, A. O. 1997, ApJ 480, 825
4. Aschwanden, M. J., et al. 1997, ApJ 487, 936
5. Aschwanden, M. J., Schwartz, R., & Dennis, B. R. 1998, ApJ 502, 468
6. Aschwanden, M. J. 2004, ApJ 608, 554
7. Aulanier, G., et al. 2000, Advances in Space Research, 26, 485
8. Bastian, T. S., Benz, A. O., & Gary, D. E. 1998, ARAA 36, 131
9. Bastian, T. S. 1999, in *Solar Physics with Radio Observations*, eds.: T. S. Bastian, N. Gopalswamy and K. Shibasaki, NRO Report 479., p.211
10. Bastian, T. S., Maia, D., Pick, M., & Kerdraon, A. 2001, ApJ 558, L65
11. Bastian, T.S., Fleishman, G. D., & Gary, D. E. 2007, ApJ 666, 1256
12. Benz, A. O. 1994, Sp. Sci. Rev. 68, 135
13. Benz, A. O., Saint-Hilaire, P., & Vilmer, N., 2002, A&A 383, 678
14. Chen, J., Marqué, C., Vourlidas, A., Krall, J., & Schuck, P. W. 2006, ApJ 649, 452
15. Christe, S., Hannah, I. G., Krucker, S., McTiernan, J., & Lin, R. P. 2008, ApJ 677, 1385
16. Cranmer, S., van Ballegooijan, A., & Edgar, 2007, ApJS 171, 520
17. Dulk, G., & McClean, D. 1978, Solar Phys. 57, 279
18. Edmonson, J. K., Antiochos, S. K., DeVore, C. R., & Zurbuchen, T. H. 2010, ApJ 718, 72
19. Fleishman, G., & Melnikov, V. 2003a, ApJ 584, 1071
20. Fleishman, G., & Melnikov, V. 2003b, ApJ 587, 823
21. Fleishman, G., Bastian, T. S., & Gary, D. E. 2008, ApJ 684, 1433
22. Fletcher, L., Metcalf, T., Alexander, D., Brown, D. S., & Ryder, L. A. 2001, ApJ 554, 451
23. Gary, D. E., & Hurford, G. J. 1994, ApJ 420, 903
24. Gary, D. E., and Keller, C. U. 2004, *Solar and Space Weather Radiophysics*, Springer/Kluwer, Astrophysics and Space Science Library, v. 314
25. Gelfriekh, G. 2004, in *Solar and Space Weather Radiophysics*, eds. D. E. Gary and C. Keller, Springer/Kluwer, Astrophysics and Space Science Library, v. 314, 115
26. Gopalswamy, N. 2000, in Radio Astronomy at Long Wavelengths, eds. R.G. Stone, K. W. Weiler, M. L. Goldstein, & J.-L. Bougeret
27. Grebinskij, A., et al. 2000, A&A 144, 169
28. Hannah, I. G., et al. 2008, ApJ 677, 704
29. Lanzerotti, L. 2004, in *Solar and Space Weather Radiophysics*, eds. D. E. Gary and C. Keller, Springer/Kluwer, Astrophysics and Space Science Library, v. 314
30. Lee, J. 2007, Sp. Sci. Rev., 133, 73
31. Lee, J., McClymont, A. N., Miki,c Z., White, S. M., & Kundu, M. R. 1998, ApJ 501, 853
32. Lee, J., & Gary, D. E. 2000, ApJ 543, 457
33. Maia, D., et al. 2007, ApJ 660, 874
34. Melrose, D. B., & Brown, J. C. 1976, MNRAS 176, 15
35. Metcalf, T. R., Derosa, M., Schrijver, C., Barnes, G., van Ballegooijen, et al. 2008, Solar Phys, 247, 269
36. Miller, J. A., Cargill, P. J., Emslie, A. G., Holman, G. D., Dennis, B. R., et al. 1997, JGR 102, 14631
37. Ofman, L., Klimchuk, J. A., & Davila, J. M. 1998 ApJ, 493, 474

38. Parker, E. N. 1988, ApJ 330, 474
39. Priest, E. R., & Schrijver, C. J. 1999, Solar Phys. 190, 1
40. Raouafi, N., Solanki, S. K., & Wiegelmann, T. 2008, ArXiv e-prints, 801, arXiv:0801.2202
41. Ryabov, B. 2004, in Solar and Space Weather Radiophysics, eds. D. E. Gary and C. Keller, 135
42. Shibasaki, K. 1998, ASP Conf. Proc., *Synoptic Solar Physics*, 140, 373
43. Trottet, G., Kerdraon, A., Benz, A. O, & Treumann, R. 1981, AA 93, 129
44. White, S. M., & Kundu, M. R. 1997, Solar Phys. 174, 31
45. White, S. M. 2004, in *Solar and Space Weather Radiophysics*, eds. D. E. Gary and C. Keller, Springer/Kluwer, Astrophysics and Space Science Library, v. 314
46. White, S. M., and Thompson, B. J. 2005, ApJ 620, L63

Particle Acceleration and Transport on the Sun

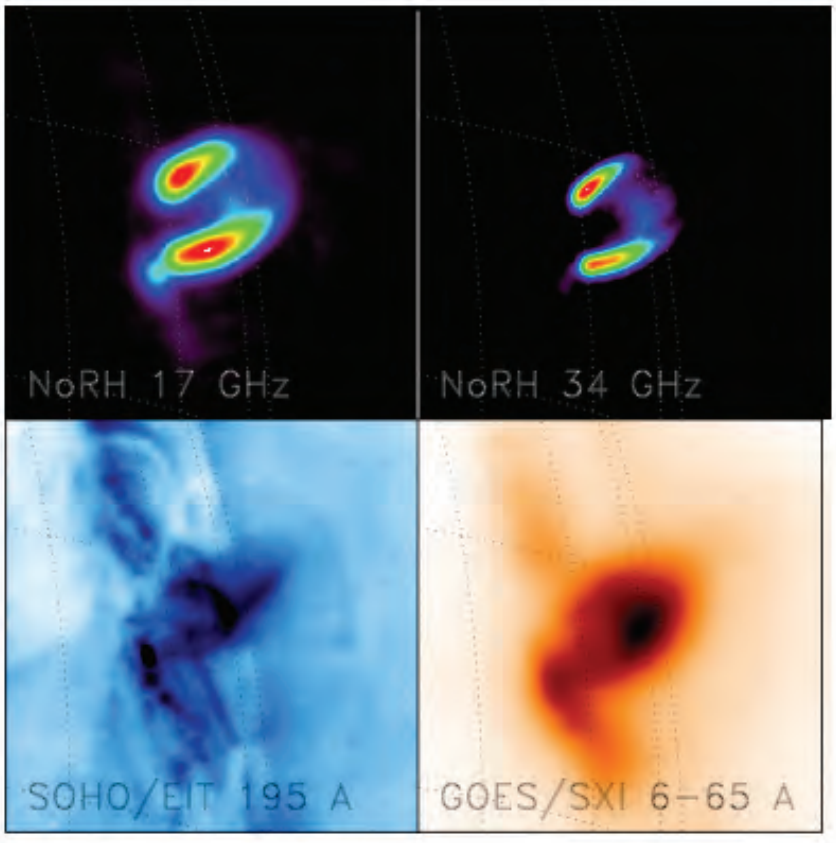
New Perspectives at Radio Wavelengths

A White Paper for the Decadal Survey of Solar and Space Physics

Prepared by

D. Gary[1], D. Alexander[2], T. Bastian[3], G. Emslie[4], G. Fleishman[1], G. Holman[5], H. Hudson[6], G. Hurford[6], S. Krucker[6,7], J. Lee[1], J. Miller[8], S. White[9]

[1]New Jersey Institute of Technology, [2]Rice University, [3]National Radio Astronomy Observatory,[4]Western Kentucky University, [5]NASA/Goddard, [6]University of California at Berkeley, [7]University of Applied Sciences, Switzerland, [8]University of Alabama at Huntsville, [9]U. S. Air Force Research Lab



Executive Summary

Particle acceleration and particle transport are ubiquitous on the Sun, planetary magnetospheres, and the heliosphere. The physical context on the Sun involves solar flares and coronal mass ejections (CMEs). These complex coupled phenomena require comprehensive and complementary observations to disentangle the relevant physical mechanisms at work. They also require modeling and theory, as discussed in the white paper entitled *Uncovering Mechanisms of Coronal Magnetism via Advanced 3D Modeling of Flares and Active Regions*, by Fleishman et al., but here we focus on observations. Radio observations are emphasized because they are positioned in the coming decade to make unique and innovative contributions to these important problems. In particular, the transformative technique of *ultra-broadband imaging spectroscopy* is discussed, which serves as the observational basis for new insights into particle acceleration and transport. These include the measurement of coronal magnetic fields, the measurement of the spatiotemporal evolution of the electron distribution function, and the observation of flares and CMEs as coupled phenomena.

Introduction

Particle acceleration and particle transport are fundamental astrophysical processes. Energetic electrons and ions are observed indirectly via the electromagnetic radiation they produce and are therefore accessible to study through observations from a variety of telescopes. Alternatively, they can be measured by in situ measurements using instrumentation on spacebased platforms. Energetic particles are produced in astrophysical contexts as diverse as planetary magnetospheres, stellar atmospheres and winds, accretion disks, supernovae and their remnants, and the lobes of extragalactic radio sources. A detailed understanding of particle acceleration and transport bears on issues ranging from astronaut safety to insights into the most energetic phenomena in our universe.

By virtue of its proximity, the Sun offers an astrophysical laboratory in which fundamental processes can be studied in considerably greater detail than is possible for those occurring in remote environments. It therefore serves as an important touchstone against which our understanding of these processes is referenced. The Sun is an efficient particle accelerator: it can release several tenths of the available energy (stored in stressed magnetic fields) in high-energy nonthermal particles: it can promptly accelerate electrons to energies in excess of 100 MeV and ions to energies ~1 GeV/nuc. It does so in a variety of energetic environments – in active regions, plasma jets, solar flares, and CMEs – and by a variety of mechanisms.

Energetic particles on the Sun emit radiation through a variety of mechanisms:

- Soft X-ray (SXR) emission arises from thermal plasma heated to ~20 MK during flares.
- Hard X-ray (HXR) emission typically arises from nonthermal bremsstrahlung resulting from electrons (~10 keV to a few x 100 keV) interacting with dense chromospheric material; trapped electrons interact with the tenuous solar corona.

- Continuum gamma-ray emission (E> few x 100 keV) arises from electron-proton and electron-electron bremsstrahlung and pion decay; gamma-ray lines are produced by electron-positron annihilation, neutron capture, and nuclear excitation.

At radio wavelengths:

- Thermal plasma at all temperatures emits radio waves via thermal bremsstrahlung, including hot, SXR-emitting plasma.

- Suprathermal electrons (few to ~20 keV) radiate coherently at radio wavelengths as a result of plasma instabilities driven by coronal shocks and particle beams.

- Energetic electrons (10s of keV to many MeV) emit nonthermal gyrosynchrotron radiation, a mechanism that is sensitive to the coronal magnetic field.

Together, X-ray, gamma-ray, and radio observations provide a powerful and complementary suite of perspectives from which to study particle acceleration. These observations provide powerful diagnostics of particle acceleration at the Sun which can in turn provide insight into the accelerated particles measured in the heliosphere (e.g. impulsive SEPs).

This white paper emphasizes opportunities becoming available in the next decade for studying particle acceleration and transport using *radio* observations of the Sun. Unlike radiation mechanisms at X-ray and gamma-ray wavelengths, radio radiation mechanisms are uniquely sensitive to magnetic fields and, moreover, can have a strong dependence on anisotropies in the electron distribution function. Furthermore, and again unlike X-ray and gamma-ray observations, radio observations can be performed using ground-based instrumentation. Radio observations therefore represent both a powerful and cost effective means of studying fundamental aspects of particle acceleration on the Sun.

Acceleration Mechanisms

The two most prominent types of energetic phenomena on the Sun that accelerate electrons and ions to high energies are solar flares and CMEs, although low-level particle acceleration occurs more or less continuously in the Sun's corona[17]. It is widely believed that solar flares are powered by magnetic reconnection and that the bulk of the energy released is deposited in non-thermal electrons although for some events the ions can also contain comparable amounts of energy. CMEs result from a large-scale destabilization in the coronal magnetic field, leading to the ejection of a magnetic flux rope from the Sun. This process frequently leads to strong localized energy release and subsequent acceleration of fast particles. Fast CMEs can drive strong shocks far into the interplanetary medium (IPM) leading to several broad classes of particle acceleration that may be relevant to both flares and CMEs[24,26]. The actual acceleration mechanisms in these events are unknown but one of three distinct mechanisms is generally proposed.

Quasi-static electric fields

Relatively large-scale field-aligned quasi-static DC electric fields may arise during the magnetic reconnection process in analogy to those observed in the Earth's auroral zone during magnetic substorms[18], although radio and X-ray observations suggest that magnetic reconnection is a more fragmentary process.[20] Super-Dreicer electric fields, which can accelerate both electrons and ions to the required energies in a relatively short distance, may be produced by mass advection into current sheets[19]. Alternatively, sub-Dreicer fields may also play a role in plasma heating and electron acceleration[15]. Finally, weak electrostatic double layers[24] distributed along the magnetic field may accelerate particles to significant energies.

Stochastic acceleration

Stochastic acceleration is a second-order Fermi process that involves particles scattering from waves, resulting in the diffusion of particles in energy and pitch angle[24,26]. A necessary ingredient is the presence of a broadband turbulent spectrum of waves (e.g., whistler waves or fast mode MHD waves) with which the particles can resonantly interact. Stochastic processes can yield plasma heating and/or particle acceleration. An injection or "seed" spectrum of particles is not necessarily required[26,28]. The spectral and angular distribution of the resulting electrons and ions depends on the details of the magnetic field, background plasma, the turbulent wave spectrum and its intensity[8,28].

Shocks

Shocks are produced in the Sun's corona and in the interplanetary medium. They are observed indirectly via coronal and interplanetary (IP) type II radio bursts[14]. IP shocks are observed directly by *in situ* measurements of particles and fields by space-based instrumentation. Coronal shocks are driven by flare blast waves and ejecta. Fast mode standing shocks associated with flare magnetic reconnection may also be observed[3]. IP shocks are driven by fast CMEs[9]. Both *shock drift acceleration* and *diffusive shock acceleration* are likely relevant.

Particle Injection and Transport

The processes of particle acceleration, injection, and transport are coupled. Access to the distribution of electrons that escape from the acceleration region – their number, anisotropy, energy distribution – is a critical link between the acceleration problem and the transport problem. Energetic particles are highly mobile. Yet a variety of factors affect their subsequent transport and, as a consequence, the transport of energy within and from the environment in which they are accelerated.

Magnetic trapping

The bulk of energetic particles accelerated in a flare remain on the Sun. They are accelerated and injected into coronal magnetic loops. Those particles with sufficiently large pitch angles mirror and remain trapped in the coronal loop for a time; those with small pitch angles "precipitate" from the loop and collide with dense material near the foot points of the loop where they liberate their energy. Trapped particles suffer Coulomb

collisions with the ambient plasma and eventually scatter to small pitch angles (weak diffusion) at which point they, too, are lost from the trap. This so-called "trap plus precipitation" model[21] and its variants[1,2] have been quite successful in accounting for the observed distribution of radio and X-ray emission from flaring loops and their evolution in time.

Wave-particle interactions

Stochastic acceleration may play an important role in particle acceleration. The magnetic reconnection process may drive a turbulent cascade[25]. Alternatively, or in addition, beamed distributions of particles are unstable to the production of plasma waves. Wave-particle interactions subsequently cause particles to diffuse in momentum and pitch angle. Consequently, scattering on turbulence may strongly affect the transport of fast particles, possibly confining them. Indeed, acceleration and transport are closely intertwined in these circumstances[6,28]. The electrons responsible for HXR emission in the impulsive phase may indeed represent energy transported as Poynting flux.[12]

Energy transport and loss

Most nonthermal electrons eventually lose their energy to Coulomb collisions with the relatively cool chromospheric plasma. Since only a small fraction of the energy is emitted as nonthermal HXR radiation ($\sim 10^{-5}$), most of the energy in nonthermal electrons goes toward plasma heating. The chromospheric plasma responds dynamically, a process given the misnomer "chromospheric evaporation", which fills the magnetic loops with hot plasma (~20 MK), which emits copious thermal SXRs. Therefore, a significant fraction of the energy going into accelerated particles in flares is ultimately radiated away in the SXR band although thermal conduction and mass motions are also important components of the energy budget. A small fraction of flare-accelerated electrons and ions escape into the interplanetary medium. Other populations of electrons and ions are accelerated to high energies by a shock driven by CME that is associated with the flare.

The problem of magnetic energy release, particle acceleration, and energy transport in flares and CMEs involves a complex coupled system. Progress in understanding this coupled system requires observational access to each of its constituent parts.

Outstanding Questions

There are a number of important questions regarding particle acceleration on the Sun. The two most fundamental questions are:

1. What basic physical processes govern the acceleration of particles in the solar atmosphere?

2. Under what circumstances does this acceleration occur?

To answer these requires understanding the following subsidiary questions in detail:

- Where precisely does magnetic energy release occur? To date, the location of magnetic energy release through 3D magnetic reconnection has in most cases

been inferred indirectly or circumstantially, although tantalizing new HXR observation are emerging[29]. Additional signatures are urgently needed.

- Where precisely are electrons accelerated? The acceleration region may not be co-located with the magnetic energy release site, for example in the case that the Poynting flux dominates the energy transport.[7] Moreover, observational signatures (e.g., HXR radiation) may be relatively remote from the energy release and acceleration site. Again, complementary and/or more direct signatures are needed.

- What is the electron distribution function and what is its spatiotemporal evolution? The essential features of the electron distribution have been established through HXR, gamma-ray, and radio spectroscopy. Progress has also been made at HXR wavelengths in imaging spectroscopy[27]. However, measuring the detailed evolution of the electron distribution function has been hampered by a lack of adequate spatial and/or temporal resolution.

- With what efficiency is magnetic energy converted into nonthermal electrons? Magnetic energy release leads to plasma heating, electron and ion acceleration, and mass motions. The energy budget needs to be refined and reassessed.

- What physical processes affect the transport of energetic electrons? Under what circumstances? It is necessary to disentangle transport effects from the observed spatiotemporal evolution of the electron distribution function in order to constrain acceleration mechanisms and to understand the energy budget and energy transport fully.

New Insights at Radio Wavelengths

Radio observations are in a strong position[13] to make progress on each of these outstanding questions using a variety of innovative and unique techniques. We emphasize here the need to exploit such observations and techniques, which are complementary to studies at optical, UV, EUV/SXR, and HXR/gamma-ray wavelengths. In order fully exploit the diagnostic potential of radio emission from energetic electrons on the Sun, the means to perform **ultra-broadband imaging spectroscopy** is urgently needed. The utility and impact of an instrument with this capability will be profound. The key attributes of such an instrument are as follows:

- Exceptionally broad frequency coverage provides access to the entire solar atmosphere from the mid-chromosphere to the upper corona, providing the means of observing the flare/CME phenomenon as a coupled system.

- The instrument will perform coronal magnetography; that is, it will use broadband frequency coverage to make quantitative measurements of the coronal magnetic field before, during, and after flares and CMEs, allowing quantitative measurements to be made of the magnetic fee energy available.

- The location of nonthermal electrons will be established wherever and whenever they occur. Magnetic energy release is accompanied by multitudes of coherent radio bursts driven by suprathermal electron beams[4], possibly produced by the reconnection process. These will be imaged for the first time. Similarly, the location and distribution – *including anisotropic distributions*[10,11] – and evolution of accelerated electrons will be observed via nonthermal gyrosynchrotron, as will their transport to remote locations.

- The polarized gyrosynchrotron spectrum of the nonthermal electrons will be observed in microwaves in every pixel with high time resolution. Forward fitting techniques will in many cases allow extraction of the electron distribution function, as well as the magnetic field, ambient density, and temperature[6]. See also the white paper entitled *Uncovering Mechanisms of Coronal Magnetism via Advanced 3D Modeling of Flares and Active Regions* by Fleishman et al.

- ANY associated erupting filament and/or fast CME will be observed at decimeter to meter wavelengths through thermal bremsstrahlung and nonthermal gyrosynchrotron emission. The physical properties of the CME, including its magnetic field and the distribution of entrained energetic electrons will be inferred[5,21]. The associated shock and locations of electron acceleration along the shock will be observed via the type II radio burst.

An instrument with these attributes has already been defined: the *Frequency Agile Solar Radiotelescope* (FASR—see the white paper of the same name, by Gary et al.) The concept is mature and it is ready for implementation. Briefly, FASR is an interferometric array of antennas that would image the Sun over a frequency range of 50 MHz (λ=6 m) to 21 GHz (λ<1.5 cm), thereby imaging the Sun in 3D from the mid-chromosphere well up into the corona. It would do so with an angular resolution as high as 1", a time resolution as high as 20 ms (for coherent bursts) or 1 s (flares), and a frequency resolution up to 0.1%. FASR will transform the study of magnetic energy release, particle acceleration, and transport, as well as addressing key questions regarding the structure of the chromosphere and corona.

Broader Implications

FASR, a next-generation radioheliograph designed to perform dynamic, ultra-broadband, imaging spectroscopy will provide new traction on outstanding questions regarding magnetic energy release, particle acceleration, and particle transport. It will provide a view of these processes as a coupled system. With the development of new observations and new techniques will come new diagnostics of physical conditions on the Sun and new ways to forecast energetic transient phenomena – flares and CMEs. While FASR will initially be a critical element of the basic research portfolio in solar and heliospheric physics, it is anticipated that key elements of its science program will transition from basic research interests to operational components in the national space weather effort.

References

1. Aschwanden, M. J., Bynum, R. M., Kosugi, T., Hudson, H. S., & Schwartz, R. A. 1997, ApJ 487, 936
2. Aschwanden, M. J., Schwartz, R., & Dennis, B. R. 1998, ApJ 502, 468
3. Aurass, H., & Mann. G. 200?
4. Bastian, T. S., Benz, A. O., & Gary, D. E. 1998, ARAA 36, 131
5. Bastian, T. S., Maia, D., Pick, M., & Kerdraon, A. 2001, ApJL 558, 65
6. Bastian, T.S., Fleishman, G. D., & Gary, D. E. 2007, ApJ 666, 1256
7. Birn, J., Fletcher, L., Hesse, M. & Neukirch, T. 2009, Astrophys. J. 695, 1151
8. Bykov, A. M., & Fleishman, G. D. 2009, ApJL, 692, 45
9. Cane, H. V. 1984, A&A 140, 205
10. Fleishman, G., & Melnikov, V. 2003, ApJ 584, 1071
11. Fleishman, G., & Melnikov, V. 2003, ApJ, 587, 823
12. Fletcher, L. & Hudson, H. 2008, Astrophys. J. 675, 1645
13. Gary, D. E., and Keller, C. U. 2004, *Solar and Space Weather Radiophysics*, Springer/Kluwer, Astrophysics and Space Science Library, v. 314
14. Gopalswamy, N. 2000, in *Radio Astronomy at Long Wavelengths*, eds. R.G. Stone, K. W. Weiler, M. L. Goldstein & J.-L. Bougeret, AGU, p. 123
15. Holman, G. 1985, ApJ 293, 584
16. Li, Y., & Fleishman, G. D. 2009, ApJL 701, 52
17. Lin, R. 1985, Solar Phys.
18. Lin, R. P., & Schwartz, R. A. 1987, ApJ 312, 462
19. Litvinenko, Y. E. 1996, ApJ 462, 997
20. Litvinenko, Y. E. 2000, Solar Phys. 194, 327
21. Maia, D., Mercier, C., Pick, M, & Kerdraon, A. 2007, ApJ 660, 874
22. Mann, G., Aurass, H., & Warmuth, A. 2006, A&A 454, 969
23. Melrose, D. B., & Brown, J. C. 1976, MNRAS 176, 15
24. Melrose, D. B. 1990, Aust. J. Phys. 32, 703
25. Miller, J. A., Larosa, T. N., & R. L. Moore 1996, ApJ 461, 445
26. Miller, J. A., Cargill, P. J., Emslie, A. G., Holman, G. D., Dennis, B. R., et al. 1997, JGR 102, 14631
27. Park, S.-H., & Fleishman, G. D. 2010, Solar Phys. 266, 323
28. Petrosian, V., & Liu, S. 2004, ApJ 610, 550
29. Sui, L., & Holman, G. 2003, Astrophys. J. 596, 251

ASTEROSEISMOLOGY: THE NEXT FRONTIER IN SOLAR-STELLAR PHYSICS

A Science White Paper for the Solar and Space Physics Decadal Survey
Submitted for consideration by the Solar and Heliospheric Physics Panel

Mark S. Giampapa
National Solar Observatory
950 N. Cherry Ave.
POB 26732
Tucson, AZ 85726-6732

giampapa@nso.edu
520-318-8236

Conny Aerts, K.U.Leuven (Belgium)
Tim Bedding, University of Sydney (Australia)
Alfio Bonanno, INAF/Catania Astrophysical Observatory (Italy)
Timothy M. Brown, Las Cumbres Observatory Global Telescope
Jørgen Christensen-Dalsgaard, Aarhus University (Denmark)
Martin Dominik, University of St. Andrews (UK)
Jian Ge, University of Florida
Ronald L. Gilliland, Space Telescope Science Institute
J. W. Harvey, National Solar Observatory
Frank Hill, National Solar Observatory
Steven D. Kawaler, Iowa State University
Hans Kjeldsen, Aarhus University (Denmark)
D.W Kurtz, University of Lancaster (UK)
Geoffrey W. Marcy, University of California Berkeley
Jaymie M. Matthews, University of British Columbia
Mario Joao P. F. G. Monteiro, University of Portugal
Jesper Schou, Stanford University

1. Introduction

The Sun is a star, and understanding the solar interior opens up new paths to understanding the interior of all stars. Conversely, understanding other stars places the Sun in an astrophysical context and increases our understanding of its likely future evolution. Stars span a range of parameter space that is inaccessible through the study of the Sun alone. Given that the Sun is the engine that drives global climate and controls the near-earth orbital and geomagnetic environments, and thereby affects all space-based activities, understanding the origin and range of solar variability becomes imperative. The study of stars similar to our Sun is a practically feasible approach toward both understanding and forecasting the long-term behavior of the Sun. Helioseismology is among the modern tools we utilize to study the Sun. Asteroseismology is the natural extension of the helioseismic study of the solar interior to the probing of the interior structure of late-type stars, including stars that are analogs of our Sun. Asteroseismology has become a very active field with the launches of *CoRoT* and *Kepler*, and the interplay between helio- and asteroseismology enriches both endeavors.

The study of the interior structure of stars historically has relied on a combination of photospheric spectral diagnostics and models. However, with the now mature development of the techniques of helioseismology, along with technological advances that yield high-precision measurements of intensity fluctuations and spectral line positions in stellar spectra at high cadence, the emerging field of asteroseismology now enables us to embark upon the direct, observational study of stellar interiors.

Asteroseismology allows us to obtain quantitative information on the internal structure of stars, to compare stellar models to real stars more extensively than heretofore possible, to explore the behavior of matter under conditions that cannot be achieved on Earth, and to confront discrepancies between stellar ages and estimates of the cosmological age from a new perspective. The foundation of modern astrophysics, most especially the distance scale to the galaxies and the ages of stellar clusters, is based on a quantitative understanding of the fundamental parameters of stars: mass, age, chemical composition, and mixing length. The detection and measurement of acoustic (*p*-mode) oscillations in late-type stars offers a new approach to measuring these parameters with significantly higher precision than has previously been possible, thus enabling quantitative comparison of data with stellar models. Thus, the direct study of stellar interiors through the techniques of asteroseismology is of broad astrophysical relevance and importance.

While this white paper is focused on the seismology of solar-like stars there are several other types of pulsating stars that are distinctly non-solar, and that have been very successfully probed with seismic methods for some time. Examples of these stellar classes are the pulsating white dwarfs, the rapidly oscillating Ap stars, and the δ Scuti stars. The diagnostic and observational approaches described herein are applicable throughout the H-R diagram.

Stars can be characterized by five intrinsic parameters, plus distance: mass, helium fraction, metallicity, age, and the mixing length ratio. In principle, measurement of the frequency separations of *p*-mode oscillations (i.e., sound waves) permits estimation of the age and mass of

a star with useful precision if the metallicity is known from other means. However, the uncertainty in stellar metallicity determinations increases the uncertainties in masses and ages determined from oscillation frequencies to significant levels. Additional information, such as the actual individual oscillation frequencies or other data which can help constrain the intrinsic parameters, is needed. In their study of the application of oscillation data to the determination of stellar parameters, Brown et al. (1994) concluded that the determination of ages, mixing lengths, and masses can be substantially improved when oscillation data can be included with traditional measures, especially for more distant stars for which astrometric distances cannot be obtained. *The application of asteroseismology techniques to stars ultimately will allow us to identify and investigate true stellar analogs of the Sun. This, in turn, will enable comparative studies of the influence of parameters such as rotation and convection zone structure on dynamo-related magnetic activity at all relevant time scales, from spot formation and evolution to cycles.*

2. Background

Stellar oscillations can be described (for spherical stars) as radial eigenmodes multiplied by a spherical harmonic, with two quantum numbers, *n* and *l*. The radial order *n* is the number of nodes in the eigenfunction between a star's center and its surface. For typical solar-type stars, *n* is of order 20. The angular quantum number *l* describes the number of nodal planes that intersect the surface of the star. For observations in integrated light, *l* is usually small ($l = 0,1,2$, or perhaps, 3). Differences in frequency of oscillation of these modes are related to the intrinsic stellar parameters. Frequency separations of the *n*-modes are typically large, and are related to the time for a sound wave to cross the star. Through further consideration of the virial theorem, it can be shown that this travel time is related to the mean density of the star. Thus, the separation of the *n*-modes, known as the “large separation”, is easily interpreted in terms of stellar structure and it is relatively straightforward to observe.

Frequency separations of the *l*-modes are smaller and are related to the radial gradient of the sound speed, particularly in the stellar core. Since these gradients change as nuclear burning alters the molecular weight distribution in and around the core, the so-called “small separation” contains information about the evolutionary state of the star. Measurement of the *l*- and *n*-mode frequency separations can yield good estimates of mass and age when combined with stellar models and other observational parameters. The information content for stars in these frequency separations is summarized in Fig. 1 below from Christensen-Dalsgaard (2004). The source of energy for the *p*-mode oscillations themselves is acoustic noise generated by high-speed convective motions within a few scale heights of the stellar surface. Hence, stars later than about spectral type F5, i.e., stars with vigorous outer convection zones, should give rise to *p*-modes that are generally similar to those observed in the Sun.

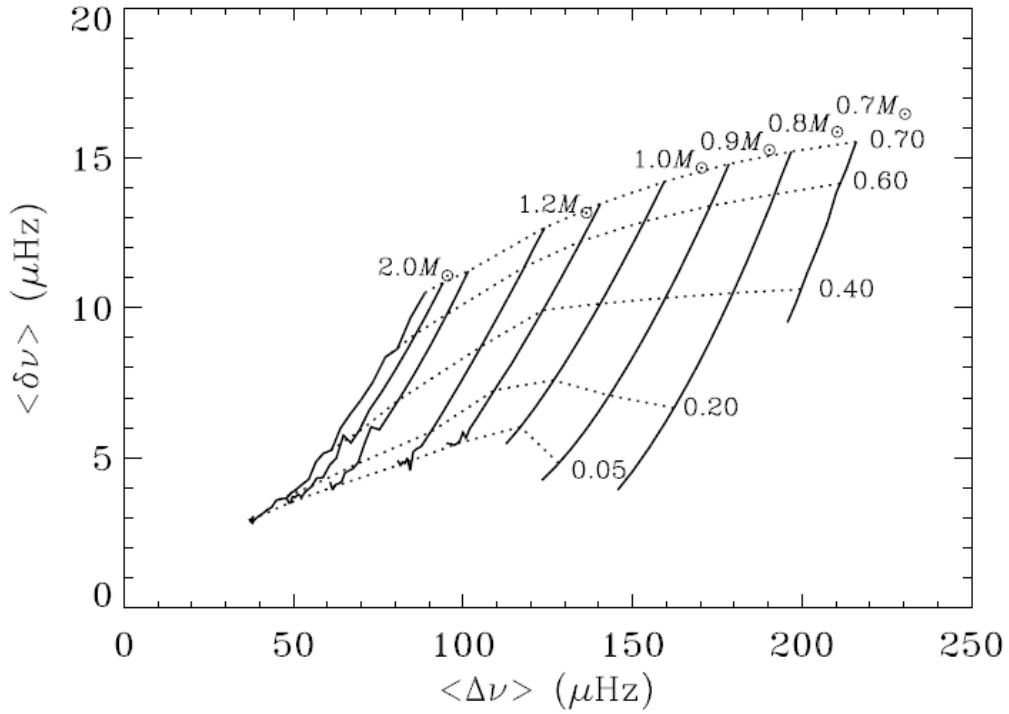


Figure 1 The "asteroseismic H-R diagram" showing the small frequency separation as a function of the large frequency separation along with stellar mass and age. Mass is constant along solid lines; age (parameterized by the central hydrogen abundance) is constant along dotted lines (adapted from Christensen-Dalsgaard 2004).

3. Results

As helioseismology celebrates its diamond anniversary, we now find asteroseismology entering a rapid phase of development, similar to that of helioseismology ~ 25 years ago. Initially, though, progress was slow. As comprehensively reviewed by Aerts et al. (2008), the search for solar-like oscillations in stars in the solar neighborhood has been ongoing since the early 1980s. The first indication of stellar power with a frequency dependence similar to that of the Sun was obtained by Brown *et al.* (1991) in α CMi (Procyon, F5 IV). The first detection of individual frequencies of solar-like oscillations was achieved from high precision time-resolved spectroscopic measurements only in 1995 for the G5 IV star η Boo (Kjeldsen *et al.* 1995; Brown *et al.* 1997); however, a confirmation of this detection from independent measurements could not be established, but it was subsequently confirmed by Carrier et al. (2003) and Kjeldsen *et al.* (2003). It took another four years before solar-like oscillations were definitely established in Procyon (Martíc *et al.*, 1999). Since then, solar-like oscillations were found in two more stars: the G2 IV star β Hyi (Bedding *et al.* 2001) and the solar twin α Cen A (Bouchy and Carrier 2001; Schou and Buzasi 2001). These important discoveries led to several more subsequent detections, a summary of which is encapsulated in Fig. 2 below.

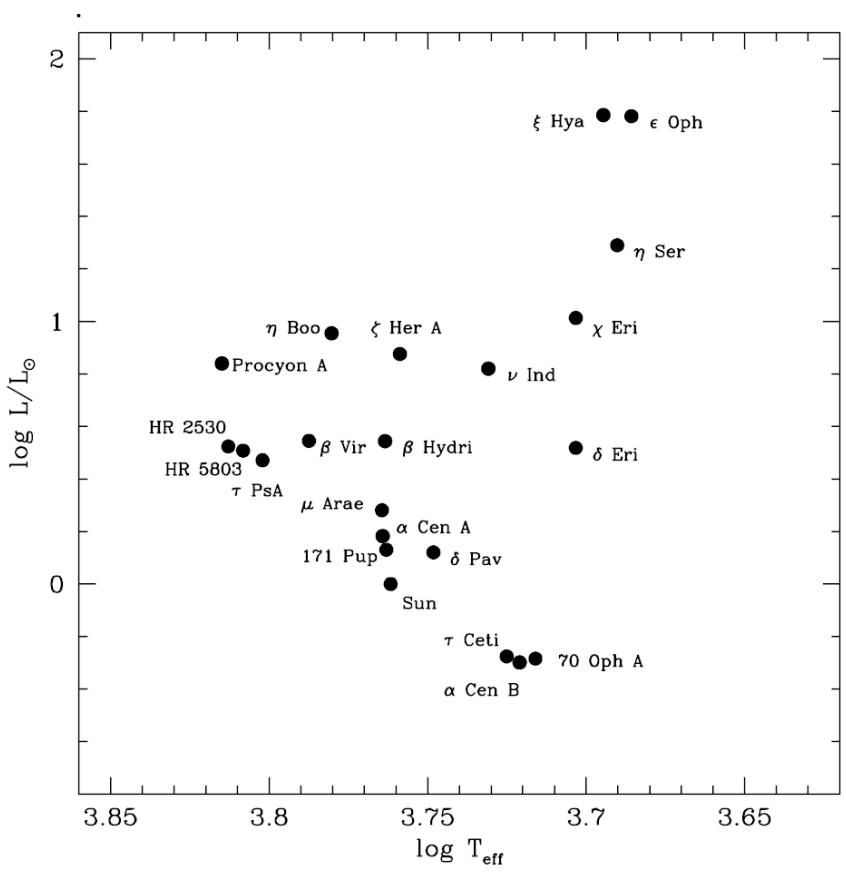


Figure 2 H-R diagram showing the stars in which solar *p*-mode oscillations have been detected (from Aerts et al. 2008, courtesy of Fabien Carrier).

.

The oscillation frequencies and frequency separations detected in solar-like "pulsators" provide additional constraints with which to test models of stellar structure and evolution in conditions slightly different from those provided by the Sun. Such studies generally involve a fit of theoretical models, characterized by a number of model parameters, to the set of seismic and non-seismic data available for a given pulsator. Theoretical modeling of solar-like pulsators using this direct fitting approach has been carried out for several stars; so far, the main results of these fits are estimates of the stellar masses, ages, and initial metallicities, even though in some cases the results are still controversial and call for better data-sets.

Among the solar-like pulsators, the binary star α Cen A and B provides a particularly interesting test bed for studies of stellar structure and evolution, given the numerous and precise seismic and non-seismic data that are available for both components of the binary. Studies of α Cen A and B, including seismic and non-seismic data for both components, indicate that the age of the system is likely to be between 5.6 and 7.0 Gyr, the value derived being dependent on the seismic observables that are included in the fits. A study by Teixeira *et al.* (in preparation) found that the mixing length was slightly *smaller* for α Cen B than for α Cen A. The latter study also found that the best-fitting model for α Cen A was on the border of having a convective core (see Christensen-Dalsgaard 2005): even a slight increase in the mass of the model led to a significant convective core and hence a model that was quite far from matching the observed properties. In another example of the power of stellar oscillations as a diagnostic of fundamental stellar properties, Teixeira et al. (2009) report the detection of *p*-mode oscillations in the G8 V star τ Ceti. The detection includes a tentative identification of modes up to $l = 3$. These investigators used the frequencies to estimate the mean density of the star to an accuracy of 0.45% which, combined with the interferometric radius, gives a mass of $0.783 \pm 0.012\ M_{\rm Sun}$ (1.6%).

Despite the case studies just outlined, in-depth seismic studies of stars with stochastically excited modes are currently still in their infancy compared with global helioseismology. However, given

the recent detections and the continuing efforts to improve them, and the new results emerging from *Kepler* and *CoRot*, we expect very substantial progress in the seismic interpretation of such targets in the coming years. As seen in the example of α Cen A, it is noteworthy that the class of main-sequence, solar-like oscillators encompasses transition objects with respect to the development of a convective core on the main sequence at masses between 1 and 1.5 solar masses. Asteroseismology can provide crucial insights on the yet poorly understood physics that occurs near the core of the objects in this transition region. Also, data of the expected quality will provide information about the depth and helium content of the convective envelope (*e.g.*, Houdek and Gough 2007), as well as more reliable determinations of stellar ages (Houdek and Gough 2008).

Additional information from solar-like oscillations is available in the cases of relatively evolved stars that are beyond the stage of central hydrogen burning. Here the frequency range of stochastically excited modes may encompass *mixed modes* behaving as standing internal gravity waves, or *g* modes, in the deep chemically inhomogeneous regions, thus providing much higher sensitivity to the properties of this region. In fact, there is some evidence that such modes have been found in the subgiant η Boo (Christensen-Dalsgaard, Bedding, and Kjeldsen 1995).

4. Observational Techniques

The two principal techniques that are utilized for the detection of oscillations in late-type stars are Doppler spectroscopy and photometry, with advantages and disadvantages associated with each approach. For example, Doppler methods work best for stars with many narrow spectral lines (implying stars with low v sin i's) while photometric techniques should work with any star. In addition, photometric techniques can be applied across a wide field of view, allowing asteroseismic data to be acquired for many stars simultaneously, while the requirement for the observation of many spectral lines at once means that Doppler spectroscopic techniques are practically feasible only for one star per observing sequence. Since vertical Doppler oscillations are naturally weighted more toward the disk center than the scalar intensity fluctuations, Doppler oscillations can be observed to degree $l = 3$ in stars while intensity oscillations can be seen only to degree 2. This difference in degree sensitivity could be used in principle to help identify the degree of modes observed on a star, which suggests that both intensity and Doppler observations should be carried out. From a purely asteroseismology perspective the Doppler observations are preferred because more modes are visible. In either case, high measurement precisions are required: of order meters per second for Doppler techniques with ~ 1000 spectral lines observed simultaneously, and sensitivities of parts per million for intensity oscillations.

The Sun exhibits non-oscillatory velocity and intensity fluctuations that produce a background "noise". The spectrum of this background noise depends on the average properties of the fluctuations and provides valuable information about these fluctuations. Nevertheless, this background limits the detection of oscillation modes. At the frequencies of *p*-modes in the Sun and sun-like stars, the background noise for both Doppler and photometric fluctuations is due almost entirely to granulation. With the Sun as a guide, we find that the signal-to-noise ratio of Doppler measurements compared to intensity measurements is ~ 40 times more favorable for Doppler measurements. Other stars may be different and the situation may differ at other frequencies, but it is reasonable to expect that Doppler observations are likely to reveal

oscillations at lower amplitudes than intensity observations across a wide range of stars. We note that the French-European *CoRoT* mission has acquired through intensity observations information on both the oscillation spectrum and the granulation "noise" spectrum of late-type stars during the course of its search for extrasolar planetary transits. In some stars, the power in the granulation spectrum has been observed with *CoRoT* to be substantially enhanced at factors ~ 5 over that of the power in the solar granulation spectrum. The *Kepler* mission already has obtained substantial quantities of intensity data that exhibits p-mode oscillations along with the pattern of granulation noise in addition to signatures of stellar microvariability due to magnetic activity (see Fig. 3 below).

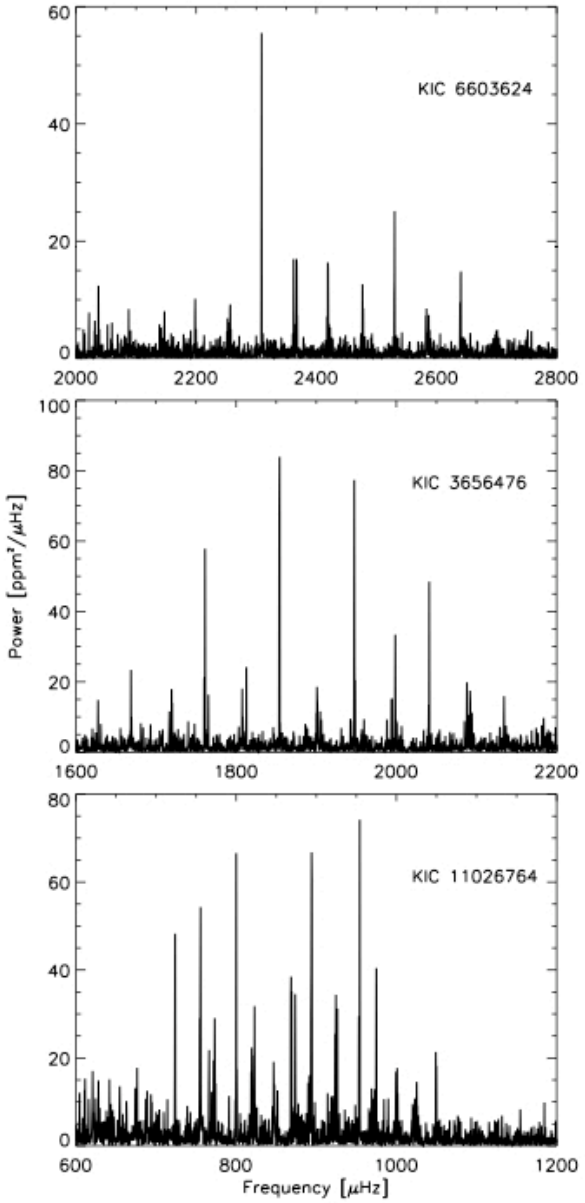


Figure 3 The detection of p-modes in three solar-type stars based on Kepler data (From Chaplin et al. 2010)

The accumulation of *Kepler* intensity data to date combined with that anticipated for the duration of the *Kepler* mission opens the possibility of searching for stellar-cycle-related variations of the

mode parameters in stars with short stellar cycle periods. García et al. (2010) recently presented the first evidence for the detection of stellar-cycle-related mode parameter variations in another star (*CoRoT* data on the F-type star HD49933); non-seismic results on another star with a short rotation period (Metcalfe et al. 2010) reinforce the idea that shorter cycle periods may go hand-in-hand with shorter rotation periods (e.g., see Böhm-Vitense 2007). Another area of frontier research is the search for g-modes in solar-type stars. The detection of g modes in solar-type stars would provide data for making much more accurate inference on the properties of stellar cores than is usually possible with p modes. The g modes are confined in cavities in the radiative interiors of stars, so that perturbations arising from these modes are extremely weak in amplitude by the time they reach the stellar photospheres (because the modes are evanescent in the convective envelopes). This has made the detection of g modes in the Sun an extremely challenging task observationally. The task might be easier in F-type stars, which have thinner convective envelopes. From our experience thus far, it is clear that complementary spectroscopic (Doppler) data are needed not just to confirm results already obtained but to significantly extend them.

5. The Next Step for Asteroseismology

It is well known that asteroseismology (and the study of stellar oscillations in general) suffers from the lack of facilities to provide long, uninterrupted time-series observations.
This is essential in order to provide adequate accuracy for the measured frequencies and to separate closely spaced frequencies. Asteroseismology, and other fields of astrophysics depending on observations in the time domain, such as the search for extra-solar planets, need dedicated facilities to achieve the required extensive and nearly continuous observations. The Stellar Observations Network Group (SONG) is a Danish initiative dedicated to overcome the problems of short observing runs and non-continuous data, by building a telescope network which will study solar-like oscillations in nearby, bright stars and search for planets around other stars. SONG will consist of 8 identical telescope nodes, with 1-m class telescopes, distributed globally to provide maximum temporal coverage, with the aim to achieve full sky coverage. The best observing duty cycle will be for the region around the celestial equator, where a duty cycle in excess of 80% will be possible. An overview of SONG can be found in Grundahl et al. (2008).

The instrumentation package will consist of a high-resolution spectrograph and an imaging camera which will be used for guiding during spectroscopic observations, but can also be used as a science imager. The high-resolution spectrograph will be optimized for measuring high-precision radial velocities and hence will be able to achieve very high precision despite the relatively modest size of the telescopes. The objective of the scientific operations is to observe targets for extended periods, up to ~ 4 months. This duration will enable exquisite frequency separation and precision to be achieved. Given an aperture of 1-m the asteroseismology targets for SONG will primarily be stars brighter than V = 6. This provides several stars in nearly all of the known classes of variables. We note here that since the targets are so bright it will be possible to measure precise radii for the SONG targets with existing interferometers. This will provide an important extra constraint on the stars under study that is not available from satellite missions such as *CoRoT* or *Kepler*. Furthermore, most stars will be quite close to the Sun, such that their parallaxes are well determined and, with spectroscopic data, oscillation modes with l =

3 can also be observed. SONG will not be limited to observing solar-like pulsators; it will also observe pulsating white dwarfs, RoAp stars, δ Scuti stars, etc.

We note that a recent campaign to measure solar-like oscillations in Procyon involved telescopes at 8 ground-based observatories (Arentoft et al. 2008; Bedding et al. 2010). This indicates the type of science that SONG would do routinely, and also the extent of international interest with more than 50 authors on the papers!

The design work for the SONG prototype unit consisting of the telescope and instrument was completed in 2008. Funding recently has been allocated for construction, assembly, integration and tests. The project anticipates that a fully tested and functioning prototype SONG node will be ready by the end of 2011, including the data reduction software. In parallel with the prototype development work, SONG wishes to establish an international collaboration, involving network sites and/or contribution of network nodes, such that the complete network can be established and ready for operation by 2014. In addition to the SONG initiative, the privately-funded Las Cumbres Observatory Global Telescope (LCOGT) is another notable example of a new facility that will pursue observations of stellar pulsations as part of its broader scientific program. LCOGT is now building a worldwide network of 1-meter telescopes that will be capable of both photospheric and spectroscopic measurements of classical pulsators, and also of stochastic pulsators that have relatively high amplitudes.

The SONG and LCOGT initiatives provide an excellent opportunity for the U.S. community to participate in, and contribute to, this forefront area of solar-stellar physics, leveraging its resources through the investment already made in the development of the prototype instruments. The rich experience of helioseismology in the development of seismic data analysis techniques and interpretive approaches, combined with the advances in technologies for high precision measurements of stellar spectra, have placed us at the threshold of discovery. By obtaining oscillation spectra of stars based on data of high quality and extending over several months, asteroseismology will enable critical tests of our understanding of interior structure, possible variations with rotation and activity, stellar evolution theory, the determination of stellar ages and helium abundance, and the identification of true solar analogs for comparison with the Sun. Without question, asteroseismology is the next frontier in solar-stellar physics.

6. References

Aerts, C., Christensen-Dalsgarrd, J., Cunha, M. & Kurtz, D.W. 2008, Sol. Phys., 251, 3
Arentoft, T. et al. 2008, ApJ, 687, 1108
Bedding, T. R. et al. 2001, ApJL, 549, L105
Bedding, T. R. et al. 2010, ApJ, 713, 935
Böhm-Vitense, E. 2007, ApJ, 657, 486
Bouchy, F. & Carrier, F. 2001, A&A, 374, L5
Brown, T. M. et al. 1991, ApJ, 368, 599
Brown, T. M. et al. 1994, ApJ, 427, 1013
Brown, T. M. et al. 1997, ApJ, 475, 322
Carrier, F. Et al. 2003, in *Asteroseismology Across the HR Diagram*, Kluwer, Dordrecht, 315
Chaplin, W.J. et al. 2010, ApJL, 713, 169

Christensen-Dalsgaard, J. 2004, Solar Physics, 220, 137
Christensen-Dalsgaard, J. 2005, in *Proceedings of the 13th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun*, eds. F. Favata et al., SP-560, ESA, Noordwijk, 81
Christensen-Dalsgaard, J., Bedding, T. & Kjeldsen, H. 1995, ApJ, 443, L29
García, R. A., Mathur, S., Salabert, D., Ballot, J., Régulo, C., Metcalfe, T. S. & Baglin, A. 2010, Science, 329, 1032
Grundahl, F. et al. 2008 in Journal of Physics: Conference Series 118**,** IOP Publishing
Houdek, G. & Gough, D. O. 2007, MNRAS, 375, 861
Houdek, G. & Gough, D. O. 2008, in *Unsolved Problems in Stellar Physics*, eds. Stancliffe, R. et al., Melville, in press
Kjeldsen, H. et al. 1995, AJ, 109, 1313
Kjeldsen, H. et al. 2003, AJ, 126, 1483
Martíc, M. et al. 1999, A&A, 351, 993
Metcalfe, T. S., Basu, S., Henry, T. J., Soderblom, D. R., Judge, P. G., Knölker, M., Mathur, S., & Rempel, M. 2010, ApJ, 723, 213
Schou, J. and Buzasi, D. 2001, in *Proceedings of the SOHO 10/GONG 2000 Workshop: Helio- and asteroseismology at the dawn of the millennium*, ed. A. Wilson, Scientific, ESA SP-464, Noordwijk,, p. 391 - 394
Teixeira, T. C. et al. 2009, A&A, 494, 237
Teixeira, T. C. et al., in preparation

CAUSES OF SOLAR ACTIVITY

A Science White Paper for the Heliophysics 2010 Decadal Survey

Submitted by

Mark S. Giampapa, National Solar Observatory
Frank Hill, National Solar Observatory
Aimee A. Norton, James Cook University (Australia) & NSO
Alexei A. Pevtsov, National Solar Observatory

1. Introduction

Our nearest star, the Sun, exhibits several processes that are common to other stars: magnetic cycles, flares, prominences, and stellar winds. Recent studies suggest that many of these processes are governed by universal laws applicable to both the Sun and other stars[1, 2]. This solar-stellar connection can be used to extend our understanding of stellar phenomena via the study of solar processes. Thus, for example, studies of the solar dynamo will yield a better understanding of stellar dynamos: for example, what are the necessary conditions for dynamo action and how are magnetic cycles maintained? The study of the magnetic topology of *solar* flares and prominences will contribute significantly to the understanding of the nature of *stellar* flares, prominences, and winds. Herein we argue that a sustained program of long-term, high-continuity observations of the solar magnetic field is required to reach a better level of understanding of the universal processes taking place throughout the Universe.

The magnetic field is a corner stone of solar activity. Sunspots, flares and coronal mass ejections (CMEs), coronal heating and solar wind acceleration, all bear on the topological properties of magnetic fields. The magnetic field itself originates in a dynamo operating deep inside the Sun. Recent progress in dynamo theory and numerical modeling brought us closer to understanding the dynamo on the Sun to the point of attempting to predict the magnitude of the next solar cycle. Nevertheless, there remain many unresolved questions about the solar dynamo(s). For example, we still know very little about the subphotospheric meridional circulation that is a key ingredient for solar subphotospheric dynamos. The magnetic field in the solar polar regions is an untapped territory. Both meridional circulation and polar magnetic fields are critical in resolving the enigma of the dynamo. Also, it appears that there are two separate dynamos operating in the Sun: a deep-seated helical dynamo, responsible for the generation of the strong magnetic fields of active regions, and a near-surface chaotic dynamo, producing weak network fields. How these two dynamos interact with each other and what their roles are in the overall solar cycle are unknown. With respect to other stars, helical dynamos are typical for solar-type dwarfs, while non-helical (distributed) dynamos can operate on other, e.g., T Tauri and low-mass, fully convective dwarf stars. Since our observational capabilities are still a long way from resolving the fundamental scales of magnetic fields on the surfaces of other stars, the Sun offers the unique opportunity to study these dynamo processes directly.

The current state of solar physics indicates an important convergence between numerical modeling and observations. The numerical capabilities are now reaching the state when realistic 3D models of solar processes can be constructed at a level allowing the direct comparison with high resolution observations. One should expect that within the next decade the progress in both computational capabilities and the observations will come close to solving the nature of many solar phenomena. With these numerical models, developed to explain solar phenomena, one can use the Sun as a giant astrophysics laboratory to study processes occurring on other stars.

To achieve the goals described above requires two key ingredients: a sustained program of long-term, high-continuity observations of the solar magnetic field coupled with a strong program of topical numerical simulations aimed on modeling solar phenomena. This white paper concentrates on the former.

2. Observational Imperatives

Briefly, what is needed to achieve a better understanding of the causes of solar activity is a long-term, comprehensive study based on high and medium resolution full-disk, vector magnetic field observations at several layers in the solar atmosphere. To understand what triggers flares/CMEs, we need to observe the reconnection site at the spatial resolution and time cadence that allows us to resolve the processes taking place at the time of the event. Neither flare locations nor times are known *a priori*, implying the need for high resolution observations of the entire solar disk. Similarly, the early stages of active region emergence (and all processes related to this) are usually missed. On longer time scales, systematic long-term observations are required to understand the change in magnetic field topology during the course of a solar cycle. The key lack on the ground is poor observational continuity with regard to measurements of vector magnetic fields at different heights in solar atmosphere.

Although some of these shortcomings could be remedied by space-borne instruments (e.g. the Solar Dynamics Observatory, SDO), reliance on a single spacecraft for synoptic data may lead to significant data gaps, as has happened in the past (e.g., the SOHO spacecraft failure). Furthermore, ground-based observational capabilities often serve as a springboard for the development of space-borne instrumentation and as a training ground for students and engineers. Maintaining a ground-based network of identical magnetographs will ensure the continuity of synoptic studies of the solar cycle. The current anomalous solar minimum teaches us that we need more than a single example of a cycle in order to not draw unjustified conclusions.

3. Origins of Solar Magnetic Activity

The broad goal of understanding of solar magnetic activity at all its spatial and temporal scales requires observations of the Sun with enough spatial resolution, sensitivity and sufficient temporal frequency and continuity to identify what processes produce the variations. This broad goal is best addressed through synoptic network observations of the following interlinked processes:

- **Origin and evolution of sunspots, active regions and complexes of activity;**
- **Sources and drivers of solar activity and disturbances;**
- **Properties of the solar dynamo on short and long time scales.**

A network of instruments with vector polarimetric capabilities, along the lines of the NSO SOLIS (Synoptic Long-term Optical Investigations of the Sun) Vector Spectromagnetograph (VSM), is required for the acquisition of high-precision magnetic field data across the solar disk through the emerging science of vector polarimetry.

Many recent scientific advances have been made through the application of vector polarimetry using, for example, the Advanced Stokes Polarimeter (ASP)/Diffraction Limited Spectropolarimeter (DLSP), Spectropolarimeter for Infrared and Optical Regions (SPINOR), SOLIS VSM and the Hinode Solar Optical Telescope (SOT). Horizontal magnetic fields in the photosphere were found to have a higher-than-expected flux density distribution in quiet-Sun regions[4,5] while granular sized horizontal fields are observed to be associated with granular

convection in plage region[6]. Meanwhile, the study of active region geometries has shown small-scale complexities such as field-wrapping around penumbral filaments[7] and the observed emergence of a helical flux rope under an active region prominence that suggests that helicity assists in maintaining prominences[8]. The magnetic topology of the quiet-Sun with respect to the height-distribution of null points is also being examined using spectropolarimeters to better understand energy deposition and heating. New results show that the majority of null points are expected in the lower photosphere (54%), less in the chromosphere (44%) and with nearly insignificant frequency (2%) in the corona[9]. Simultaneous efforts by Deluca et al.[10] examine data from coordinated observations with Solar Optical Telescope, X-Ray Telescope, and the SOLIS VSM to determine the characteristics of the magnetic field topology that create bright points.

The last few years of solar spectropolarimetric measurements have furthered interactions between theory, modeling and data analysis. For example, observations of kG field strengths appearing in the presence of strong downflows are evidence for the convective instability mechanism of flux tube creation in the photosphere[11]. Observed topologies of Stokes asymmetries in network cell interiors may be the first evidence of boundary layers predicted years ago by numerical simulations[12]. Also, spatially resolved spectropolarimetric observations in the quiet Sun support the downward revision in the solar oxygen abundance[13], in conflict with the success of traditional composition used by helioseismologists in solar interior models. These observations have re-invigorated discussions between modelers and observers. In brief summary to this point vector magnetic field data are an asset to the solar physics and solar-terrestrial communities, and extensive databases are in short supply. A global network of vector instruments will produce for the community an archive that is long overdue that contains temporally and spatially extensive coverage of photospheric and chromospheric vector magnetic field data.

4. Why a Global Network?

The Sun varies on a wide range of time scales from billions of years to milliseconds and on spatial scales from the entire heliosphere to less than a kilometer. The variations are manifested in electromagnetic radiation from radio frequencies of a few kHz to powerful gamma rays. Particle emission ranges from the relatively gentle solar wind to blasts of charged and neutral particles that reach Earth at virtually the speed of light. A synoptic program concentrates on observations of the Sun over time scales of minutes to decades and spatial scales from about 1 Mm to the full solar disk. Why are synoptic network observations needed? Fundamentally, we cannot understand the Sun by single or sporadic observations, no matter how good they are. Here are some of the reasons:

1. Observations from a single site on Earth are a poor match to the frequency and duration of many forms of important solar activity such as magnetic field evolution leading to flares and coronal mass ejections. A network of similar instruments is the solution to this problem. Sustained observations from a global network are required to catch short-lived but infrequent phenomena such as flares and coronal mass ejections.

2. Solar activity is not simply the endless repetitions of the same events. The statistical properties of the various phenomena convey vital information about their origins and can be utilized to forecast energetic events.
3. Solar data at a regular cadence is often critical in helping to interpret short-lived campaign observations, i.e., it is important to put high-resolution observations into context.
4. Sampling that is too isolated in time can give misleading results. An example is the observation made from Skylab for several months in the 1970s, from which it was concluded that coronal holes rotate nearly rigidly. Later observations showed that this was true only for that particular, brief phase of the solar cycle.
5. Some phenomena are intrinsically periodic such as the magnetic activity cycle. We still do not have enough well-observed periods or the proper kinds of observations to understand the degree to which the cycle is quasi-periodic or chaotic.

4.1. Origins of Solar Activity

Our understanding of solar magnetic activity is far from complete. It is commonly believed that the principal driver of solar activity is stressed magnetic field. The stresses are released in the solar corona, producing explosive flare events. Flares usually occur in areas where the magnetic configuration is complex, with strong shears, high gradients, long and curved neutral lines, etc.[16] This implies that the trigger mechanisms of flares are controlled by critical properties of the magnetic field that lead eventually to MHD instabilities. But what kinds of instability actually govern, and under what conditions they are triggered, are unknown[17] Vector magnetic field measurements are necessary to infer field topology and vertical electric current, both of which are essential to understand the flare process[18]. Observations are required that can continuously track changes in the magnetic field and electric current with sufficient spatial resolution to reveal changes of field strength and topology before and after flares. The potential appearance of recurrent patterns in the local vector field evolution is the first step in the development of empirical predictors of flare activity. A vector network can provide these unique measurements of the vector magnetic field in specific flare active regions with high accuracy and at high cadence.

In addition to flares, coronal mass ejections (CMEs) are energetic eruptive events associated with heliospheric and geomagnetic disturbances. CMEs and flares were at first thought to be directly connected, with the flare driving the CME. However, only 60% of larger flares are associated with CMEs[21]. Similarly, many CMEs are not associated with flares. It is now thought that CMEs and associated flares are caused by a common event (i.e., the CME peak acceleration and the flare peak radiation often coincide). In general, all of these events (including the CME) are thought to be the result of a large-scale restructuring of the magnetic field. In particular, CMEs apparently interact with the global-scale magnetic field[22], but many CMEs, especially fast CMEs, are associated with flares, which are considered local phenomena. Model-based reconstruction of 3D magnetic structure is one way to estimate the field from observations[25]. Models using vector field data in active regions provide the best match to the observations. More realistic MHD coronal models[26] based on high-cadence vector-field maps as boundary conditions will greatly enhance our understanding of how the corona responds to evolving, non-potential active regions.

The origin of solar transient events and quasi-static structures ranging from active complexes to bipolar spot groups is intimately related to the emergence of magnetic flux. Flares are associated with the emergence of magnetic flux within active regions. Emerging magnetic regions near filaments leads to the eruption of the filaments[27]. CMEs are also found to accompany emerging flux regions. However, emergence of isolated active regions can proceed without any eruptive events. This suggests that magnetic flux emerging into the atmosphere interacts with pre-existing fields leading to a loss of magnetic field stability. Observations of electric current and magnetic topology differences between emerging and existing fields may lead to an understanding of why emerging flux causes transient events[28].

The high sensitivity, for example, of the SOLIS VSM instrument allows us to see much of this flux, extending from the ubiquitous small regions of mixed polarity, termed the 'magnetic carpet', to the larger magnetic structures. In addition, a network of such instruments would allow global-scale observations of the small-scale element distribution, their interactions on short and long time scales, and the resulting transformation of the large-scale field.

Many groups are actively engaged in research utilizing vector magnetogram data to gain further insight, and therefore predictive capability, concerning eruptive solar phenomena. It is generally accepted that the source of the energy released in solar transient events such as flares and CMEs results from the stored magnetic energy in an active region, where the solar magnetic field departs from a simple potential configuration. The best measure of magnetic non-potentiality can be provided by vector magnetograms. Global MHD models that reconstruct the 3D nature of the corona show improved ability to reproduce the eruption and evolution traits of specific CME events[30]. These models are sufficiently comprehensive to incorporate the non-potentiality observed by vector magnetograms into the initial configuration of the magnetic field prior to eruption.

The application of vector magnetograms specifically to active regions that could be the sites of coronal mass ejections (CMEs) is especially crucial to understanding the origins of these powerful events. The high sensitivity of SOLIS/VSM-equivalent synoptic observations will enable the investigation of whether the magnetic evolution of an active region as represented by the time-rate-of-change of the total non-potentiality can provide a higher CME-prediction success rate[34]. Therefore, extending these pioneering studies with a VSM-like network as the cycle is rising to solar maximum becomes particularly timely. In brief summary, we expect that the combination of global MHD modeling, CME and flare prediction capability, and routine vector magnetogram observations from a vector magnetic field network will result in more accurate models of these explosive solar events.

4.2. Sources and Mechanisms of Solar Irradiance Variations

Besides sustaining life on Earth, it is well known that spectral irradiance from the Sun has strong effects on the upper and middle terrestrial atmosphere. Effects of variations of total solar irradiance are harder to detect in the troposphere. The ultimate source of solar spectral and total irradiance variations is the changing solar magnetic field. Most of the variations can be explained by magnetic fields in the form of cool sunspots and hot faculae and plages with occasional transients from flares[35]. Attempts have been made to model solar irradiance variations, starting

with the sunspot and facular symptoms of the magnetic field, with impressive success[36]. But what works for one cycle does not work for another[37], pointing to a need to better understand how the Sun distributes its magnetic field between sunspots and faculae and plages. The long-term availability of high-quality vector magnetograms of solar activity will help solve this dilemma.

5. Estimated Usage

A global network of vector instruments will serve a large international community of users. In order to estimate the usage of the VSM network data, consider the history of the line-of-sight magnetic field instruments. Mapping of the solar magnetic field started with one instrument and a few investigators. More instruments and scientists rapidly followed with sustained observations of the LOS field, starting at NSO in 1973 and continuing today. These observations are freely available with a result that from 1974–2004 more than 1,206 research papers and theses (involving more than 983 unique users from every country with a solar research program, including a Nobel Laureate and a current member of Congress) used data from the NSO synoptic magnetic field program. The program has been supported by both NOAA's Space Weather Prediction Center and NASA for nearly 30 years in order to provide near-real-time data in support of those agencies' missions. We see a similar path emerging for the full vector field.

Observations of the vector magnetic field are now moving from the exploratory to sustained effort phase. We project a large growth in users of vector data as it becomes available on a regular basis. We studied the usage of three exploratory instruments to set a baseline for future growth. These include the spectrograph-based Advanced Stokes Polarimeter located at NSO and operated by the High Altitude Observatory, and filter-based vector magnetographs of the Solar Flare Telescope (SFT) located at the National Astronomical Observatory of Japan, and the Imaging Vector Magnetograph (IVM) located at Haleakela and operated by the Institute for Astronomy of the University of Hawaii. Other notable filter-based instruments were operated by Big Bear Solar Observatory and NASA's Marshall Space Flight Center. Only the SFT has a readily available public archive.

Table 1. Usage of Three Exploratory Vector Magnetographs (Source: ADS)

Instrument	*Service Dates*	*Refereed Papers*	*Citations*	*Unique Authors*	*Countries*
ASP (spectro)	1991 – 2006	57	1626	81	9
SFT (filter)	1991 - 2007	42	491	67	5
IVM (filter)	1992 - 2006	50	577	92	9

As expected, these usage values are small fractions of those of the sustained NSO LOS synoptic program. The SOLIS VSM is essentially a full-disk version of the ASP. We estimate that a VSM network would have usage at a rate up to an order of magnitude greater than ASP.

5.1. *Community Support.* A VSM network was the first-ranked, ground-based project in the Small Projects category in the 2000 NAS/NRC report *Astronomy and Astrophysics in the New Millennium,* i.e., the Decadal Survey. Specifically, it was the primary recommendation for ground-based initiatives in the small size category: *"The expansion of SOLIS to a 3-station network around the globe -$4.8 million. To gain near continuous coverage in full-disk solar*

vector magnetic field monitoring, as the backbone of an assessment of the solar magnetic field flux budget over the solar cycle."

The NSF Division of Astronomical Sciences (AST) Senior Review Committee report (issued in 2006) commissioned by the NSF Mathematical and Physical Sciences Advisory Committee and entitled, *"From the Ground Up: Balancing the NSF Astronomy Program,"* recommended the following with respect to the SOLIS VSM: *"The SR [Senior Review] supports the proposal to construct two additional copies of SOLIS [VSM] in order to ensure more complete temporal coverage."*

6. Summary

The advantages of a VSM-like global network over the single-site currently in operation include:

- **Nearly continuous data on magnetic field evolution.**
- **Two-three times better chance of capturing rare events.**
- **Improved potential for short-term activity forecasts.**
- **Improved coverage for comparison with helioseismology.**
- **Better contextual data for space missions such as Hinode, Solar Orbiter, Solar Dynamics Observatory, etc.**
- **Ability to detect and correct systematic data errors.**
- **Robustness against a single-site failure.**
- **Fosters international scientific collaboration.**
- **Stimulates stronger research programs on solar activity.**
- **Better constraints on theoretical models of activity.**
- **As a ground-based system, has a potentially much longer lifetime than any space experiment can ever offer.**

7. References cited

1. Pevtsov A.A. et al, 2003, ApJ, 598, 1387-1391.
2. Schwadron, N. A. & McComas, D.J, 2008, ApJ, 686, L33.
4. Harvey, J.W. et al., 2007, ApJ, 659, L177-L180.
5. Lites, B.W. et al., 2008, ApJ, 672, 1237-1253.
6. Ishikawa, R. et al. 2008, A&A, 481, L25-28.
7. Borrero, J.M. et al. 2008, A&A, 481, L13-16.
8. Okamoto, T.J. et al. 2008, ApJ, 673, L215-218.
9. Regnier, S. et al. 2008, A&A, 484, L47-50.
10. Deluca, E.E. et al. 2007, AGU, SH53A, 1058.
11. Nagata, S. et al. 2008, ApJ, 677, L145-147.
12. Rezaei, R.et al. 2007, A&A, 476, L33-36.
13. Socas-Navarro, H. and Norton, A.A., 2007, ApJ, 660, L153-156.
16. Sakurai, T. and Hiei, E., 1996, AdSpR, 17, 91-100.
17. Sammis, I., Tang, F. and H. Zirin, 2000, ApJ, 540, 583-587.
18. Canfield, R.C. et al. 1993, ApJ, 411, 362-369.
21. Andrews, M. D., 2003, Sol. Phys., 218, 261-279.
22. Luhmann, J.G. et al. 1998, JGR, 103, 6585-6594.
25. Falconer, D.A., Gary, G.A., Moore, R.L., Porter, J.G., 2000, ApJ, 528, 1004-1014.
26. Lee, J., White, S.M., Kundu, M.R., Mikic, Z., and McClymont, A.N., 1999, ApJ, 510, 413- 421.
27. Wang, Y.-M. and Sheeley, N.R., Jr., 1999, ApJ, 510, L157-L160.
28. Priest, E.R. and T.G. Forbes, 2002, A&ARev, 10, 313-377.
30. Riley, P., Lionello, R., Mikic, Z., and Linker, J., 2008, ApJ, 672, 1221-1227.
34. Falconer, D. A., Moore, R. L. and Gary, G. A. 2007, BAAS, 210**,** 135.
35. Woods, T.N. and Lean, J., 2007, EOS Transactions, AGU, 88, 44, 457.
36. Fröhlich, C. and Lean, J., 2004, A&ARv, 12, 273-320.
37. de Toma, G. et al. 2004, ApJ, 609, 1140-1152.

What Composition Measurements Could Have Done For Solar Probe Plus

A white paper submitted to the Solar and Space Physics Decadal Survey

Jason A. Gilbert
Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan

I. The Solar Probe Plus Mission

The Solar Probe Plus (SPP) mission will travel closer to the Sun (9.5 Solar Radii) than any previous spacecraft. This provides an unparalleled opportunity to explore the structure and origin of the solar wind, including its acceleration and heating. Several reports commissioned by the National Research Council (National Research Council 2003) and NASA (2009 Heliophysics Roadmap Team 2009; Solar Probe Science and Technology Definition Team 2005; Solar Probe Science and Technology Definition Team 2008) have detailed the wealth of scientific information expected from SPP.

The SPP Announcement of Opportunity (AO) (National Aeronautics and Space Administration 2009) includes an Ion Composition Analyzer in the list of reference instruments considered as "being of highest priority". The SPP payload selected in September 2010 by NASA is capable of measuring energetic ions, electric and magnetic fields, radio emissions, electrons, alphas, and protons, and includes an imager. Unfortunately, the ability to measure the composition of solar wind plasma or pickup ions is absent. Composition measurements are vital to the success of the SPP mission, and without them the science return will be severely restricted. The AO offers this summary of Science Questions (Q) that must be answered to address the Science Objectives (SO) of such a mission:

1. SO: Determine the structure and dynamics of the magnetic fields at the sources of the fast and slow solar wind
    a. Q: How does the magnetic field in the solar wind source regions connect to the photosphere and the heliosphere?
    b. Q: How do the observed structures in the corona evolve into the solar wind?
    c. Q: Is the source of the solar wind steady or intermittent?
2. SO: Trace the flow of energy that heats the solar corona and accelerates the solar wind
    a. Q: How is energy from the lower solar atmosphere transferred to, and dissipated in, the corona?
    b. Q: What coronal processes shape the non-equilibrium velocity distribution observed throughout the heliosphere?
    c. Q: How do the processes in the corona affect the properties of the solar wind in the heliosphere?

3. SO: Explore mechanisms that accelerate and transport energetic particles
    a. Q: What are the roles of shocks, reconnection, waves, and turbulence in the acceleration of energetic particles?
    b. Q: What are the source populations and physical conditions necessary for energetic particle acceleration?
    c. Q: How are energetic particles transported radially across magnetic field lines from the corona to the heliosphere?
4. SO: Explore dusty plasma phenomena in the near-Sun environment and their influence on the solar wind and energetic particle formation

II. The Strength of Composition Measurements

The ability to measure ion composition has been at the heart of space plasma research for decades, and has provided fundamental scientific results (Gloeckler et al. 1993; Gloeckler and Geiss 1996). Ion composition measurements would have greatly contributed to the mission of SPP, and would have been used in collaboration with the other SPP measurements to address all of the SPP Science Objectives. The elemental, isotopic, and charge state composition of the major elements in the solar wind (from He-Fe) have never been measured inside of 1 AU.

One area of active research in heliophysics involves the identification of the origins of the different types of solar wind. Whether it is coming from coronal hole boundaries (Wang, Ko, and Grappin 2009), near active regions (Neugebauer et al. 2002), being released from magnetic loops that interact with the closed corona (Fisk and Zurbuchen 2006), or any other source, acceleration and deceleration in the heliosphere prevent accurate identification of the "slow" wind by its speed alone. The compositional charge state signature of the plasma, the velocity distribution functions (VDFs) of heavy ions, as well as a measure of the ratio of elements with a low vs. high First Ionization Potential (FIP) such as Mg/O, are all used to characterize the different types of wind (Zhao, Zurbuchen, and Fisk 2009; Geiss, Gloeckler, and von Steiger 1995).

The wealth of scientific knowledge that can be gained by having complete radial profiles of the solar wind and pickup ion populations to 9.5 Rs is vast. Strong constraints could be placed on theories of solar wind and inner-source pickup ion origins; the evolution of the solar wind could be traced from SPP out to the several spacecraft located at farther radial distances. There would also be the mutual benefit of composition measurements on ESA's concurrent Solar Orbiter mission, which carries the NASA-funded Heavy Ion Spectrometer and will reach a perihelion of 50 Rs. While these composition measurements would be an invaluable resource for heliophysics in general, is it useful to examine how they would impact the SPP mission specifically.

III. SPP Science Questions Addressed by Composition Measurements

Q1a: How does the magnetic field in the solar wind source regions connect to the photosphere and the heliosphere? By using solar wind type as a tracer, the charge state composition and FIP measurements would provide information on the source regions of the measured plasma. This information, when combined with the magnetic field data and remote sensing of the corona, would achieve an accurate mapping of the source regions of the solar wind.

Q1b: How do the observed structures in the corona evolve into the solar wind? Composition anomalies in the solar wind can be caused by large-scale structures that evolve in the solar corona. Coronal Mass Ejections (CMEs), for example, are characterized not only by their magnetic signatures, but also by their ion charge states (Lepri and Zurbuchen 2004), which are established near the Sun and remain unchanged as the CME propagates through the heliosphere, unlike other plasma properties. The in-situ signatures of CME plasma, when compared to coronal structures observed via remote sensing, would detail the evolution of these structures in the solar wind and provide data toward answering this question.

Q1c: Is the source of the solar wind steady or intermittent? Solar wind is injected into the heliosphere with variations in its composition, temperature, and density. The VDFs and composition of the solar wind will provide the solar wind type, while magnetic field and electron data can provide information on the location of emitted plasma. When these measurements are used together, the sources of the different types of solar wind can be identified, and it can be determined whether they are steady or intermittent.

Q2a: How is energy from the lower solar atmosphere transferred to, and dissipated in, the corona? To fully characterize energy transfer and dissipation in the corona, it is useful to know which structures are major contributors to the dissipation (e.g., coronal holes, streamers, etc.). Low FIP elements are enhanced in the slow wind relative to their photospheric values; the effect is not as strong in the fast wind. This FIP effect is a key indicator of the acceleration and heating that occurred in the corona. Composition measurements of heavy ions in the acceleration region of the solar wind would provide data on the solar wind's heating and acceleration, and would distinguish between the mechanisms that create fast and slow wind.

Q2b: What coronal processes shape the non-equilibrium velocity distribution observed throughout the heliosphere? By measuring VDFs and the magnetic field over the radial distances covered by the SPP orbit, anisotropies in the plasma can be characterized. There are significant differences in the VDFs of heavy ions, protons, and helium in the different types of solar wind, however. To determine the extent to which turbulence and wave-particle interactions affect the VDFs, the distribution functions of not only H^{+}, but also of He^{++} and heavy ions need to be measured. Question 2b can be addressed more rigorously when heavy ions are included in the data.

Q2c: How do the processes in the corona affect the properties of the solar wind in the heliosphere? Because of its close proximity to the corona, it is likely that SPP will encounter light FIP elements (e.g., sodium), which have not yet been measured due to their scarcity. The low-FIP enhancement found in the slow solar wind can be explained by assuming ions spend time trapped in coronal loops, heating and rising in charge state before being released through the opening of the flux loop via magnetic reconnection (Fisk, Schwadron, and Zurbuchen 1998). Once the FIP effect is characterized for these additional elements, the details of the mechanisms that cause the FIP effect can be narrowed down.

Q3a: What are the roles of shocks, reconnection, waves, and turbulence in the acceleration of energetic particles? The composition and energy spectra of suprathermal tails and energetic

ions will be made by the SPP energetic particle instrument down to energies of 20 keV/nuc. A composition instrument that can measure the core and low-energy suprathermal tail of the velocity distribution will allow the energetic particle detector to map its findings to lower energies over a range of ion species.

Q3b: What are the source populations and physical conditions necessary for energetic particle acceleration? An important factor in the identification of SEP seed populations is their composition. Enrichments in ^{3}He or differences in isotopic composition are distinguishing signatures (Cohen et al. 1999). It is essential for SPP to have composition measurements of the solar wind to compare to that of the suprathermal tails measured by the energetic particle instruments. Heavy ion VDFs from the core out to the suprathermal tails, along with a careful measure of charge state ratios, will also contribute to the understanding of how CMEs accelerate ions. The importance of inner-source pickup ions to the seed populations cannot be understated (Gloeckler and Geiss 1998; Gloeckler et al. 1998; Chotoo et al. 2000), and composition measurements would allow an unprecedented radial profile of inner source pickup ions inward to 9.5 Rs.

Q3c: How are energetic particles transported radially across magnetic field lines from the corona to the heliosphere? While not essential for this particular question, a solar wind composition instrument can still be beneficial by providing charge state and elemental abundance ratios to characterize any CMEs that are present with the energetic particles.

SO4: Explore dusty plasma phenomena in the near-Sun environment and their influence on the solar wind and energetic particle formation. The sources of inner-source pickup ions have been postulated to include such things as dust grains that absorb and re-emit the solar wind (Schwadron and Gloeckler 2007). A composition instrument would measure pickup ions of various species in the solar wind and correlate with the measurements of dust. By this collaboration, clues to the precise nature and sources of inner-source pickup ions will be found.

IV. Summary

The benefits that composition measurements could have provided to the SPP mission have been briefly summarized. Measurements of VDFs and solar wind composition are beneficial for all of the SPP science questions; for some questions, they are essential to achieve science closure. An expert panel could quickly determine if the mission funded to address the Science Objectives of SPP can still achieve those objectives without the originally planned solar wind composition measurements. Instruments modeled after successful composition spectrometers (Gloeckler et al. 1998; Andrews et al. 2007) such as the Solar Wind Ion Composition Spectrometer (~8 kg, $15M-$20M) or the Fast Imaging Plasma Spectrometer (~2 kg, $5M-$10M) would be sufficient to make the necessary measurements of SPP, and each design has advantages over the other in terms of field of view or energy range. Carbon-foil based mass spectrometers have accumulated over 100 years of data collection in space, and the robustness of these instruments is well-documented (McComas et al. 2004). If the science return of a composition instrument is found to be fundamental to the success of SPP, then possible solutions need to be investigated soon.

References

2009 Heliophysics Roadmap Team. 2009. “Heliophysics: the solar and space physics of a new era : recommended roadmap for science and technology 2009-2030.”.

Andrews, G. B., Thomas H. Zurbuchen, B. H. Mauk, H. Malcom, Lennard A. Fisk, George Gloeckler, G. C. Ho, J. S. Kelley, Patrick L. Koehn, T. W. Lefevere, Stefano S. Livi, Robert A. Lundgren, and Jimmie M. Raines. 2007. “The Energetic Particle and Plasma Spectrometer Instrument on the MESSENGER Spacecraft.” *Space Science Reviews* 131:523-556.

Chotoo, K., Nathan A. Schwadron, G. M. Mason, Thomas H. Zurbuchen, George Gloeckler, A. Posner, Lennard A. Fisk, Antoinette B. Galvin, D. C. Hamilton, and M. R. Collier. 2000. “The suprathermal seed population for corotating interaction region ions at 1 AU deduced from composition and spectra of H+, He++, and He+ observed on Wind.” *Journal of Geophysical Research* 105:23107-23122.

Cohen, C. M. S., R. A. Mewaldt, R. A. Leske, A. C. Cummings, Edward C. Stone, Mark E. Wiedenbeck, Eric R. Christian, and Tycho T. von Rosenvinge. 1999. “New observations of heavy-ion-rich solar particle events from ACE.” *Geophysical Research Letters* 26:2697-2700.

Fisk, Lennard A., Nathan A. Schwadron, and Thomas H. Zurbuchen. 1998. “On the Slow Solar Wind.” *Space Science Reviews* 86:51-60.

Fisk, Lennard A. and Thomas H. Zurbuchen. 2006. “Distribution and properties of open magnetic flux outside of coronal holes.” *Journal of Geophysical Research* 111(A10):9115-+.

Geiss, Johannes, George Gloeckler, and Rudolf von Steiger. 1995. “Origin of the Solar Wind From Composition Data.” *Space Science Reviews* 72:49-60.

Gloeckler, George, J. Cain, Fred M. Ipavich, E. O. Tums, Peter Bedini, Lennard A. Fisk, Thomas H. Zurbuchen, Peter Bochsler, J. Fischer, Robert F. Wimmer-Schweingruber, Johannes Geiss, and Reinald Kallenbach. 1998. “Investigation of the composition of solar and interstellar matter using solar wind and pickup ion measurements with SWICS and SWIMS on the ACE spacecraft.” *Space Science Reviews* 86:497-539.

Gloeckler, George and Johannes Geiss. 1996. “Abundance of 3He in the local interstellar cloud.” *Nature* 381:210-212.

Gloeckler, George and Johannes Geiss. 1998. “Interstellar and Inner Source Pickup Ions Observed with SWICS on ULYSSES.” *Space Science Reviews* 86:127-159.

Gloeckler, George, Johannes Geiss, H. A. Balsiger, Lennard Fisk, A. B. Galvin, Fred M. Ipavich, Keith W. Ogilvie, Rudolf von Steiger, and B. Wilken. 1993. “Detection of Interstellar Pick-Up Hydrogen in the Solar System.” *Science* 261:70-73.

Lepri, Susan T. and Thomas H. Zurbuchen. 2004. “Iron charge state distributions as an indicator of hot ICMEs: Possible sources and temporal and spatial variations during solar maximum.” *Journal of Geophysical Research (Space Physics)* 109(A18):1112-+.

McComas, David J., Frederic Allegrini, Craig J. Pollock, Herbert O. Funsten, S. Ritzau, and George Gloeckler. 2004. “Ultrathin (~10 nm) carbon foils in space instrumentation.” *Review of Scientific Instruments* 75:4863-4870.

National Aeronautics and Space Administration. 2009. “Announcement of Opportunity: Solar Probe Plus Investigations.”.

National Research Council. 2003. “The Sun to the Earth -- and Beyond: A Decadal Research Strategy in Solar and Space Physics.”.

Neugebauer, Marsha, Paulett C. Liewer, Edward J. Smith, Ruth M. Skoug, and Thomas H. Zurbuchen. 2002. “Sources of the solar wind at solar activity maximum.” *Journal of Geophysical Research* 107:1488-+.

Schwadron, Nathan A. and George Gloeckler. 2007. “Pickup Ions and Cosmic Rays from Dust in the Heliosphere.” *Space Science Reviews* 130:283-291.

Solar Probe Science and Technology Definition Team. 2005. “Solar Probe: Report of the Science and Technology Definition Team.”.

Solar Probe Science and Technology Definition Team. 2008. “Solar Probe Plus: Report of the Science and Technology Definition Team.”.

Wang, Y.-M., Y.-K. Ko, and R. Grappin. 2009. “Slow Solar Wind from Open Regions with Strong Low-Coronal Heating.” *The Astrophysical Journal* 691(1):760-769.

Zhao, Liang, Thomas H. Zurbuchen, and Lennard A. Fisk. 2009. “Global distribution of the solar wind during solar cycle 23: ACE observations.” *Geophysical Research Letters* 36:14104-+.

RAM: The Reconnection and Microscale Mission

Understanding Ideal and Resistive Processes in the Dynamic Corona

Authors: L. Golub, J. Bookbinder, E. DeLuca, SAO; S. Christe, GSFC;
J. Cirtain, MSFC; K. Kobayashi, UA-Huntsville; C. Korendyke, NRL

This paper outlines a new approach to understanding the dynamic activity of hot, magnetized astrophysical plasma by studying the best example available to us, the solar corona. As a result of recent progress, both in theory and in observations such as those from the NASA TRACE Small Explorer satellite, we now understand that it is the skeletal structure of the space occupied by the magnetized coronal plasma – the way in which the regions of different connectivity divide up the space - that determines the nature and location of the observed dynamics.

One of the key factors in improving our understanding of coronal physics has been the continued development of high resolution imaging, particularly in the X-ray and EUV spectral regime. Beginning in the 1960s, direct on-disk imaging of the X-ray corona became available, leading to a clear understanding of the close connection between surface magnetic fields and X-ray emission (Vaiana et al 1973). The improved spatial resolution of the Skylab instruments (Zombeck et al 1978) led to the view that the corona is composed of loop-like structures representing isolated mini-atmospheres (Rosner et al 1978), with thermodynamic-type scaling laws describing their physical properties.

Further improvements in spatial resolution with the Yohkoh SXT led to an acceptance of the role of magnetic reconnection in coronal dynamics (Tsuneta 1996) and the additional improvements in both spatial and temporal resolution from TRACE (viz. Figure 1) have led to a far more elaborate model of coronal configuration and dynamics set in three-dimensional topologies defined by null points, separators and bald patches among other things (Démoulin 2005). Moreover, recent work on determining the true coronal density in the observed structures (Dere 2008, 2009; Warren etal. 2009; Tripathi et al. 2009) indicates that the coronal filling factor is in the range of 1-10%. Assuming linear structures, this means that a 3-10X improvement beyond our current best spatial resolution of 0.5 arcsec pixel size will finally be enough to resolve the basic structures in the corona.

At the heart of this new understanding is the concept of a Quasi-Separatrix Layer. QSLs are 3D geometrical objects that define narrow volumes across which magnetic field lines have strong, but finite, gradients of connectivity from one footpoint to another. QSLs extend the concept of separatrices, topological objects across which the magnetic connectivity is discontinuous. Based on analytical arguments, and on magnetic field extrapolations of the Sun's coronal force-free field above observed active regions, it has long been conjectured that QSLs are favorable locations for current sheet formation, as well as for magnetic reconnection, and therefore are good predictors for the locations of magnetic energy release in flares and coronal heating.

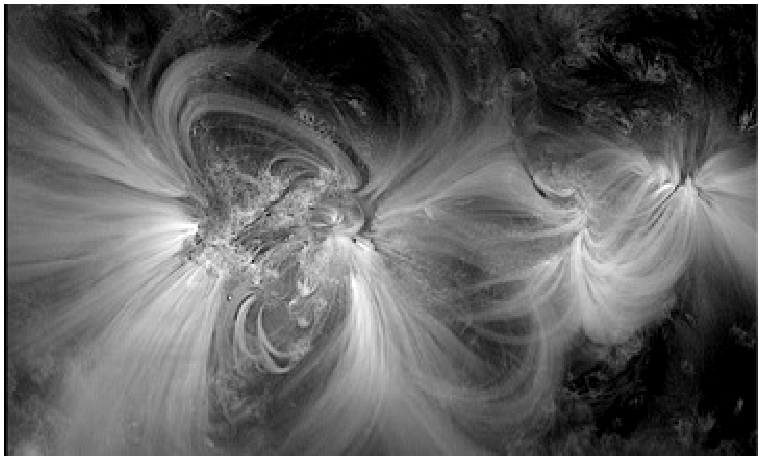


Figure 1 Current EUV imagers see coronal fine structure at their resolution limit. Recent studies indicate that the proposed RAM spatial resolution will finally resolve these transverse scale sizes.

Recent numerical MHD simulations, direct solar observations, and laboratory experiments, support the conjecture that QSLs play an important role (e.g., Titov 2007, Chandra *et al.* 2009, Rappazzo *et al.* 2008, Schrijver *et al.* 2010). Together, they suggest that QSL reconnection is involved in the EUV and SXR brightenings seen in chromospheric flare ribbons, that it is implicated in the heating of EUV coronal loops, and that the dissipation of QSL-related current sheets may be the cause of coronal heating in initially homogeneous, braided and turbulent flux tubes, as well as in

coronal arcades rooted in small-scale photospheric magnetic flux concentrations. The apparent ubiquity of QSL-related current sheets in the Sun's corona also suggests that QSLs play an important role in the dynamics of stellar atmospheres and in hot, magnetized plasmas throughout all of astrophysics.

Our understanding of coronal dynamics will be further advanced by understanding how localized reconnection fits into the three-dimensional topology of the field. For this to be accomplished, we must compare coronal structures to the relevant calculated topological features which are predicted to play a role in coronal activity. We must then be certain that we are observing the elementary structures in the corona which participate in the interactions between the plasma and the magnetic field. Analysis of the TRACE and AIA data shows that the scale size of coronal structures extends down to the resolution limit of the instrument (Golub et al 1998), and there is substantial evidence that additional fine structure is present. This discussion therefore takes the view that the first step in understanding the corona is to see what is there and then to analyze the physical processes occurring in the observed structures.

Scientific objectives. The main objectives of RAM are to determine the geometric configuration and topology of the structures making up the inner corona and to establish the relationship of those structures to coronal dynamics. The mission is designed to study the effects of magnetic field emergence, diffusion and reconnection, and to help understand the coupling of small-scale dynamic and eruptive processes to the large-scale structure and dynamics. The specific detailed objectives are:

- Determine the coronal configurations that lead to energy release.
- Identify energy release locations, their relation to surrounding plasma structures and their physical properties (such as size, geometry, structural and topological changes).
- Confront theory (e.g., presence and effects of QSLs, breakout configurations, Alfvén and magnetosonic waves) with the relevant observations.
- Locate possible sources of particle production within dynamic coronal events and determine the conditions leading to such events.

Observables to be measured. The scientific objectives of the RAM mission will be met by obtaining three types of observational data:

- High resolution imaging of the coronal structures involved in reconnection events.
- Soft X-ray or EUV spectroscopy of the reconnection and energy release processes.
- Hard X-ray imaging to determine the location and nature of high energy particle acceleration.

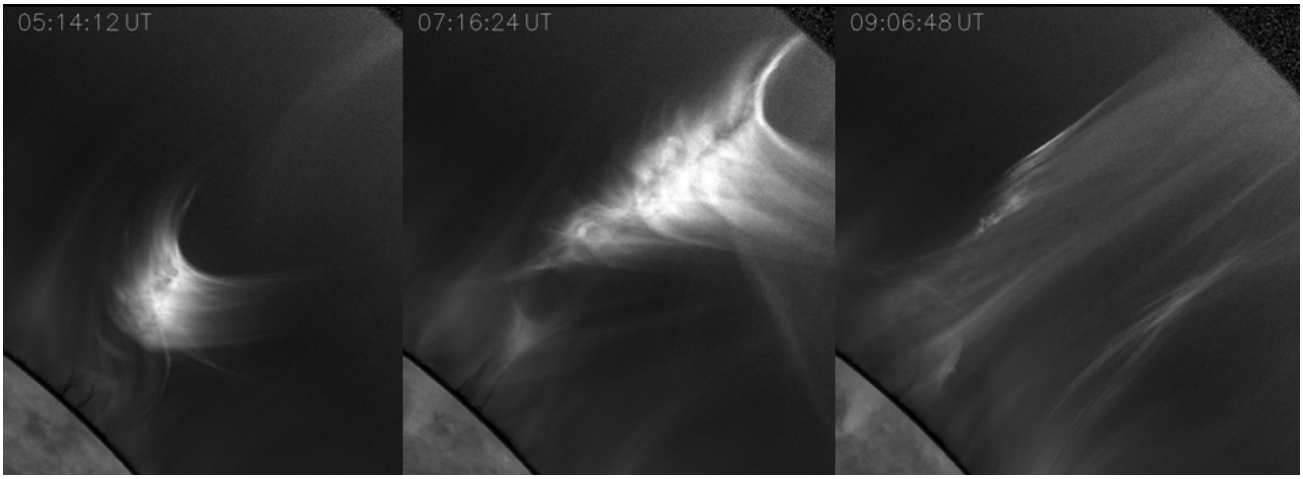


Figure 2 An erupting CME viewed in the EUV by SDO/AIA; the solar disk is seen at lower left in these images. The movie clearly shows a pre-existing flux tube with coronal material residing at the bottom, which slowly lifts upward while magnetic field lines underneath reconnect at a thin, dark vertical region. The effects of the eruption are seen at locations on the solar surface distant from the reconnection site, apparently related to downflows at the ends of the flux rope.

In the following we will describe in some detail the coronal imager and the X-ray spectrometer option. The other instruments under consideration are described in their respective white papers, which we reference but whose detailed discussion we will not repeat here. A strawman payload for the RAM mission, containing three instruments for obtaining the necessary data, would consist of:

A high-resolution coronal imager, similar to the Hi-C (shown below) sounding rocket telescope, with resolution designed to meet the goals of the RAM mission. The telescope, operating at 193Å (Fe XII/XXIV) has a 50 cm diameter primary, a single-extension deployable tube for 7.5m total length, 10X secondary for an efl of 75m. The single-element extendable bench used for Hi-C will also be available to the spectrometer and the HXR imager. With a 12μm CCD pixel this yields 1/30 arcsec pixel resolution, more than enough to resolve clearly the corona substructures, as discussed above. A minimum focal plane array of 4Kx4K is needed for encompassing target regions of interest. The estimated instrument mass is 50 kg, power 50W, and telemetry 3Mbps avg. rate to the CPU.

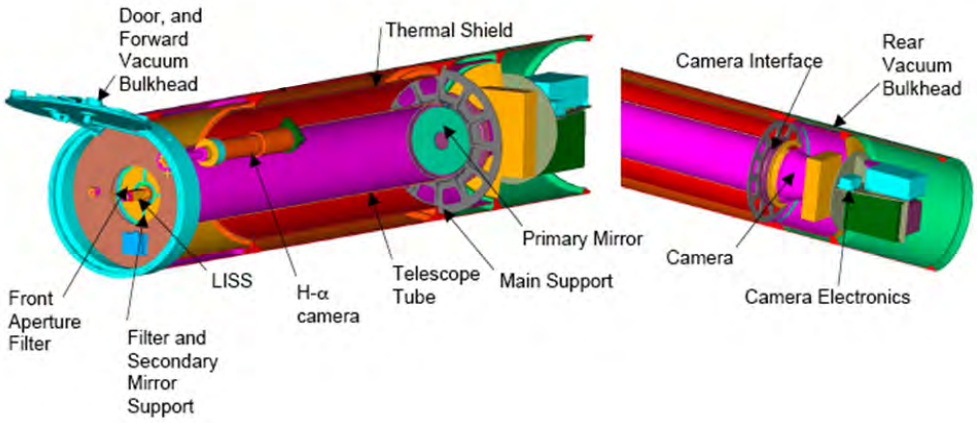


Figure 3 The Hi-C sounding rocket payload. The RAM telescope uses a single-extension optical bench to increase the instrument length from the 4m launch length to 7.5m on-orbit deployed length. All of the co-observing instruments on RAM will be able to utilize this total length in their designs.

A UV, EUV or Soft X-ray (SXR) spectrometer. The purpose of this instrument is to provide diagnostics of the physical conditions in the target regions observed by the high resolution coronal imager and the hard X-ray imager. The spectrometer should, at a minimum, provide temperature, density and Doppler shift (bulk flow velocity) information for both the large-scale structures surrounding the energy release sites, and for the expected energy release locations with their associated high temperature emission, as well as locations of moving shock fronts. The required temperature coverage is 1 MK to 15 MK, spatial resolution of 1 arcsec or better, and temporal resolution is 1 second or better. There are several candidate instruments for providing the necessary spectroscopy. We will list three possibilities, noting that there may be others.

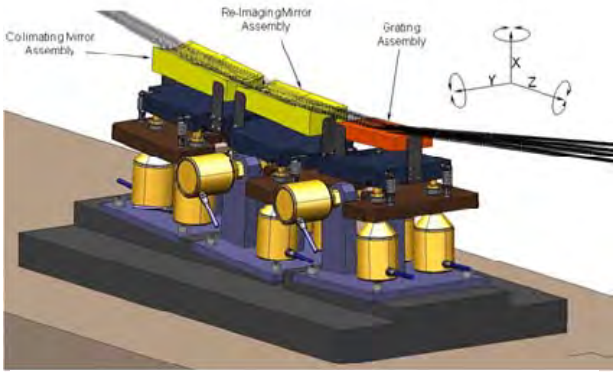
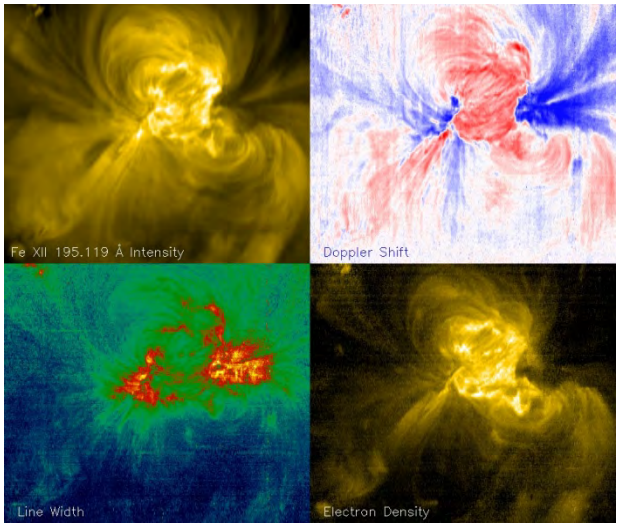


Figure 4 Prototype test assembly for the MaGIXS spectrometer, showing the key components of the design.

MaGIXS (Marshall Grazing Incidence X-ray Imaging Spectrometer): A novel design for an imaging spectrograph operating at soft x-ray wavelengths has been developed at MSFC, and a prototype is under construction jointly with SAO. The key feature of this instrument, illustrated in Figure 4, is a pair of collimating and refocusing mirror behind the slit which correct the major aberrations to allow for high quality imaging over a large (10') field of view. Anticipated parameters for the RAM version are: 7m L, 20cm diam., 0.5 arcsec pixels, spectra 6-24A, full Sun X-ray context imaging 2-60A, 35kg, 40W, 2Mbps.

XRC: The X-Ray Calorimeter is described in the white paper by Bandler etal. The key features are: 1000-element focal plane array, soft x-ray spectroscopy with spectral resolution of a few eV, photon counting at very high rates and consequent ability to carry out high temporal and spectral resolution flare and active region diagnostics, including absolute densities, flows and non-equilibrium temperatures.



FACTS (Fine scale Advanced Coronal and Transition Region Spectrometer): see the separate white paper from Korendyke for detailed discussion. We assume a baseline version of FACTS here, with 0.1" pixels, 170-210Å passband, 30cm diam. mirror, 4Kx2K CCD, 7m length.

Figure 5. Physical parameters of an active region determined by imaging and spectroscopy combined. The UL image is Fe XII, the UR shows flows measured via Doppler shifts. The LL shows line broadening with turbulent a component and the LR shows electron density in the active region.

Hard X-ray Imager. We baseline the FOXSI (Focusing Optics X-ray Solar Imager; Krucker et al, 2009). A hard x-ray imaging spectrometer for locating and analyzing the non-thermal particle properties of solar events. Operates at 4-80 keV, 8 arcsec resolution (FWHM), 7m length, 70kg mass, 60W power, 5 Mbps telemetry. See the Christe white paper for additional details and status.

Table 1 RAM Mission Description

Primary Elements	Mass (kg)	Datarate (Mbps)
Hi-res (0.03") Imager	**80**	**10**
XRC	**300**	**25**
0.1" UV/EUV Spectrometer	**150**	**20**
FOXSI	**100**	**5**
Spacecraft (with EOB)	**325**	
Reserves	**170-230**	**15**
Total (includes 30% reserves)	**850-1110**	**50**
LV Capability to GEO (Falcon 9)	**2000**	**165**
Margins (% of total)	**45-55%**	**70%**

RAM Mission ROM Cost Estimate		
Instrument	**Option**	**ROM Cost**
High-res coronal imager	Hi-C	$75M
SXR/EUV Spectrometer (options shown)	MaGIXS	$75M
	EUV-HI	$85M
	X-ray Calorimeter	$95M
Hard x-ray imager	FOXSI	$50M
Spacecraft	TBD	$125M
30% contingency		$100M
Launch	TBD	$180M
MO&DA	Polar, Geosynch	$20-50M
Total		**$605-655M**

NASA Research Focus Areas. The RAM mission directly addresses a number of NASA strategic goals:

- *F1: Magnetic Reconnection.* Locate reconnection regions and determine their boundary properties (geometry/topology, inflow/outflow speeds, and timing within overall event).
- *F2: Particle Acceleration and Transport.* Find likely locations of particle production, determine their relation to the large-scale structure of the event.
- *H3: Role of the Sun in driving changes in the Earth's atmosphere.* Determine physical and magnetic properties of CMEs and other eruptions and flows leaving the corona.
- *J3: Capability to predict the propagation and evolution of solar disturbances.* Determine the accuracy of theories and models of solar eruptive phenomena.

References:

Antiochos, S. K., 1998, ApJ, 502, L181.
Aulanier, G., Pariat, E., Démoulin, P. 2005, A&A, 444, 961
Aulanier, G., Pariat, E., Démoulin, P., DeVore, C.R. 2006, Sol. Phys., 238, 347
Aulanier, G., Golub, L., DeLuca, E.E, Cirtain, J.W., Kano, R., et al. 2007, Science, 318, 1588
Aulanier, G., Tor¨ok, T., D´emoulin, P., DeLuca E.E. 2010, ApJ, 708, 314
Baker, D., van Driel-Gesztelyi, L., Mandrini, C.H., et al. 2009, ApJ, 705, 926
Beveridge, C., Longcope, D. L., Priest, E. R., 2003, Sol. Phys., 216, 27.
Bogachev, S.A., Somov, B.V., Kosugi, T. Sakao, T. 2005, ApJ, 630, 561
Cargill, P. & Klimchuk, J., ApJ, v605, 2004.
Carroll, T.A. 2011, Proceedings of IAU Symposium 273.
Chandra, R., Schmieder, B., Aulanier, G., Malherbe, J.-M. 2009, Sol. Phys., 258, 53
Démoulin, P., Hénoux, J.-C., Priest, E.R., Mandrini, C.H. 1996a A&A, 308, 643
Démoulin, P., Priest, E.R., Lonie, D.P. 1996b JGR, 101, A4, 7631
Démoulin, P., Bagala, L.G., Mandrini, C.H., Hénoux, J.C., Rovira, M.G. 1997 A&A, 325, 305
Démoulin, P. 2006 Adv. Sp. Res., 37, 7, 1269
Dere, K. P., Landi, E., Mason, H. E., Monsignori Fossi, B. C., Young, P. R., 1997, A&A, 125, 149.
Fan, Y. 2010, ApJ, 719, 728
Fletcher, L., López Fuentes, M.C., Mandrini, C.H., et al. 2001, Sol. Phys., 203, 255
Fletcher, L., Hudson, H.S 2002, Sol. Phys., 210, 307
Galsgaard, K., Titov, V.S., Neukirch, T. 2003, ApJ, 595, 506
Galsgaard, K. 2011, Proceedings of the IAU Symposium 273.
Gómez, D.O., DeLuca, E.E., McClymont, A.N. 1995, ApJ, 448, 954
Gudiksen, B.V., Nordlund, A. 2005, ApJ, 618, 1031
Hendrix, D.L., van Hoven, G. 1996, ApJ, 467, 887
Hesse, M., Schindler, K. 1988 JGR, 93, 5559
Krucker, S., et al. 2009, Proc. SPIE, Vol. 7437, 743705.
Kobayashi, K., et al. 2010, Proc. SPIE, Vol. 7732, 773233.
Jardine, M., Donati, J.-F.et al. 2011 Proceedings of the IAU Symposium 273
Lawrence, E.E, Gekelman, W. 2009, PRL, 103, 10, 105002
Mandrini, C.H., Démoulin, P., Klimchuk, J.A. 2000 ApJ, 530, 999
Masson, S., Pariat, E., Aulanier, G., Schrijver, C.J. 2009, ApJ, 700, 559
Masson, S., Aulanier, G., Pariat, E., Klein, K.L. 2010, ApJ (Letters), in press
Mikic, Z., Schnack, D.D., van Hoven, G., 1989, ApJ, 338, 1148
Milano, L.J., Dmitruk, P., Mandrini C.H., G´omez, D.O., Dmoulin, P. 1999, ApJ, 521, 889
Priest, E.R., Forbes, T.G. 1992, JGR, 97, A12, 1521
Priest, E.R., D´emoulin, P. 1995, JGR, 100, A12, 23443
Priest, E.R., D´emoulin, P. 1997, Sol. Phys., 175, 123
Rappazzo, A.F., Velli, M., Einaudi, G., Dahlburg, R.B. 2008 ApJ, 677, 1348
Restante, A.L., Aulanier, G., Parnell, C.E. 2009, A&A, 508, 433
Schmieder, B., Aulanier, G., Démoulin, P., van Driel-Gesztelyi, L., et al. 1997, A&A, 325, 1213
Schrijver, C.J., DeRosa, M.L., Title, A.M. 2010, ApJ, 719, 1083

Titov, V.S. 2007, ApJ, 660, 863
Titov, V.S., Hornig, G., D´emoulin, P. 2002, JGR, 107, A8, SSH 3-1
Vaiana, G. S., et al. , 1973, 185, L47.
van Ballegooijen, A.A. 1986, ApJ, 311, 1001
Wiegelmann, T. 2011, Proceedings of the IAU Symposium 273
Wilmot-Smith, A.L., Hornig, G., Pontin, D.I. 2009, ApJ, 696, 1339
Wilmot-Smith, A.L., Pontin, D.I., Hornig, G. 2010, A&A, 516, 5
Van Ballegooijen, A. A., 1985, ApJ, 298, 421.
Young, P. R., Del Zanna, G., Landi, E., Dere, K. P., Mason, H. E., Landini, M., 2003, ApJS, 144, 135.
Zombeck, M. V., Vaiana, G. S., Haggerty, R., Kreiger, A. S., Silk, J. K., Timothy, A., ApJS, 38, 69.

The 1.6 m Clear Aperture Optical Solar Telescope in Big Bear – The NST

Summary

The first facility-class solar observatory in the US in more than a generation is presently in operation at Big Bear Solar Observatory (BBSO). The telescope is producing diffraction limited images with its 1.6 m clear aperture. This largest aperture solar telescope will provide an essential complement to SDO, Hinode and other satellite data, especially as a probe of the origins of space weather.

Center for Solar-Terrestrial Research (CSTR) -- New Jersey Institute of Technology (NJIT)

Big Bear Solar Observatory (BBSO) of New Jersey Institute of Technology (NJIT) is located high in the San Bernardino Mountains of Southern California. BBSO operates a 1.6 m off-axis solar telescope (the NST) that is world's largest aperture solar telescope. BBSO is collaborating with NSO in building multi-conjugate adaptive optics hardware and a fully cryogenic near infrared spectrograph for the NST. Both of these instruments will be unique. The NST is also the ideal solar telescope to provide the NASA solar community leading-edge observations to complement NASA observations of the Sun, especially for the low solar atmosphere to provide the most precise, observational lower boundary condition for all studies of the heliosphere and most especially the origin of space weather.

The essential improvement is resolution that is needed in probing the origins of solar dynamics is already available from the NST. In Fig. 1, the NST clearly reveals individual bright points that are the fundament scale of solar magnetism. The small scale turns out to be much more dynamic than imagined.

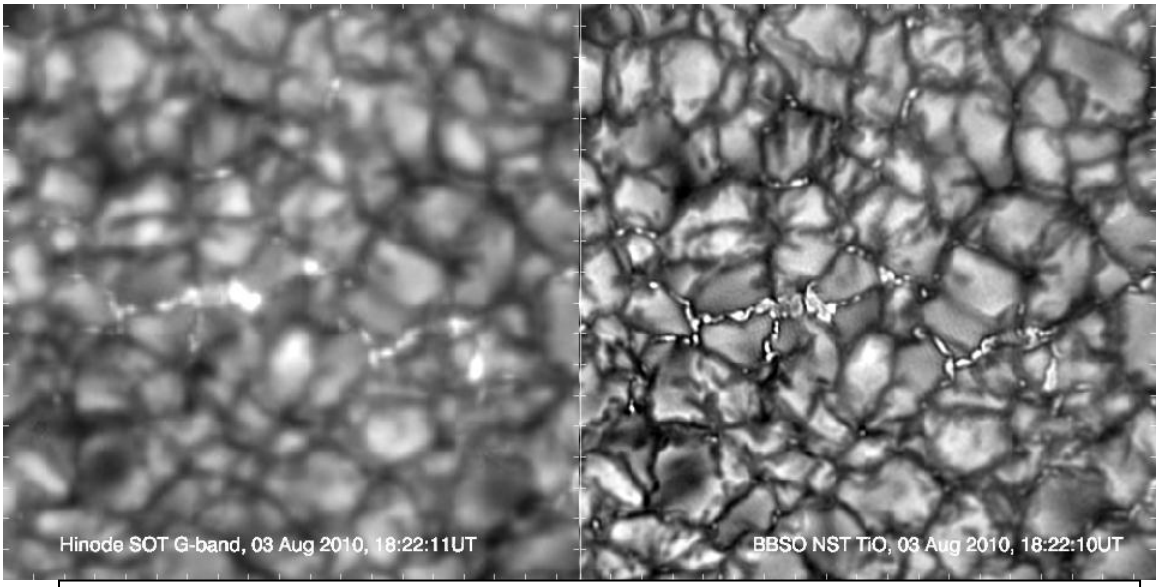


Fig. 1 Hinode (right) and NST (left) intensity images at the same time and place.

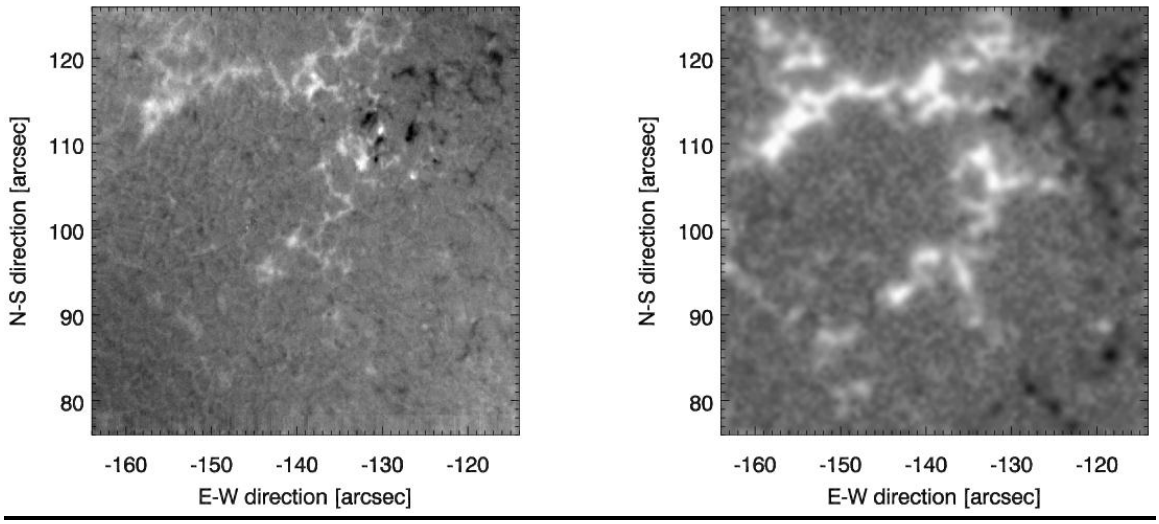


Fig. 2. NST (left) and SDO (right) line-of-sight magnetograms at the same place and a few minutes apart.

Fig.2 shows how sharp NST magnetograms can be. In particular they show small magnetic sources that are otherwise invisible but also highly dynamic. Such information is critical in space weather studies and an essential complement to NASA missions to probe the heliosphere.

The primary mission of the CSTR is to observe and model the solar system environment from the Sun to the Earth. To accomplish its mission, CSTR operates several significant and unique ground-based facilities, including Big Bear Solar Observatory (BBSO), the Owens Valley Solar Array (OVSA) and Automated Geophysical Observatories (AGO) across Antarctica. As well, CSTR is involved as co-investigators or principal investigators in other ground-based and space projects. NJIT is designated as New Jersey's Science and Technology University. CSTR's recent postdoctoral associates and graduate students have gone on to make their marks in universities, NASA, national centers and industry. Our successes have made NJIT one of the most significant universities in the U.S. in solar astrophysics and space sciences.

Fig.3. The NST showing its active mirror cell



BBSO is located in Big Bear Lake, California in the San Bernardino Mountains of Southern California. The mountain lake is characterized by sustained atmospheric stability, which is essential for BBSO's primary interests of measuring and understanding solar complex phenomena utilizing our dedicated telescopes and instruments. BBSO new solar telescope is a 1.6 m clear aperture, off-axis telescope, the NST. Roger Angel has pointed out that the NST is also a pathfinder for a new generation of nighttime telescopes like the Giant Magellan Telescope with huge off-axis mirrors

The broad scientific goals of the NST are to (i) make high cadence, high resolution studies of solar flares, (ii) measure the structure and evolution of magnetic fields and flow fields in flaring active regions and (iii) observe the dynamics of kilogauss "flux-tubes". To accomplish these goals, sophisticated post-focus instrumentation is required.

Early in the three year commissioning phase, Adaptive Optics (AO) and the near-infrared (IRIM) polarimeter were brought on line. TiO and Halpha images from the NST are shown in Figure 4.

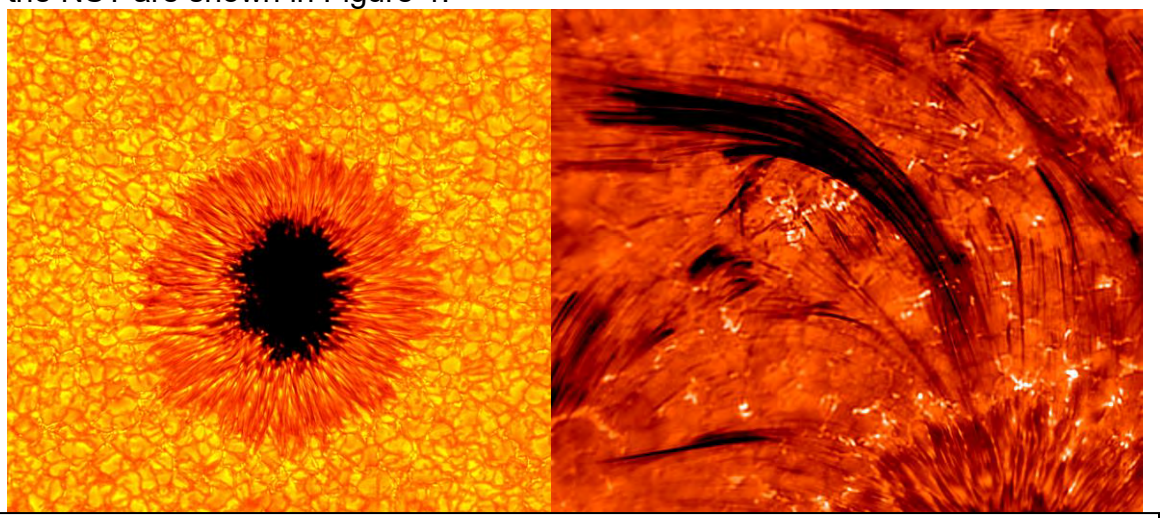


Fig. 4.An active region (NOAA 1084) taken on July 2, 2010 in TiO (706 nm) using the1.6 m aperture New Solar Telescope in Big Bear Solar Observatory with realtime correction for atmospheric distortion (adaptive optics). For perspective, the Earth is slightly smaller than the whole sunspot including the dark umbra and the daisy petal-like penumbra. The spot is surrounded by the Sun's ubiquitous granular field in which the small individual bright points in intergranular lanes are near the diffraction limit of the telescope.
Halpha off-band image of the same active region taken a day earlier reveals dynamics in the chromospheric layer overlying the surface region shown in the full sunspot image. Most apparent are long dark streaks across the image that are ~50km/s "jets" arising from the bright magnetic areas immediately outside the main body of the penumbra.

NJIT was founded in 1882 and is currently the Science and Technology University of New Jersey. It is one of two state universities of New Jersey. NJIT's efforts in astronomy and astrophysics focused in CSTR directed by Phil Goode. The Center was founded in 1997 and has graduated more than twenty PhD's since its founding. As well, more than ten of our former postdoctoral associates have moved on to tenure/track positions at universities or national centers.

Earth-Affecting Solar Causes Observatory (EASCO): A New View from Sun-Earth L5

N. Gopalswamy, J. M. Davila, O. C. St. Cyr, T. Duvall, E. C. Sittler, R. J. MacDowall, A. Szabo, and M. R. Collier (NASA/GSFC), J. Johnston, D. F. Webb, E. W. Cliver, S. W. Kahler (AFRL), F. Auchère (Orsay/France), M. Maksimovic, J.-L. Bougeret (Meudon/France), B. Heber (Kiel/Germany), E. Valtonen (Turku/Finland), S. Vennerstrom (Copenhagen/Denmark), C. de Koning, D. Biesecker, V. Pizzo (NOAA/SWPC), R. L. Moore, A. Sterling (NASA/MSFC), A. Vourlidas, R. A. Howard (NRL), J. Zhang (GMU), B. Vrsnak (Zagreb), S.P. Rajaguru (IIA/Bangalore)

Abstract: This white paper outlines the concept of a mission that will make remote-sensing and in-situ measurements from the Sun – Earth Lagrange point L5 to understand the origin and evolution of large-scale solar disturbances such as coronal mass ejections (CMEs) and corotating interaction regions (CIRs). The L5 mission is named the Earth-Affecting Solar Causes Observatory (EASCO), which will make measurements of solar variability from the solar interior, atmosphere, and the interplanetary (IP) medium. The L5 vantage point provides an unprecedented view of the solar disturbances and their solar sources that can greatly advance the science behind space weather. A coronagraph/heliospheric imager combination at L5 will be able to view CMEs broadsided, so space speed of the Earth-directed CMEs can be measured accurately and their radial structure discerned. In addition, an inner coronal imager and a magnetograph from L5 can give advance information on active regions and coronal holes that will soon rotate on to the solar disk. Radio remote sensing at low frequencies can provide information on shock-driving CMEs, the most dangerous of all CMEs. Coordinated helioseismic measurements from the Sun-Earth line and L5 provide information on the physical conditions at the base of the convection zone, where solar magnetism originates. Finally, the in situ measurements at L5 can provide information on CIRs heading towards Earth that potentially result in adverse space weather. Observing energetic particles from L5 and combining them with observations along the Sun-Earth line will help disentangle contributions from flares and CMEs.

EASCO will build upon the advances made during the last decade from heliospheric missions such as the Solar and Heliospheric Observatory (SOHO), Wind, and the Advanced Composition Explorer (ACE) that observe the Sun from the L1point.

1. Introduction

This white paper outlines the concept of a mission that will make remote-sensing and in-situ measurements from the Sun – Earth Lagrange point L5 to understand the origin and evolution of large-scale solar disturbances that affect Earth's space environment. Coronal mass ejections (CMEs) and corotating interaction regions (CIRs) are the large-scale disturbances that concern us, which can be best observed from the L5 vantage. We call the proposed mission as the Earth-Affecting Solar Causes Observatory (EASCO). The mission will contain instruments that will probe the solar interior where the magnetic field originates, to the solar atmosphere where the magnetic energy is released to the interplanetary medium into which the disturbances propagate.

More details on the L5 vantage point can be found in Akioka et al. (2005), Webb et al. (2010) and Gopalswamy et al. (2010).

A wealth of knowledge on CMEs and CIRs has been accumulated over the past three decades. However, many fundamental questions remain unanswered: How are CMEs initiated in closed magnetic regions? How do CMEs accelerate particles, alone and in combination with flare reconnection? Where and when do shocks form in the corona and how do they evolve? What is the radial profile of shock-driving CMEs? What is the internal magnetic structure of CMEs that cause magnetic storms? How do CMEs and CIRs evolve in the inner heliosphere? What are the changes that occur in the convection zone before active regions erupt? Answering these questions requires making accurate measurements of physical parameters from the solar interior to the atmosphere and into the heliosphere using a variety of techniques. The measurements include the magnetic and plasma properties of the source regions (active regions and filament regions for CMEs; coronal holes for CIRs) as well as the disturbances themselves. The observations involve multi-wavelength imagery and time series.

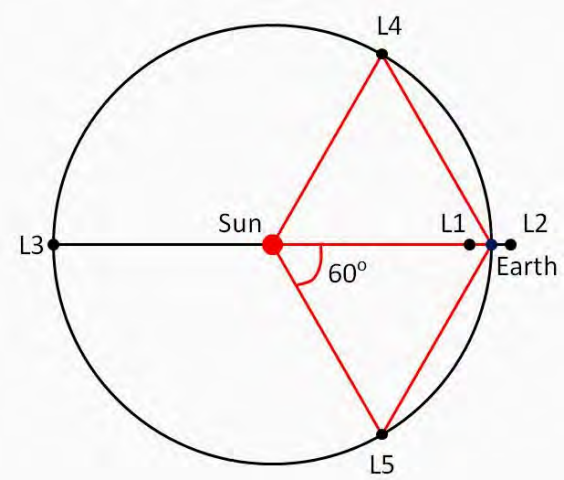


Figure 1. A sketch showing the five Sun – Earth Lagrange points, L1 – L5 with respect to Earth's orbit around the Sun. All the Lagrange points view the Sun continuously save L2, which is located on the night side of Earth. L5 is located at 60° away from the Sun-Earth line, trailing Earth. There have been many spacecraft located at L1. The Behind spacecraft of the STEREO mission went past the L5 point recently (end of 2009).

Coronagraphs have an occulting disk to block the direct photospheric light so the faint coronal structures can be imaged. This is a problem for observing CMEs originating on the solar disk along the Sun-Earth line: the occulting disk obscures the nose of the CME and CME-driven shock, where the shock is strongest and likely to accelerate particles. CME measurements are subject to severe projection effects, which prevent us from obtaining accurate Sun-Earth travel times. The case of CIRs is worse because we can observe them only when they are about to hit Earth (except for the white light observations using Heliospheric Imagers and SMEI). These difficulties can be readily overcome by EASCO because the L5 vantage point provides the full view of the CMEs when they are still close to the Sun and the CIRs well before they arrive at Earth (see Fig. 1).

2. Earth-affecting Solar Disturbances

CMEs from closed magnetic regions and CIRs from coronal holes are the two large-scale disturbances that seriously impact Earth's space environment. These large-scale disturbances are characterized by enhanced density and magnetic field with respect to the ambient solar wind. The enhanced magnetic field, if it has a southward component, interacts with Earth's magnetosphere producing geomagnetic storms. Both CIRs and CMEs drive shocks, but at different heliocentric distances: CME-driven shocks are known to form very close to the Sun

(~300,000 km above the solar surface), while the CIR shocks form only occasionally within 1 AU, but not within 0.3 AU. The CME-driven shocks accelerate particles from the corona and IP medium to very high energies. These solar energetic particles (SEPs) present a huge radiation hazard to humans and their technology in space in addition to affecting the conductivity of the ionosphere and ozone level in the atmosphere.

2.1 CMEs

CMEs responsible for SEPs at Earth and geomagnetic storms have different paths in the heliosphere (see Fig. 2). To drive a shock the CME speed V_{CME} above the solar wind speed V_{SW} must exceed the magnetosonic speed V_{MS} of the ambient medium. SEP production typically starts when the CME-driven shock is within a few solar radii from the Sun. The shock continues to produce SEPs in the IP medium so long as it remains strong enough. When the shock arrives at the observing spacecraft, a sudden increase in SEP intensity is observed (the so-called energetic storm particle event). The shock-driving ability of a CME is crucial for the production of SEPs and the shocks need to be strong enough to efficiently accelerate SEPs. Apart from the CME speed, the shock strength depends on the magnetosonic speed (or Alfven speed) in the ambient medium. Geomagnetic storms, on the other hand, occur only upon the arrival of the CME and/or its sheath at Earth's magnetosphere, provided these structures contain southward magnetic field components. The shock is also important for geoeffective CMEs in compressing the magnetosphere, but the intensity and duration of the southward component of the magnetic field in the shock sheath and CME determine the strength of the resulting magnetic storm. Thus the magnetic structure of the CME is important for geomagnetic storms, but not for SEPs.

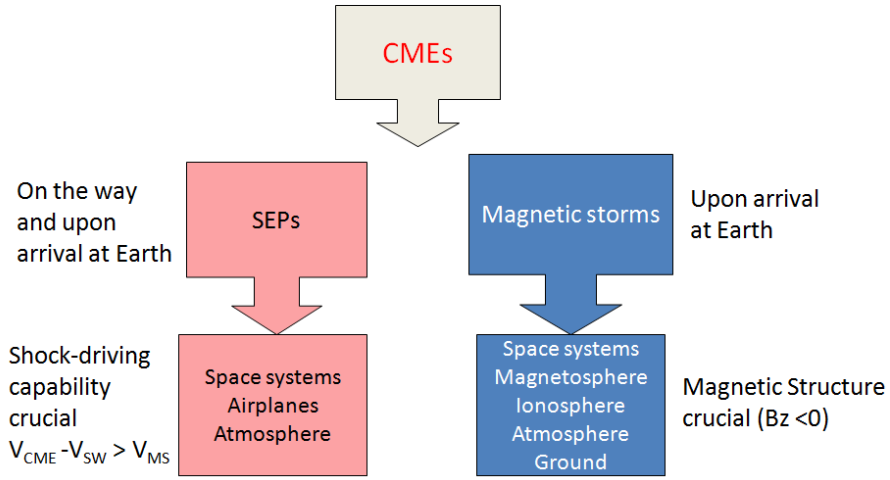


Figure 2. CME-driven shocks produce large SEP events and intense geomagnetic storms. CIRs also drive shocks, accelerate particles and produce geomagnetic storms. Only the magnetic storms are significant for CIRs because the CIR shocks accelerate particles typically at several AU.

The importance of an L5 view can be appreciated from the solar source locations of geoeffective and SEP-producing CMEs during solar cycle 23 (see Fig. 3). The source locations of geoeffective CMEs tend to cluster near the central meridian because only these CMEs head directly to Earth. There is a slight western bias to these source locations (average around W15) because of the eastward deflection due to the solar rotation. W15 in Earth view corresponds to W75 from L5. From Earth view, CMEs from W15 generally appear as halo CMEs, so it is difficult to measure their speeds accurately from a coronagraph viewing along the Sun-Earth line. On the other hand, these CMEs are limb events for an L5 view and hence the sky-plane speed from an L5 view is close to the true speed. The average direction of geoeffective CMEs is

marked as “GEO” in Fig. 3. One other difficulty with the halo CMEs from an Earth view has been deciding whether a CME is front sided or back-sided. A front-sided CME with weak or no obvious near-surface signatures may be mistakenly classified as a backside event. An L5 view can resolve this ambiguity, so one can readily say whether it is headed towards Earth or away from it. Resolution of such ambiguity is very important for space weather applications.

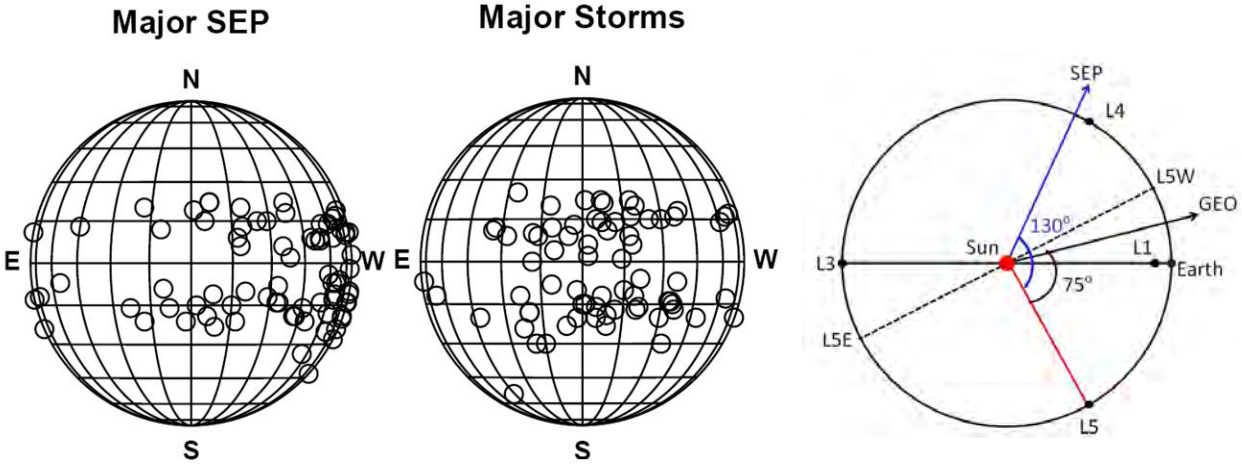


Figure 3. Heliographic coordinates of the solar sources of CMEs that produces major SEPs (left) and geomagnetic storms (middle). GEO and SEP denote the average directions of geoeffective and SEP-producing CMEs (right). The east (L5E) and west (L5W) limbs from L5 view are also shown.

CMEs associated with large SEP events typically originate from E45° to W90° (some even beyond W90). For fast (speed > 900 km/s) and wide (width > 60°) CMEs, the SEP association rate peaks for CMEs originating from W40° – W50° in an Earth view (Gopalswamy et al., 2008). If we consider CMEs accompanied by type II radio bursts in the decameter-hectometric (DH) wavelengths, then the SEP rate peaks for CMEs originating from ~W70. This direction is marked as “SEP” in Fig. 3. The direction “SEP” in Fig. 3 corresponds to about 40° behind the west limb for an L5 view and is at the limit of CME measurement without projection effects. The CMEs can also be observed with minimal “snow-storm” effects [coronagraph detector saturated by SEP events] because SEP events well connected to Earth are poorly connected to L5.

Simultaneous white-light and low-frequency radio observations will help reveal the shock signature ahead of the flux rope for geoeffective and SEP-producing CMEs. The radial structure of CMEs with shocks recognized in coronagraphic images (Gopalswamy, 2009; Gopalswamy et al., 2009a; Ontiveros and Vourlidas, 2009) and confirmed by radio dynamic spectra at DH wavelengths can directly provide accurate estimates of the coronal Alfven speed (Va) and hence the magnetic field if the electron density is inferred. Since the projection effects are minimal for Earth-directed CMEs, one can measure the shock standoff distance ΔR from the images. Since the CME speed (V) and distance (R) are known, one can get the Alfvénic Mach number (M = V/Va) from the following well-known relation that connects ΔR to R, M, and γ, the ratio of specific heats (Bennett et al., 1997):

$$\Delta R = 1.1\ R\ [(\gamma-1)M^2 + 2]/(\gamma+1)M^2].$$

This is especially important because there is no easy way of measuring magnetic field strength in the outer corona and near-Sun IP medium. Finally, a heliospheric imager at L5 can provide the density estimates in IP CMEs, which can be combined with Faraday rotation measurements to get the CME magnetic field strength and orientation.

2.2 CIRs

When a high speed stream from a coronal hole overtakes a slow wind ahead of it, a CIR is formed. This is because CIRs align themselves along the Parker spiral and rotate in the direction of planetary motion, so one observes a CIR structure at L5 ~4 days before its arrival at Earth. Figure 4 shows the Parker spiral structure when STEREO B was at L5 on 2009 October 25. The spiral field lines at STEREO B connect to the central meridian in an Earth view. Recently, Simunac et al. (2009) used in-situ observations from STEREO B and A separated by ~60° to show that it is possible to predict the properties of a CIR at STEREO A based on the observations at STEREO B. This separation is similar to the L5 – Earth separation and, hence, demonstrates the usefulness of the method. CIRs cause geomagnetic storms because of the amplified Alfvénic fluctuations and compression contained in the stream interface, with a magnetic field strength similar to that in magnetic clouds (Gopalswamy, 2008). CIRs are primarily identified using plasma signatures such as proton density (peaks at the stream interface), proton thermal pressure (peaks at the interface), solar wind speed (peaks about half a day after the interface), and the flow direction, which changes from west to east at the interface (see Gosling 1996 for a review). Simultaneous magnetic field measurements additionally characterize the CIR magnetic content and structure, which decide the strength of the geomagnetic storm. Although the STEREO mission's Heliospheric Imagers can identify CIRs, the in situ measurements provide information on the magnetic field, density, and speed of CIRs, which are crucial in predicting the strength of the geomagnetic storm that would occur when the CIR arrives at Earth.

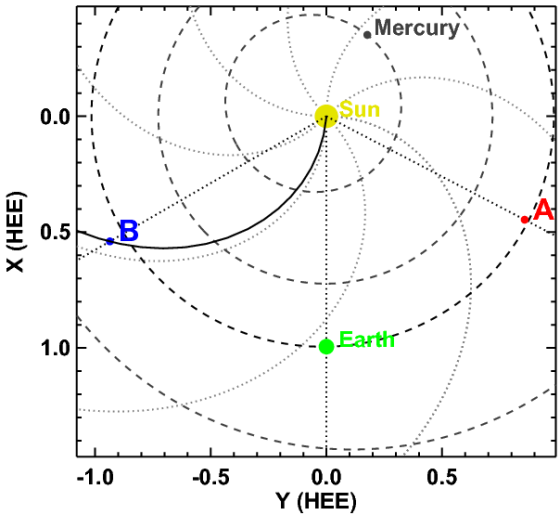


Figure 4. Solar magnetic field in the IP medium represented by the Parker spiral for 2009 October 25 (in heliocentric Earth Ecliptic (HEE) coordinate system). STEREO B and A were located near the Sun-Earth L5 and L4, respectively. The CIR observed at L5 (marked by the solid spiral field line) would rotate to Earth in about 4 days (neglecting the orbital motion of Earth, which is very small compared to the solar rotation).

3 Solar Sources

An inner coronal imager similar to the EUV instruments on board SOHO and STEREO can readily observe active regions and coronal holes, which are the solar sources of CMEs and CIRs, respectively. EUV images show several signatures of an eruption: the flare arcades, EUV waves, coronal dimming surrounding the flare location, and the eruptive filament (see Gopalswamy et al., 2009b for details). In addition, the EUV images from L5 can provide advanced warning of active regions and flaring activity that would be rotating on to the disk to face Earth. Active regions at the east limb in L5 view are almost at the backside of the Sun in Earth view (E150). Backside CMEs and their solar sources observed from L5 thus provide advance warning of

activity centers moving on to Earth view in a few days. An L5 EUV instrument will be able to provide information on regions that are up to 60° behind the east limb in Earth view. EUV images also provide similar information on coronal holes that are likely to rotate to the front side. Magnetograms obtained from L5 view can also track the complexity and evolution of the active regions and coronal hole regions well before they come to the Earth view. Surface magnetic field measurements from L5 will be helpful to modeling efforts that depend on extrapolation of the photospheric field. Note that the east limb contains the least known (oldest) fields and weakest input to the heliospheric models and hence the L5 input will greatly improve this situation.

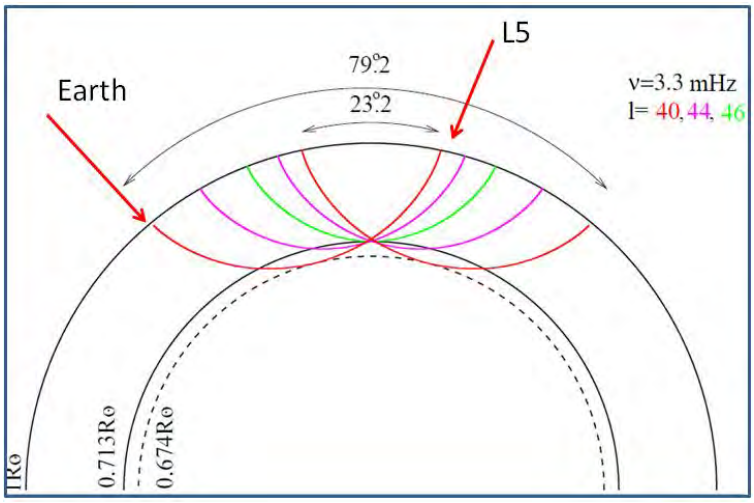


Figure 5. Simulation results show that acoustic modes with a frequency of ~3.3 mHz with l values between 40 and 46 reach the tachocline (to heliocentric distances of 0.67 and 0.71 Rs). Two views from Earth and L5 can study the l = 40 mode using correlation techniques. The computations were performed using the standard solar model and ray theory calculations.

4 Helioseismic Investigation of Solar Magnetism

The origin of all the magnetic field in the heliosphere, starting from the active region fields that result in violent eruptions to the field clumps in the heliosphere that modulate galactic cosmic rays, is thought to be at the tachocline. The generation and maintenance of the solar magnetic field and its evolution are central to the solar variability that affects Earth's space environment and the entire heliosphere. Helioseismic techniques aid in probing various layers of the solar interior using acoustic modes trapped inside the Sun (see e.g., Christensen-Dalsgaard et al. 1996). Of particular interest are the modes trapped between the base of the convection zone and the solar surface. It is expected that the modes that reach the bottom of the convection zone bounce off at the surface at 60° intervals (see Fig. 5), rendering it attractive to make Doppler measurements from L5 and Sun-Earth line. Existing helioseismic measurements from the Sun-Earth line (ground and space) can be combined with those from L5 for studying the base of the convection zone. Subsurface flow pattern changes (indicative of twisted magnetic fields) preceding large flares have been reported recently (Reinard et al. 2010), showing that local helioseismology may become useful in predicting large solar eruptions.

5 The Baseline Mission

Table 1 shows the elements of the straw man EASCO mission to make both remote-sensing and in-situ measurements along with nominal spatial (SR) and temporal (TR) resolutions. The mission will not need much instrument development and will use instruments similar to those on existing L1 missions. The mission could be launched around 2020 to begin observations in cycle 25. The Magnetic and Doppler Imager (MADI) will measure the photospheric magnetic and velocity fields. A white light coronagraph (WCOR) and a heliospheric imager (HI) will image

the corona in the range 2.5 – 45 Rs. Combination of polarized brightness images from WCOR and those at L1 will be useful in predicting Earth-arriving CMEs. The Inner Coronal Imager in EUV (ICIE) will provide coronal images at EUV wavelengths. A combination of MADI and ICIE images can characterize the solar surface and corona containing magnetic regions with eruptive potential. The low-frequency radio telescope (LRT) will isolate CMEs driving shocks near the Sun using the type II radio bursts, providing valuable information on the shocks that eventually impact geospace. The Solar Wind Plasma Instrument (SWPI) and a magnetometer (MAG) will make in-situ measurements of the solar wind providing information on CIRs that will arrive at Earth ~4 days after being detected at L5. The energetic particle detector (EPD) will detect protons in several energy channels that will be useful in addressing flare and CME-shock contributions to large SEP events.

Table 1 Baseline Payload of the EASCO mission

Instrument	**Measurements**	**FOV**	**SR**	**TR**
Magnetic and Doppler Imager (MADI)	Photosph. B, V	Full disk	2”	90 min
White-light Coronagraph (WCOR), HI	Coronal images	2.5 – 45 Rs	1-2’	10, 60 min
Inner Coronal Imager at EUV (ICIE)	Coronal images	0 – 1.2 Rs	2”	5 min
Low-frequency Radio Telescope (LRT)	Dyn. spectrum	1 – 100 Rs	--	1 min
Solar Wind Plasma Instrument (SWPI)	Plasma param.	In situ	--	1 min
Solar Wind Magnetometer (MAG)	Magnetic field	In situ	--	1 min
Energetic Particle Detector (EPD)	SEP intensity	In situ	--	1 min

References

Akioka,M., et al., AdSR 35, 65, 2005

Christensen-Dalsgaard et al., Science 272, 1286, 1996

de Koning, C. A., V. J. Pizzo, and D. A. Biesecker, Solar Phys. 256, 167, 2009

Gopalswamy, N., JASTP 70, 2078, 2008

Gopalswamy, N. et al., Ann. Geophysicae 26, 3033, 2008

Gopalswamy, N. In CAWSES book, Terrapub, Japan, 2009

Gopalswamy, N., Thompson, W. T., Davila, J. M., et al., Solar Phys. 259, 227, 2009a

Gopalswamy, N. et al., Earth Moon and Planets 204, 295, 2009b

Gopalswamy, N. et al., JASTP, accepted for publication

Gosling, J. T., Ann. Rev. Astron. Astrophys. 34, 35, 1996

Ontiveros, V. and Vourlidas, A., ApJ 693, 267, 2009

Reinard, A. A., Henthorn, J., Komm, R., and Hill, F., ApJ 710, L121, 2010

Simunac, K. D. C., L. M. Kistler, A. B. Galvin, et al., Ann. Geo. 27, 3805, 2009

Webb, D. F. et al., Space Research Today 178, 10, 2010

Value and Need of Helio and Space Physics Summer Schools

Nicholas Gross, Center for Integrated Space Weather Modeling, Boston University
Jan Sojka, Utah State University

Value and Need Extra-Institutional Courses in Space Weather

Helio and Space Physics are specialized subfields that do not fall neatly into the typical disciplinary departments and can draw researchers and students from many departments including: physics, astronomy, engineering, earth science and atmospheric science. Because of this, a 'standard curriculum' that will provide students with the important underlying principles or a broad overview of the field and its broader impacts has not emerged. In fact, there is strong consensus among many faculty members who do research in space physics that a strong undergraduate foundation in a traditional discipline (physics, engineering, geoscience, atmospheric science, etc.) is more important then particular knowledge of the space physics field. At the undergraduate level, space physics content can play a role in providing interesting concrete examples of fundamental principles, but it will not be the primary focus of a course or even a unit.

Graduate school is usually the first exposure that perspective space physicists have explicitly to space physics content. Even in graduate school, the topic is usually covered only as a special research topic and then with a focus on teaching a traditional discipline rather then on conceptual or systems level understanding of space weather or helio-physics. Thus students often gain an introduction to the field only through their particular specialization and so may develop a narrow view of the system.

For this reason a number of summer schools have been organized to fill this gap, most recently the Center for Integrated Space Weather Modeling (CISM) "Space Weather Summer School" (SWSS) [Simpson, 2004], supported by the NSF, and the NASA "Heliophysics Summer School" (HSS). These schools are designed to expose students to an overview of the field as well as present work at the frontiers of research. The CISM summer school starts with the assumption that students have an understanding of physics at the undergraduate level but have little knowledge or experience of solar and space physics and its effects. It primarily services graduate students new to the field as well as space weather professionals who need a broad overview. The NASA HSS assumes some experience in the field and targets more experienced students and postdoctoral researchers already in the field. Together these summer schools are training a cadre of the next generation of space physics researchers and professionals in a way that cannot be done in a traditional department setting.

The Center for Integrated Space Weather Modeling - Space Weather Summer School

The Center for Integrated Space Weather Modeling (CISM) offers a two-week Space Weather Summer School (SWSS; http://www.bu.edu/cism/SummerSchool/overview.html) intended primarily for first-year graduate students, though it is taught at a level and in a manner suitable

for well-prepared upper division undergraduates, and also space weather professionals. Over the last 10 years over 300 participants have attended the summer school. Graduate students from CISM and non-CISM institutions attend from across the country and internationally. These participants are interested in research topics from across the space physics disciplines (solar physics, the heliosphere, the magnetosphere and the ionosphere) as well as space craft design, forecasting, and other operational issues. Participants come from a wide variety of departments including, astronomy, physics, engineering, earth science, and meteorology. Approximately 20% of the participants have been professions with a space weather focus. These include forecasters and other staff from the NOAA's Space Weather Prediction Center (SWPC), engineers from NASA's Space Radiation Analysis Group (SRAG), researchers and managers from the Air Force Research Lab (AFRL), officers from the Air Force Weather Agency (AFWA), and aerospace engineers and airline pilots concerned about operational and safety issues related to space weather. These space weather professionals are 'trained' in their particular specialty, but attend the summer school to obtain a broad overview that puts their expertise and operational knowledge into context. These participants not only gain from the summer school, but also add to it by providing other students with lessons from their experience working in the field.

During the two weeks of the summer school the focus begins at the Sun and then follows the connections through the solar wind, the magnetosphere, and finally the ionosphere and thermosphere. A typical daily schedule of the summer school involves three lectures connected to the broad daily topic in the morning, and an afternoon 'lab' session. Experts in the field deliver the morning lectures and each lecture falls into one of three series: the space environment ("reality"), its effects on technological systems and humans ("harsh reality"), and models used to specify or predict the space environment ("virtual reality"). During the morning lectures modern pedagogy techniques are applied such as 'Peer Instruction' [Mazur], in which students respond to a question posed by the lecturer and talk to each other about it, and the use of question cards [Angelo and Cross], where students are asked to pose anonymous questions to the lecturers who then answer them after lunch.

In afternoon computer labs, students use results from these models and data to explore the concepts discussed in the morning sessions. The students are divided up into heterogeneous groups of three or four (mixing backgrounds, interests, and experiences). Each group is provided a computer and a common workspace to sketch pictures, write down results and argue about interpretation. Through guided inquiry, students use the computer to access data and visualize simulation results so that they may explore the concepts discussed during the morning lectures.

The value of the student centered focus of the SWSS structure is two fold. First the students are given an opportunity to follow their own path towards constructing their knowledge base in the field. Second, the curriculum as a whole and the afternoon labs in particular allow students to interact with the instructors and with each other on an individual basis. As described already, participants bring a wide variety of backgrounds and perspectives to the material presented, so they can learn a significant amount from each other in addition their interactions with the instructors.

NASA Living With a Star - Heliophysics Summer Schools
The Heliophysics Summer School (HSS; http://www.vsp.ucar.edu/Heliophysics/summer-about-over.shtml) is hosted by the UCAR Visiting Scientist Programs. This series of summer schools, commencing in 2007, helps graduate students and scientists learn and develop the science of heliophysics as a broad, coherent discipline that reaches in space from the Earth's troposphere to the depths of the Sun, and in time from the formation of the solar system to the distant future. Attendees of this summer school are all graduate students or early postdoctoral fellows working in space physics at both US and internationally institutions.

Each summer the HSS teaches a new cohort of students and develops the complementary materials that support the teaching of helio-physics at both graduate and undergraduate levels. This summer school has also adopted a schedule involving morning lectures and afternoon labs and also includes some time for recitation sections to discuss the homework. The afternoon labs have been based on materials developed at the Community Coordinated Modeling Center (CCMC; http://ccmc.gsfc.nasa.gov/). The HSS coordinators have also begun to work with CISM staff to adapt some of the lab materials developed for their summer school.

Outcomes and Products
The CISM Space Weather Summer School has run since 2001 and almost 300 participants have attended it. The NASA Heliosphereic Summer School has run for the last four year and over 100 students have attended. Some of students have attended both, but we can expect that by the end of 2012, well over 500 participants will have attended one or the other school. Both schools also have significant international participation providing students with the opportunity to network globally. Participant evaluations have rated both schools highly.

Both summer schools have produced valuable education products for the space physics community. The SWSS had developed and made available a series of lab exercises that follow the flow of space weather from the sun to the Earth [Gross, 2009; http://www.bu.edu/cism/SummerSchool/Labs/Lab_Index.htm]. These labs begin by looking at the structure of the solar coronal magnetic field, then follow the connections out to the solar wind, the earth's magnetosphere and then the ionosphere. Some materials are available on the web, but much of the data and visualization tools are available on a bootable DVD set for distribution. Extracts of these are being used for an REU program at LASP [http://lasp.colorado.edu/reu/] and at the HSS.

The original organizers of the HSS have developed three books [Schrijver and Siscoe] with contributions by the lecturers from the summer school. In addition the HSS organizers are beginning to collect sample problems for students to practice on. These three space science text books target an advanced clientele whose background needs to include plasma physics and advanced E&M. The chapters are designed for future practitioners of heliophysics. The texts follow an ordered sequence:

- Heliophysics: Plasma Physics of the Local Cosmos;
- Heliophysics: Space Storms and Radiation: Causes and Effects;
- Heliophysics: Evolving Solar Activity and the Climates of Space and Earth.

As the HSS continues, instructors have been asked to develop problem sets to accompany these books.

The Road Forward

The continued popularity of the CISM and NASA summer schools strongly suggests the need for what these summer schools provide. Both summer schools are fully subscribed. For the CISM summer school between 1/3 and 1/2 of the applicants to the CISM summer school over the last several years have been from advisors or organizations that have sent previous participants. The NASA summer school has not yet been running long enough to build up a significant number of 'repeat customers' but based on the CISM model it is expected that it will.

CISM funding will end in 2012 and the summer school is funded through then. Initial planning for continuing the summer school after 2012 has begun, but at this time no funding has been secured. We hope that funding for the Heliophysics Summer School will continue, but clearly this will only handle about half the demand and serves a different the more advanced portion of the audience. An opportunity to replace the CISM Summer School should be available in order to provide a valuable introduction to heliophysics and space weather for students early in the graduate career and space weather professionals who need a broad overview. Direct costs for this summer schools is modest. It includes travel costs for the participants and instructors, honorarium for instructors, cost for the room and computer equipment. Direct costs total approximately $50,000. The hidden costs involve time for a school directors and coordinator.

Further expansion of both efforts could include: evolving and expanding the existing lab activities, developing homework sets and estimation activities for both schools, and providing more or all of the materials in an 'online' format. In fact, with new online meeting technologies ("GotoMeeting" or "Adobe Connect") that are available, some or all of the activities could be online, where students and instructors are in different locations, but collaborating synchronously or asynchronously. These extended efforts would keep the summer schools fresh and allow for broader opportunities of participation for both students and instructors.

A small investment in these summer schools pays significant dividends both in training the next generation of space physics researchers and professionals, but also in providing an opportunity for new members of the diverse community to interact with each other and with more senior researchers and professionals. The authors here encourage the committee to include a call in the decadal survey for continued funding of these summer schools that are vital to the support of the space physics and helio-physics community.

Bibliography

Catherine H. Crouch and Eric Mazur, 'Peer Instruction: Ten years of experience and results', AJP, September 2001, Volume 69, Issue 9, pp. 970-977

Simpson, S. (2004), A Sun-to-Mud Education in Two Weeks, Space Weather, 2, S07002, doi:10.1029/2004SW000092.

T. Angelo, K. Cross; Classroom assessment techniques: a handbook for college teachers; Jossey-Bass Publishers, 1993

Lopez, R. E. and N. A. Gross (2007), Active Learning for Advanced Students: The Center for Integrated Space Weather Modeling Graduate Summer School, Adv. Space Res., in press, doi:10.1016/j.asr.2007.06.056. (update this)

N. A. Gross, N. Arge, R. Bruntz, A.G. Burns, W.J. Hughes, D. Knipp, J. Lyon, S. McGregor, M. Owens, G. Siscoe, S. C. Solomon, M. Wilterberter; EOS, TRANSACTIONS AMERICAN GEOPHYSICAL UNION, VOL. 90, NO. 2, doi:10.1029/2009EO020001, 2009

Schrijver, C. and Siscoe, G.; Heliophysics: Evolving Solar Activity and the Climates of Space and Earth; Cambridge University Press, 2010

Schrijver, C. and Siscoe, G.; Heliophysics: Space Storms and Radiation Causes and Effects; Cambridge University Press, 2010

Schrijver, C. and Siscoe, G.; Heliophysics: Plasma Physics of the Local Cosmos; Cambridge University Press, 2009

White Paper for the 2010 Survey of the Decadal Strategy for Solar and Space Physics

Determination of optical spectra and g-values for negative ions of low-mass atoms and molecules

Emmanuel B. Grotheer[a, b], Stefano A. Livi[b, a], Rosemary M. Killen[c]

a) Department of Physics and Astronomy, University of Texas at San Antonio (UTSA), San Antonio, TX 78249, USA
b) Space Science & Engineering Division, Southwest Research Institute (SwRI), San Antonio, TX 78238, USA
c) Goddard Space Flight Center (GSFC), National Aeronautics and Space Administration (NASA), Greenbelt, MD 20771, USA

Disciplines: SHP – Solar and Heliospheric Physics
SWM – Solar Wind-Magnetosphere Interactions
AIM – Atmosphere-Ionosphere-Magnetosphere Interactions

Abstract

Currently, when studying ions in the field of heliophysics, this is usually done with a focus on positive-charged ions. In the case of the planet Mercury, one has the unique combination of a tenuous atmosphere, more properly called a surface-bounded exosphere, and the existence of an intrinsic planetary magnetosphere. The interaction of the solar wind and the Hermean magnetosphere can cause sputtering at Mercury's surface, thereby releasing regolith material into the exosphere. Through this, and other processes, the exosphere around Mercury is created. Photoelectrons at the surface released via interactions with solar radiation could potentially be captured by just-released particles, such as oxygen which has a high electronegativity. The resulting negatively charged ions would then have their paths altered not only by Mercury's gravity, but also by interactions with the planet's magnetic and electric fields. At the time of this writing, there is a lack of emission spectra and g-values, i.e. light scattering probabilities, for negatively charged ions. Hence, spectrometers can not be used at present to determine the abundance of negatively charged ions in Mercury's exosphere or other solar system bodies with tenuous atmospheres, such as comets. If research into determining the spectra and g-values of negative ions were made available, a new area of research in the field of solar and space physics would be opened up.

Introduction

The constituents of planetary exospheres can be detected remotely using optical techniques. One such method involves making intensity measurements for certain wavelengths known to pertain to an element's or molecule's emission spectrum. In order to determine the density of the measured species, one can use a relation of the form $4\pi N \alpha I/g$, where N is the density, and I the intensity. The g-value in this equation is the probability for light to be scattered by an atom for all directions, i.e. the g-value is a measure of probability for omni-directional emission, and a more detailed description of this parameter can be found in a 2009 paper by Killen et al. [1].

This technique was utilized by the UV airglow spectrometer on Mariner 10 as it examined the planet Mercury [2]. The MASCS instrument on the MESSENGER spacecraft, which has now completed 3 flybys of Mercury and will go into orbit around that planet in March 2011, utilizes the same data analysis principle [3]. Naturally, the application of this technique is not limited to the planet Mercury, and has been applied to the data gathered by various missions, such as the Voyager probes, Cassini-Huygens, and New Horizons.

At present, while there is a wealth of data on the spectra and g-values of neutral and positively ionized species, there is an unfortunate dearth of measurements for negatively charged ions. Some data on negatively charged species is available, such as in the text by Stemmler and Hites, but this is limited to more complex molecules with molecular weights of 94 and above, and thus not highly applicable in the field of space physics [4]. Still, similar techniques to those utilized for the creation of those spectra may be applied to determine the spectra and g-values for negative ions of species that are more commonly detected in solar and space physics missions.

The Giotto spacecraft detected negative ions, including O^- and OH^- among others, at comet Halley in 1986 [5]. For the case of Titan's ionosphere, the Cassini spacecraft was able to detect heavy negative ions, which are interpreted to include several heavy hydrocarbons, as well as the lighter O^- [6]. Note that in both these cases, the instruments used for detecting the negative ions utilized in-situ particle detection techniques, namely electrostatic analyzers. This paper focuses on expanding our capability of detecting negative ions in space environments through the use of optical techniques, which could thus potentially enable remote sensing of negative ions, assuming that their spectral lines do not coincide with those of other species. Such a technique could potentially be applicable for ground-based observations as well, so long as the negative ion emission lines coincide with the Sun's Fraunhofer lines, though choosing the proper viewing conditions can alleviate this concern by cleverly utilizing the Doppler shift effect [7].

Thus, we recommend that funding be made available for laboratory and theoretical research to determine the spectra and g-values of negative ions, such as O^- for example. With the resulting data, it may prove possible to utilize currently active space flight spectrometers to search for negative ion constituents in the exospheres of the instruments' respective planetary or other astronomical targets. Development of new instruments can then begin with the express purpose of looking for the heretofore largely-ignored negatively charged constituents of planetary exospheres.

The case for Mercury

Mariner 10 completed three flybys of Mercury, and proved that there is great scientific insight to be gained from studying this planet [8]. The first Mariner 10 flyby of Mercury found evidence of a planetary magnetosphere, which ignited new interest in the study of this previously neglected planet [9]. As Siscoe pointed out, the study of other planets helps us to better

understand our own planet. This becomes especially true when a particular type of process is more pronounced or less encumbered with complex interactions with other processes compared to the case at Earth. Mercury's magnetosphere, out of all the others in our solar system, is the most like Earth's, but its smaller size and lack of a dense atmosphere make it an excellent test case for studying magnetospheres and perfecting computer models thereof [10].

One of the processes that can be responsible for releasing atoms and molecules from Mercury's surface into the exosphere is ion sputtering. Sputtering refers to the process where a particle, often referred to as the projectile, impacts on a target surface with sufficient energy to cause a cascade of collisions. The resulting transfers of momentum end up giving some of the surface's constituent particles a trajectory and enough kinetic energy to overcome the binding energy of the target material and thus are ejected from the target, an example of this is shown in Figure 1 below. In addition to momentum transfer, there is another mechanism known as potential sputtering which can be quite efficient at releasing particles from the surface as well.

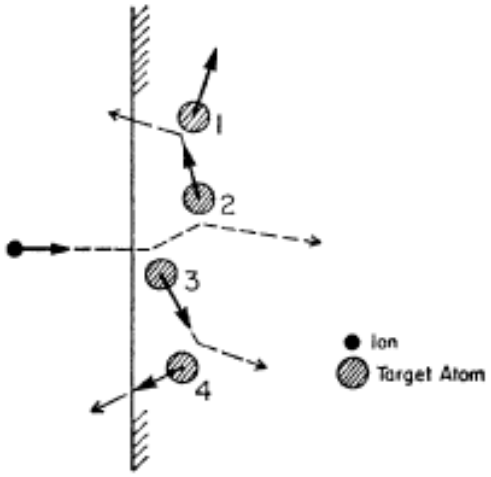


Figure 1: Ion sputtering schematic [11]

At Mercury, the intrinsic magnetic and electric fields provides an acceleration mechanism which guides ions, whether they originated from the solar wind or from Mercury's surface via some previous release process, to the surface and those with sufficient energy then cause sputtering when they impact.

The solar system is filled with micrometeoroid debris which can impact on larger bodies such as planets. At Mercury, the lack of a dense atmosphere allows micrometeoroids to impact at a fairly constant rate. Larger meteoroids impact at lower frequencies and are thus often considered fairly negligible at Mercury. Micrometeoroids with a radius in the range of 10^{-8} to 10^{-1} m are considered to be of the greatest importance [12]. When these micrometeoroids impact on the Hermean surface, the resulting local temperatures can reach roughly 5000 K. With such high temperatures and the resulting pressures, these impacts vaporize portions of the surface and can release even refractory elements into the exosphere [13].

Currently, research efforts into the constituents of Mercury's exosphere focus on neutral species, especially sodium which has been studied extensively in the literature {cf. [14], [15], [16]}, and positive ions [17]. The sole in-situ data of oxygen, from Mariner 10 as re-examined by Shemansky, suggests a near-ground density of 4.4×10^4 cm^{-3}. In this same paper, the sodium near-ground density is given as $1.7 - 3.8 \times 10^4$ cm^{-3}, as can also be seen in table 1 below [18].

If one includes even the weak intensity emission lines, the spectra for O I (i.e. neutral oxygen) extend from 685.544 to 26,173.56 Å, or from the ultraviolet to the infrared portion of the electromagnetic radiation spectrum. For positive ionization, the spectra are known to shift, so that for O IV (i.e. O^{3+}) the spectra are confined to the ultraviolet region from 195.86 to 3744.89 Å. Similarly, for O V (i.e. O^{4+}) the spectra covers the ultraviolet to visible range from 124.616 to 6500.24 Å [19]. Thus, the spectra for negative ions of oxygen would probably fall within the ultraviolet to infrared band regions as well.

Table 1: Revised g-values at Mercury [18]

Species	Wave-length (A)	g(s^{-1})	N^0 (cm^{-3})
H*	1216	5.3×10^{-3}	23,230
He†	584	5.1×10^{-5}	6×10^3
OI‡	1304	2.1×10^{-5}	4.4×10^4
Ar§	867	5.5×10^{-8}	$<6.6 \times 10^6$
N_α	5890, 5896	2.45, 1.22	$1.7\text{-}3.8 \times 10^4$
K	7664, 7699	3.24, 1.67	5×10^2

*I = 2.25×10^{11} Ph cm^{-2} s^{-1} at 1 AU; $I\nu/I = 1.42 \times 10^{-2}$ cm.
†I = 1.28×10^9 Ph cm^{-2} s^{-1} at 1 AU; $I\nu/I = 3.52 \times 10^{-2}$ cm.
‡I = 5.19×10^9 Ph cm^{-2} s^{-1} at 1 AU; I(1,2)/I = 0.25; I(1,1)/I = 0.333; I(1,0)/I = 0.417; $I\nu/I = 7.9 \times 10^{-2}$ cm.
§$I\nu$=5.86×10^5 Ph cm^{-2} s^{-1} $(cm^{-1})^{-1}$ at AU.

Note that the g-values in table 1 are specific to the case of zero velocity, and g-values are velocity dependent, as is discussed by Killen et al. [1].

Given that oxygen is expected to account for 60% (by weight) of the Hermean crust, one would expect it to be significantly more abundant than sodium [13]. McGrath et al. estimated the sodium abundance in Mercury's crust to be about 0.2% [20]. As a rough calculation, if one neglects the difference in mass between oxygen and sodium, and we increase the estimate of sodium abundance to 1%, oxygen should be about 60 times as abundant at Mercury than sodium, with a density of roughly 1.8 x 10^6 cm^{-3}, so why is all this oxygen not detected?

Since oxygen has an electronegativity of 3.44 on the Pauling scale [19], one possibility is that the oxygen that is released from the surface, whether it be through sputtering or impact vaporization, actually captures an electron from the Hermean regolith and is thus released as a negative ion. The electron component of the solar wind may also contribute slightly to the population of electrons available at or near Mercury's surface for capture by just-released oxygen atoms. With the current suite of instruments available on MESSENGER and the upcoming BepiColombo missions, the negatively charged constituents of Mercury's exosphere will not be measured, and thus we may miss out on fully understanding why oxygen does not appear to be the major constituent of the Hermean exosphere even though it is the most abundant constituent of the regolith [21].

Conclusion

Since the ions in the solar wind are positively charged, it is not too surprising that in the field of solar and space physics not too much emphasis has been placed on negative ions. The spectra and g-values of negative ions of the relatively light species often encountered in heliophysical research are not known at this point in time. Since the solar wind contains electrons as well as ions, it is plausible that in interactions with solar system bodies with tenuous atmospheres these electrons and others released at the surface by solar radiation may actually be captured by electronegative atoms or molecules, such as oxygen for example. In the case of Mercury, one must further take into consideration the interactions with the planet's intrinsic magnetic and electric fields, which can both guide where solar wind electrons and ions impinge on the surface, as well as alter the course of subsequently released regolith particles, be they positively or negatively charged.

We hereby recommend that funding be made available for determining the emission spectra of negatively charged ions, such as oxygen and other species commonly measured in space missions, as well as their corresponding g-values. Based on these spectral measurements, from laboratory and theoretical research, the heliophysics community can then begin work on creating spectrometers that are capable of measuring negatively charged ions. Subsequently, the g-values can be utilized to interpret the intensity measurements from this new generation of spectrometers as corresponding densities of the measured negative ions. The ability to utilize remote optical sensing techniques to measure negative ions, whether it be through ground-based observations or spacecraft-mounted instruments, could open up a new area of study in the field of solar and space physics. This new approach could help us to gain a better understanding of the mysteries of the innermost planet, Mercury, as well as the atmospheres of other planets and even comets. When one considers that the combined partial pressures of the detected species in Mercury's exosphere lead to a total that is roughly 2 orders of magnitude lower than the total exospheric pressure measured by Mariner 10 [22], determining the abundance of negatively charged ions in Mercury's exosphere could be an important step in finding all the constituents of the Hermean exosphere.

References

1. Killen, R.M., D.E. Shemansky, and N. Mouawad, *Expected Emission from Mercury's Exospheric Species, and Their Ultraviolet-visible Signatures.* The Astrophysical Journal Supplement Series, 2009. **181**(2): p. 351-359.
2. Broadfoot, A.L., D.E. Shemansky, and S. Kumar, *Mariner 10: Mercury Atmosphere.* Geophysical Research Letters, 1976. **3**(10): p. 577-580.
3. Gold, R.E., et al., *The MESSENGER mission to Mercury: scientific payload.* Planetary and Space Science, 2001. **49**(14-15): p. 1467-1479.
4. Stemmler, E.A. and R.A. Hites, *Electron Capture Negative Ion Mass Spectra of Environmental Contaminants and Related Compounds*. 1st ed. 1988, New York: VCH Publishers. 390.
5. Chaizy, P., et al., *Negative ions in the coma of comet Halley.* Nature, 1991. **349**(6308): p. 393-396.
6. Coates, A.J., et al., *Discovery of heavy negative ions in Titan's ionosphere.* Geophys. Res. Lett., 2007. **34**(22): p. L22103.

7. Potter, A.E., R.M. Killen, and T.H. Morgan, *Solar radiation acceleration effects on Mercury sodium emission.* Icarus, 2007.
8. Clark, P.E.e., *Mercury: The Elusive, Enigmatic, and/or Endmember Planet.* The Mercury Messenger, 1987. **1**(1): p. 1-2.
9. Ness, N.F., et al., *Magnetic Field Observations near Mercury: Preliminary Results from Mariner 10.* Science, 1974. **185**(4146): p. 151-160.
10. Siscoe, G.L., *Mercury's Magnetospheric Laboratory.* The Mercury Messenger, 1989. **1**(3): p. 3.
11. Sigmund, P., *Theory of Sputtering. I. Sputtering Yield of Amorphous and Polycrystalline Targets.* Physical Review, 1969. **184**(2): p. 383-416.
12. Killen, R.M., et al., *Processes that Promote and Deplete the Exosphere of Mercury.* Space Science Reviews, 2007. **132**(2-4): p. 433-509.
13. Killen, R.M., T.A. Bida, and T.H. Morgan, *The calcium exosphere of Mercury.* Icarus, 2005. **173**(2): p. 300-311.
14. Ip, W.-H., *The Sodium Exosphere and Magnetosphere of Mercury.* Geophysical Research Letters, 1986. **13**(5): p. 423-426.
15. Potter, A.E. and T.H. Morgan, *Variation of sodium on mercury with solar radiation pressure.* Icarus, 1987. **71**(3): p. 472-477.
16. Baumgardner, J., J. Wilson, and M. Mendillo, *Imaging the sources and full extent of the sodium tail of the planet Mercury.* Geophysical Research Letters, 2008. **35**(3): p. L03201.
17. Vervack, R.J., Jr., et al., *Mercury's Complex Exosphere: Results from MESSENGER's Third Flyby.* Science, 2010. **329**(5992): p. 672-675.
18. Shemansky, D.E., *Revised Atmospheric Species Abundances at Mercury: The Debacle of Bad g Values.* The Mercury Messenger, 1988. **1**(2): p. 1.
19. Lide, D.R., *CRC Handbook of Chemistry and Physics*. 88th ed. 2008, New York: CRC Press. 2640.
20. McGrath, M.A., R.E. Johnson, and L.J. Lanzerotti, *Sputtering of sodium on the planet Mercury.* Nature, 1986. **323**(6090): p. 694-696.
21. Wurz, P., et al., *Self-consistent modelling of Mercury's exosphere by sputtering, micro-meteorite impact and photon-stimulated desorption.* Planetary and Space Science, 2010. **58**(12): p. 1599-1616.
22. Wurz, P. and H. Lammer, *Monte-Carlo simulation of Mercury's exosphere.* Icarus, 2003. **164**(1): p. 1-13.

Exploring the Physics of the Corona with Total Solar Eclipse Observations

Shadia R. Habbal[1], John Cooper[2], Adrian Daw[3], Adalbert Ding[4]
Miloslav Druckmüller[5], Ruth Esser[6], Judd Johnson[7] and Huw Morgan[1]
[1] *Institute for Astronomy, University of Hawaii*
[2] *NASA/GSFC, Heliospheric Physics Laboratory, Code 672*
[3] *NASA/GSFC, Solar Physics Laboratory, Code 671*
[4] *Institute of Optics and Atomic Physics, Technische Universitaet Berlin, Germany*
[5] *Faculty of Mechanical Engineering, Brno University of Technology, Czech Republic*
[6] *University of Tromso, Norway*
[7] *Electricon, Boulder*

Abstract

This white paper is a call for a concerted effort to support total solar eclipse observations over the next decade, in particular for the 21 August 2017 eclipse which will traverse the US continent. With the recent advances in image processing techniques and detector technology, the time is ripe to capitalize on the unique diagnostic tools available in the visible and near infrared wavelength range to explore the physics of the corona. The advantage of coronal emission lines in this wavelength range, over their extreme ultraviolet counterparts, is (1) the significant radiative component in their excitation process (in addition to the collisional excitation), which allows for observations out to a few solar radii, (2) the higher spectral selectivity available for imaging, giving well-defined temperature responses for each bandpass (one line as opposed to many), and (3) the capability of polarization measurements in a number of spectral lines. Consequently, the evolution of the thermodynamic and magnetic properties of the coronal plasma can be explored starting from the solar surface out to a few solar radii, namely the most important region of the corona where the expansion of the solar magnetic field and the acceleration of the solar wind occur. Since the planning of eclipse observations will not be possible without the invaluable NASA-published total solar eclipse bulletins by Espenak and Andersen, a call is also made to ensure continued support for these efforts.

Impetus: Examples from Recent Eclipse Observations

The realization that imaging the corona in the visible wavelength range provides unique diagnostic capabilities, compared to their EUV counterpart, emerged from the 2006 eclipse observations of the Fe XI 7892 Å line. In the composite of four eclipse observations of this spectral line (Figure 1), one outstanding feature is the unexpected extent of the Fe XI emission, out to at least 3 $R_\odot$ in streamers. The other striking feature is the appearance of regions with localized intensity enhancements at distances ranging from 1.2 to 1.5 $R_\odot$. These enhancements are often visible in some wavelengths and not others, as they are a function of the underlying plasma temperature. Curiously, there are also features in Fe XI that are much more striking than in their white light counterpart, as seen in Figure 2 (see Habbal et al. 2007, 2010a for details). Also striking are the twists in the large scale coronal structures, which are often associated with the magnetic field and density.

The observed radial extent of the emission from the coronal lines is a consequence of the dominance of radiative relative to collisional excitation in the formation of these lines, within a few tenths of a solar radius above the limb. Beyond that distance the intensity of the emission decreases with radial distance in proportion to the ion density. This is most pronounced for the Fe XI line (see Figure 7 in Habbal et al. 2007a) in comparison with the more 'standard' red (Fe X 637.4 nm) and green (Fe XIV 530.3 nm) lines. The extreme ultraviolet lines observed by EIT/SOHO, EUVI/STEREO and SDO, on the other hand, are

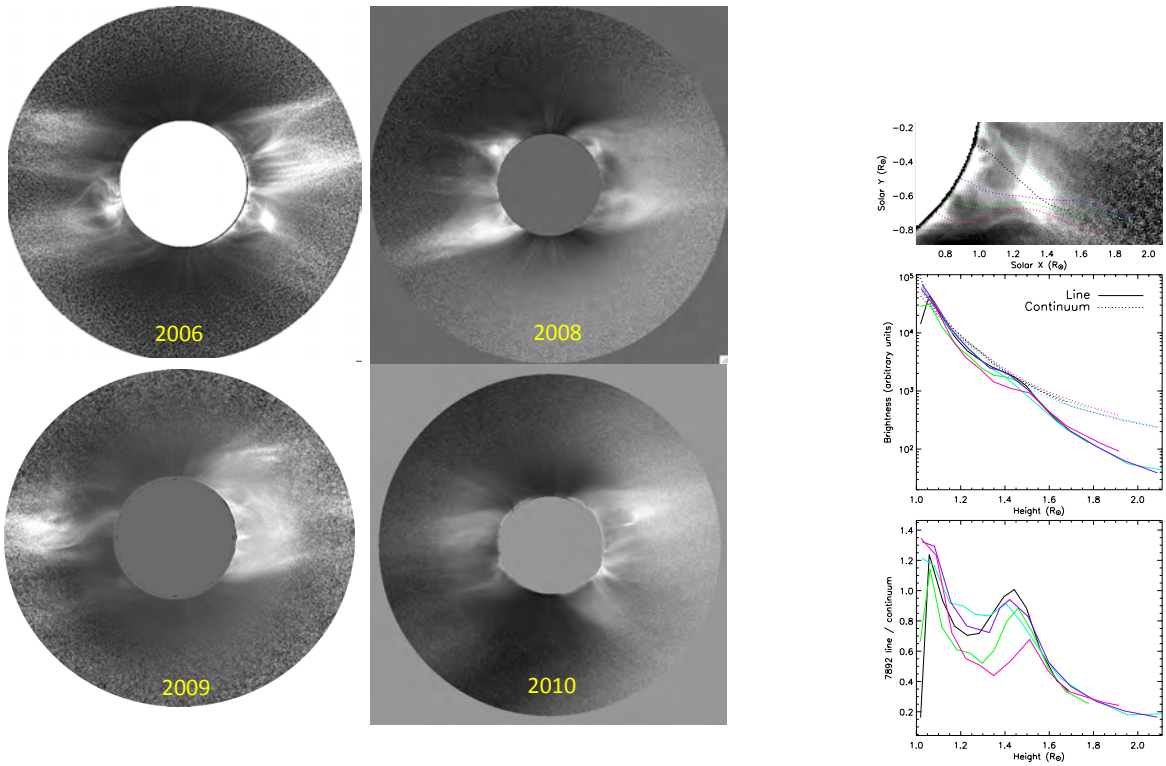


Fig. 1.— Left four panels: Fe XI 789.2 nm images of the corona taken during the eclipses of 2006, 2008, 2009 and 2010, and processed with NRGF (Morgan et al. 2006). This technique compensates for the radial falloff of intensity with distance without introducing any spurious artifacts in the images. Right panel: Trapping of ions in select magnetic structures as evidenced by the profiles of the Fe XI 789.2 nm line intensity (middle) and intensity ratio of Fe XI 789.2 nm and neighboring continuum at 787 nm (bottom) for the southwest streamers in the 2006 observations. Each colored curve is derived from a different coronal structure as indicated by the traces in the top panel. The transition to a predominantly resonantly scattered emission becomes evident when the slope of the ratio changes from a rapid falloff to an almost constant value as a function of radial distance.

dominated by collisional excitation. Their intensity thus falls off very quickly with radial distance as the electron density squared. While the maximum extent of the detectable UV emission is 1 $R_\odot$ above the limb, emission from the visible and near infrared Fe lines extends beyond 2 $R_\odot$, and emission lines of other elements can be observed as well.

The localized intensity enhancements, first observed in Fe XI, are a signature of increases in the Fe^{+10} density relative to that of electrons (see details in Figure 1, right panel). Model studies show that the localized ion abundance enhancements have implications for the distribution of heating in the corona, which also depends on the underlying magnetic geometries (Lie-Svendsen and Esser 2005).

Comparison with continuum observations yields the distance at which the emission transitions from being dominated by collisions to one dominated by resonant scattering, as shown in the bottom right panel of Figure 1. Such a transition thus corresponds to the transition from a collision-dominated to a collisionless plasma. As shown in Figure 3, such a determination can be made for each of the spectral lines observed. Since this transition does not occur at the same place for all the lines, implies a dependence on the density

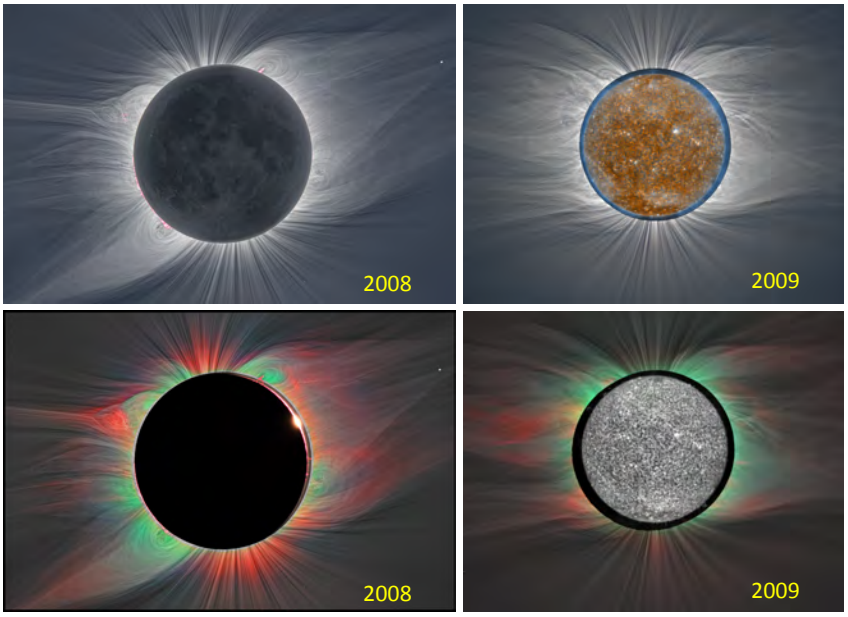


Fig. 2.— White light (top panels) and overlays of Fe XI 789.2 nm (red), Fe XIV 530.3. nm (green), and white light (lower panels) for the 2008 (left) and 2009 (right) eclipses. He II 304 SoHO/EIT are overlaid on the solar disk for the 2009 eclipse. Images have been processed using the technique developed by Druckmüller (2009).

and temperature of the underlying magnetic structure (Habbal et al. 2010a). The locus of the transition, R_t, between a collision-dominated regime to a collisionless one for a given emission line, ranging between 1.1 and 2 $R_\odot$, also determines the locus of the plasma's departure from ionization equilibrium (Habbal et al. 2010c).

Another unique feature of these two-dimensional multi-wavelength observations is the inference of the two-dimensional distribution of the electron temperature in the corona. As seen in Figure 2, there is a clear demarcation between emission in Fe X/FeXI (covering a temperature range of 0.8 to 1.2 $\times$ 10^6 K) and Fe XIII/FeXIV (characteristic of 1.6 - 2.2 $\times 10^6$ K plasmas) for most coronal structures. There is also a distinct dominance of cooler Fe X and Fe XI emission in polar regions. The presence of emission from all spectral lines, such as in some streamers in both 2008 and 2009 observations, is a reflection of a multi-temperature plasma there. The two-dimensional multi-wavelength composites can thus be regarded as maps of the coronal electron temperature placed in the context of density/magnetic structures. The 2006 - 2009 observations indicate that the expanding corona is dominated by cooler electron temperatures averaging 10^6 K (Habbal et al. 2010a).

The characteristic distance R_t where the transition from a collisional to a collisionless plasma occurs for each spectral line also represents the distance beyond which the emission reflects the distribution of its corresponding charge state. In the examples of 2006, 2008 and 2009, the emission was dominated by Fe XI

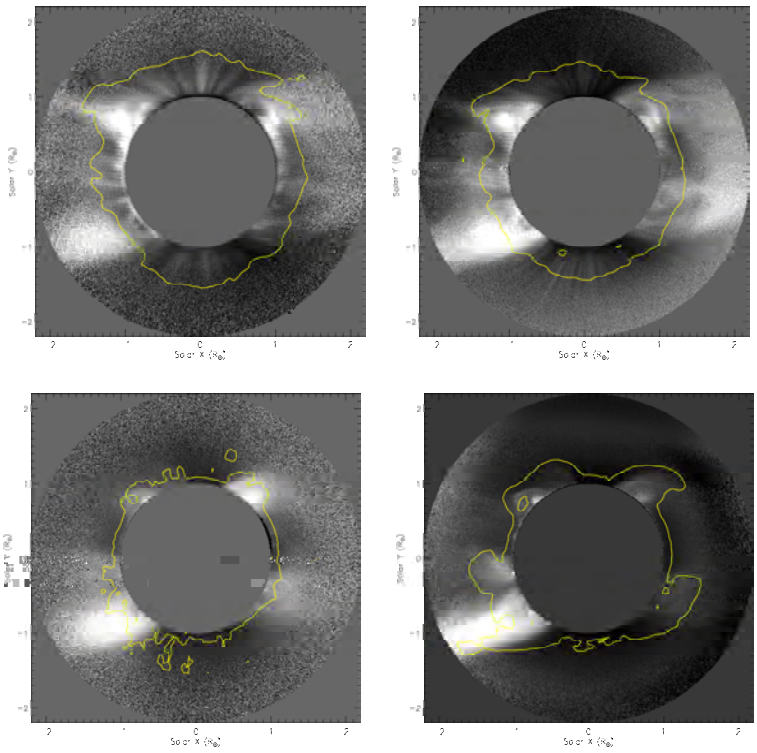


Fig. 3.— Contours of R_t overlayed on the Fe X (top left), Fe XI (top right), Fe XIII (bottom left) and Fe XIV (bottom right) 2008 eclipse NRGF-processed images. (From Habbal et al. 2010a).

implying a dominance of Fe^{+10} charge states in the expanding corona. The only charge state measurements available for comparison are those from interplanetary space, which show a distinct peak in the ion fraction of iron at Fe^{+10} in both the fast and slow wind during the periods covered by the eclipse observations. As shown in more detail in Figure 4, the dominance of an electron temperature of 10^6 K and the Fe^{+10} charge state in the corona match remarkably well the dominance of Fe^{+10} charge state measured in interplanetary space. This correspondence establishes an unambiguous link between conditions in the inner corona and interplanetary space (see Habbal et al. 2010b).

The multi-wavelength eclipse observations also revealed the role of prominences in defining large scale coronal structures and their underlying thermodynamic properties. As evident in the examples of Figure 2, and described in detail in Habbal et al. (2010c), prominences are enshrouded in hot plasmas of twisted magnetic structures (formally referred to as 'coronal cavities') which form the bases (or bulges) of streamers. The observed twisted helical structures within the bulges of streamers provide the most direct evidence for the emergence of helicity with prominences that is not limited to the prominences themselves but extends to their immediate surroundings. The fact that twisted structures surrounding cool prominence material are tenuous and hot, often exceeding 2×10^6 K, should have significant implications for coronal heating mechanisms in the neighborhood of magnetic polarity reversal regions.

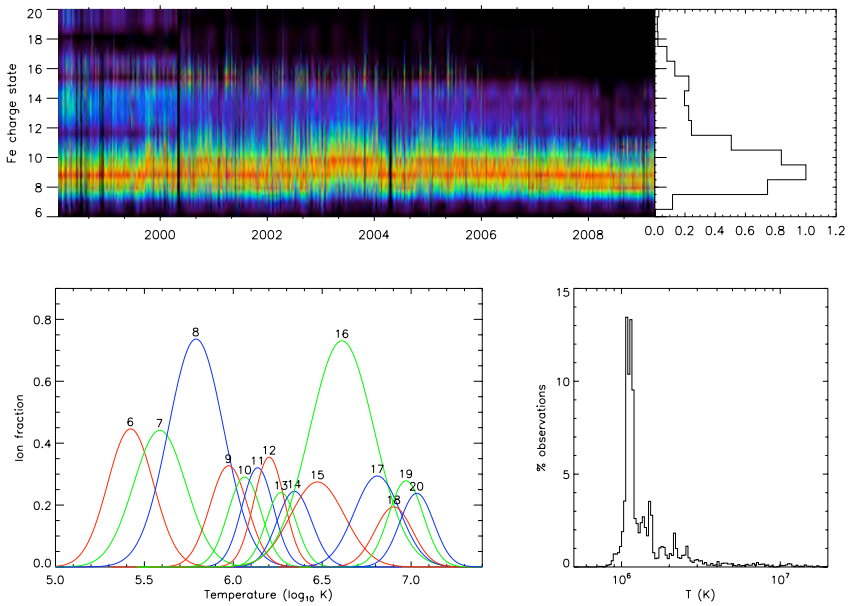


Fig. 4.— Top: 12 hour averages of the SWICS/*ACE* Fe^{n+} (with n = 6 to 20) charge state measurements as a function of time, from 1998 to 2009 (left), and their histogram averaged over the entire time interval, and normalized to unity (right). Lower left: Calculated Fe^{n+} ion fraction, versus temperature for the measured charge states, using data published by Mazzotta et al. (1998) cited in the CHIANTI database. Lower right: Calculated distribution of coronal electron temperatures that best reproduces the SWICS/*ACE* Fe charge state data above. The vertical axis is the percent of observations in each temperature bin. (From Habbal et al. 2010b).

The Case for Support for Total Solar Eclipse Observations in the Next Decade

As described above, the multi-wavelength eclipse observations in the visible and near IR lines of Fe taken between 2006 and 2010 led to a number of key insights into the magnetic and density structures of the corona, starting from the solar surface out to a few solar radii. The key property that enabled these discoveries is the radiative component in the emission process for these spectral lines, which dominates over the collisional component as the view moves away from the solar limb. The other important property is the well-defined temperature diagnostic offered by these coronal emission lines, as thin-film interference filters for the visible and near infrared are available with bandpasses even narrower than the Doppler-broadened coronal line widths. Consequently, spectral lines are well-separated and line profiles can even be explored. Such is not the case for EUV lines currently used in space-based observatories for temperature and density diagnostics of the corona, as multilayer coatings for EUV imaging provide wide bandpasses with multiple spectral lines in them. The recent eclipse observations provided new insight into the role of selective heating of minor ions in the inner corona, and how it leads to localized ion density enhancements, and is linked to the underlying magnetic structures. They established the locus of the transition from a collision-dominated to a collisionless regime. They also provided the first two-dimensional maps of the coronal electron temperature and coronal Fe charge states.

In addition to these examples, two very important diagnostic capabilities of the visible and near IR spectral ranges have to be kept in mind, namely the role of spectroscopy and polarimetry. Despite the proliferation of space-based observations of the solar corona starting with the OSO missions and Skylab in the late 60's and early 70's, spectra of the corona in the visible wavelength range remain rare, with the



Fig. 5.— Path of the 21 August 2017 eclipse across the US, with a total duration of 1h 33min 16.8s.

most comprehensive ones dating back to the 1960's. However, spectroscopy without contextual images falls short in yielding its full potential, as is evident from the different published accounts of some spectral lines recorded at some eclipses but not at others (e.g., Jefferies et al. 1971). Indeed, it is only when the underlying magnetic and density structures are identified in the corona, that one can make sense of the corresponding spectrum, i.e. the presence or absence of a given spectral line can be accounted for when the context of the spectrum is known, such as Hα being present in prominences, but not in other regions of the corona. Polarimetric observations in select spectral lines (Judge 1998) offer a powerful tool for coronal magnetometry (Lin et al. 2000). Although they are difficult measurements, primarily because of the stringent requirements on observing conditions, they are the only tools currently available for inferring the direction and strength of the coronal magnetic field. Polarization measurements carried out with different spectral lines will enable the identification of the coupled temperature-magnetic structures. Furthermore, the coupled polarization and imaging measurements can yield the locus of the transition from closed to open magnetic structures. Knowledge of this transition is essential for model studies of coronal heating and solar wind acceleration.

Unfortunately, no space-based coronagraph at present has any capability for narrowband imaging, spectroscopy or polarimetry in the visible and near IR wavelength range. Furthermore, even with technological advances in formation flying and/or deployable boom technology that would support a larger occulter at a greater distance, the level of stray light suppression would not match that available from occultation by the moon. At ground-based coronagraphs there is the additional issue that sky conditions are no match to the reduction in sky brightness during eclipses. A total solar eclipse provides a unique opportunity to observe the extent of the emission from the solar limb out to a few solar radii, the key to exploring the most critical region of space involving the evolution of the coronal magnetic field and the acceleration of the solar wind.

A space-based alternative to coronagraphs that would overcome the stray light issues would be an observatory in lunar orbit, providing a daily sequence of eclipse periods that would allow for coronal observations by a small instrument which could be part of an observatory for lunar and heliophysical science.

Given the limitations of our current space- and ground-based observatories, it seems that eclipse observations offer unique opportunities for further exploring the physics of the corona in the least ambiguous manner. The impetus for these observations can also drive the development for new detector technology for observations in the near IR. Furthermore, there is no doubt that such scientific endeavors cannot be carried out without the underlying support provided by the NASA eclipse bulletins that continue to be prepared by

F. Espenak and J. Andersen, which needs to be secured. The most desirable eclipse for organizing serious and coordinated observations is the 21 August 2017 eclipse which will cross the United States (see Figure 5). The total duration of this eclipse across the US of over an hour provides a further unprecedented opportunity for detailed observations of the spatial scales of coronal structures, and their evolution and variation in time.

REFERENCES

Druckmüller, M. 2009, ApJ, 706, 1605

Druckmüller, M., Rušin, V., & Minarovjech, M. 2006, Contributions of the Astronomical Observatory Skalnate Pleso, 36, 131

Habbal, S. R., Morgan, H., Johnson, J., Arndt, M. B., Daw, A., Jaeggli, S., Kuhn, J., & Mickey, D. 2007a, ApJ, 663, 598

Habbal, S. R., Morgan, H., Johnson, J., Arndt, M. B., Daw, A., Jaeggli, S., Kuhn, J., & Mickey, D. 2007b, ApJ, 670, 1521

Habbal, S. R., et al. 2010a, ApJ, 708, 1650

Habbal, S. R., Morgan, H., Druckmüller, M., and Ding, A. 2010b, ApJ, 711, L75

Habbal, S. R., et al. 2010c, ApJ, 719, 1362

Jefferies, J. T., Orrall, F. Q., and Zirker, J. B. 1971, Sol. Phys., 16, 103

Judge, P. G., ApJ, 500, 1009-1022, 1998

Lie-Svendsen, O & Esser, R. 2005, ApJ, 618, 1057

Lin, H., Penn, M. J., and Tomczyk, S. 2000, ApJ., 541, L83

Morgan, H., Habbal, S. R., & Woo, R. 2006, Sol. Phys., 236, 263

Pasachoff, J. M., Rušin, V., Druckmüller, M., Aniol, P., Saniga, M., & Minarovjech, M. 2009, ApJ, 702, 1297

This preprint was prepared with the AAS LaTeX macros v5.2.

Helioseismology

F. Hill[6], S. Basu[13], D. Braun[7], Y. Elsworth[1], L. Gizon[5], I. González Hernández[6], S. Jefferies[9], S. Korzennik[2], A. Kosovichev[8], J. Leibacher[6], C. Lindsey[7], E. Rhodes[12], M. Roth[4], P. Scherrer[8], M. Thompson[3], J. Toomre[11], R. Ulrich[10], J. Zhao[8]

[1]Birmingham University
[2]Harvard-Smithsonian Center for Astrophysics
[3]High Altitude Observatory
[4]Kiepenheuer-Institut für Sonnenphysik
[5]Max Planck Institute for Solar System Research
[6]National Solar Observatory
[7]Northwest Research Associates
[8]Stanford University
[9]U. Hawaii
[10]UCLA
[11]U. Colorado
[12]U. Southern California
[13]Yale University

Summary
Virtually all solar phenomena originate in the Sun's interior, thus a full understanding of solar physics requires a comprehensive description of the conditions below the photosphere. This can only be done with helioseismology, which has already provided many significant advances in solar physics. Progress towards current helioseismology science goals would be best accomplished with ground-based multi-wavelength observations, and space-based multi-viewpoint measurements.

Introduction and Goals
All solar phenomena begin with the generation of energy by the fusion reactions in the core of the Sun. During the 10^7-year long journey of a photon from the core to the surface, the energy it carries can be transformed into thermal, kinetic, or magnetic forms. In turn, this energy drives large-scale motions in the convection zone creating differential rotation and generating the magnetic fields that eventually erupt on the solar surface where they cause the solar activity which controls the atmosphere and affects modern society.

The only way that we can probe the solar interior is with helioseismology, the study of the waves that are mainly trapped below the photosphere. These waves are primarily acoustic in nature, and were discovered in 1960. Helioseismology has advanced solar physics in many ways, including

- Demonstrating that the interior was not rotating sufficiently rapidly to call into question predictions of General Relativity.
- Demonstrating that the neutrino flux deficit did not arise from errors in our knowledge of conditions in the solar core
- Substantially validating the standard solar model
- Precisely determining the depth of the solar convection zone
- Mapping the internal solar rotation rate leading to the discovery of the tachocline (Figure 1)
- Detecting activity on the solar far side
- Measuring the surface torsional oscillation deep into the convection zone
- Showing that meridional and zonal flows in the solar convection are linked to activity cycle
- Showing that the oscillation frequencies track the solar cycle
- Determining the structure beneath sunspots
- Discovering that strong subsurface vorticity underlies flare-productive active regions

Scientific Motivation
The Sun is a star, and understanding the solar interior opens up new paths to understanding the interior of all stars. Conversely, understanding other stars places the Sun in astrophysical context and increases our understanding of its likely future evolution. Asteroseismology has become a very active field with the launches of CoRoT and Kepler, and the interplay between helio- and asteroseismology enriches both endeavors.

The most promising research topics for the coming decade for understanding the solar interior are:

- Understanding large-scale flows and constraining solar dynamo models
- Studying convection and rotational shear immediately below the photospheric boundary
- Understanding the near-surface excitation, damping, and *p*-mode coupling to the atmosphere
- Determining the internal structure and evolution of stellar active regions
- Reconciling the new abundances and stellar interior models

- Understanding the physical basis of flare and CME production and the development of space weather predictions

We discuss three of these topics in some more detail.

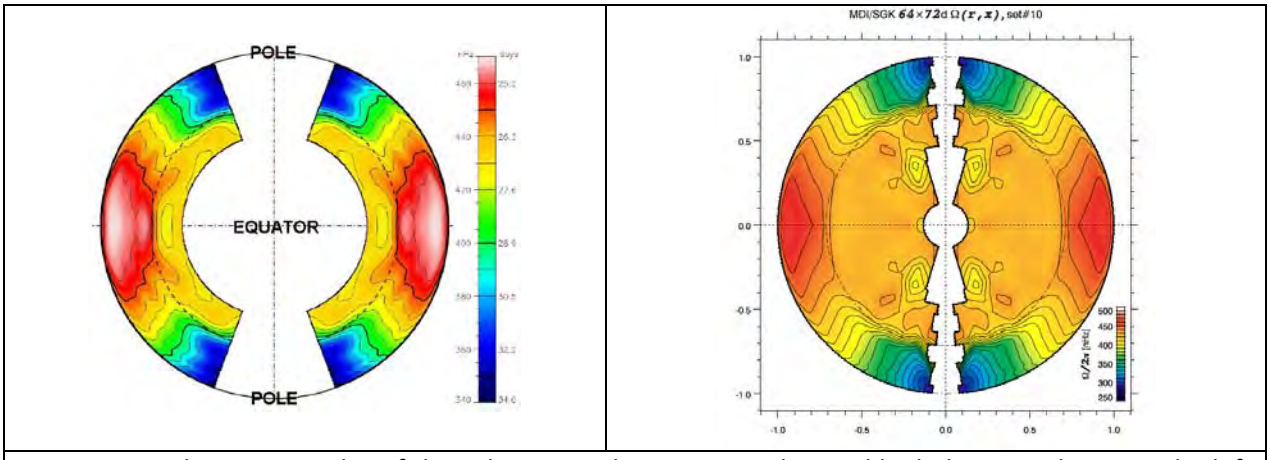


Figure 1: A color contour plot of the solar internal rotation as obtained by helioseismology. On the left is a plot using a few months of data obtained in the early 1990s; the right panel shows a modern inversion from S. Korzennik using 12.6 years of data, demonstrating the improvement obtained with long data sets. These observations revealed that the surface differential rotation extends downward completely through the convection zone, below which it is replaced with solid body rotation. The transition region from differential to solid-body rotation at the base of the convection zone (indicated by the dashed line in the plot) is known as the tachocline. This observation was in contradiction to the then-current dynamo theories of solar activity, which required that the internal rotation be constant on cylinders parallel to the solar rotation axis.

Solar dynamo — The Sun's 22-year Hale magnetic cycle dominate the long-term measures of solar variability. Essential aspects of our star's convective motions, meridional and zonal flow, and differential rotation that drive the cycle dynamo and the formation of individual active regions are still unknown. These motions are intimately connected with the solar cycle as shown by the recent discovery of the correlations between the speed of the meridional flow, the strength of the polar magnetic field and the amplitude of the subsequent solar cycle, as well as between the torsional oscillation migration rate and the temporal length of the recent deep and extended minimum.

Surface measurements of velocities and magnetic fields combined with helioseismic measurements of differential rotation and discovery of the tachocline and near-surface shear layer have provided important constraints on dynamo theories. However, even increasingly realistic models fall far short of adequate treatment of the basic physical processes, relying on parameterization of the widely varying physical and temporal regimes involved. Accurate measurements of the meridional flows throughout the Sun along with better characterization of non-axisymmetric dynamics over long intervals are essential ingredients for the development of adequate models to enable prediction of the solar cycles and long-term activity. In particular, the following key problems need to be addressed during the next decade:

- Determine the structure of the meridional flow at high latitudes and in the deep interior, and characterize the return meridional flow

- Determine the structure and dynamics of supergranulation in the polar regions, the role of supergranular diffusion in the magnetic flux transport, and search for large-scale organization of the supergranular convection (giant cells)
- Determine solar-cycle variations of the deep convection zone, tachocline and the nature of the possible 1.3- and 2-year periodicities
- Determine the 3D structure of the zonal flows ("torsional oscillation") and large-scale converging flows around active regions ("solar subsurface weather") including the relationship to surface flows; investigate the relationship between these flows and the formation of active regions and complexes of activity.

To solve these problems it is necessary to:

- Maintain the existing helioseismic observations provided by GONG and SDO
- Develop new capabilities for out-of-ecliptic observations of the polar dynamics and 3D stereo-type helioseismology for the reconstruction of the global 3D structure and dynamics of the Sun; these capabilities include proposed space missions: Solar Orbiter, Solar-C (Plan A), and Solar Polar Imager (SPI);
- Develop new methods of local and global helioseismology (supported by realistic numerical simulations), and a synergy of these methods, for determining the 3D structure and dynamics of the whole Sun.
- Maintain existing systems that measure surface velocities and magnetic fields.

Internal structure and evolution of active regions — Active regions apparently emerge quickly through the outer several megameters of the convection zone. Current state-of-the-art local-area helioseismology analysis with sufficient precision to resolve their subsurface characteristics requires observations collected over several hours, comparable to the emergence time. Once a sunspot complex has emerged, accurately measuring its characteristics requires the same precise knowledge to resolve what happens on the time scale of its evolution. There is good evidence from local helioseismology that the temporal evolution of subsurface helicity below a specific active region is related to the timing and strength of solar flares, thus local helioseismic techniques should be able to provide useful predictions of space weather. Over the course of the solar cycle, active region characteristics may change – we already know something about variations in location, orientation, and helicity, but not yet about complexity and ultimate source. In order to determine these variations, a large number of regions over the course of a cycle at various depths and ages must be observed, preferably with surface vector magnetic field measurements which provide the best estimates of the field characteristics. More sensitive and more deeply probing measurements require sufficient spatial resolution over a large part of the solar disk and better understanding of how magnetic, velocity and other physical conditions along the ray paths traveled by the sound waves affect their propagation. That includes the surface effects where the waves reach the solar surface. The Solar Dynamics Observatory (SDO) will help with the issue of spatial resolution, although multiple spacecraft observing from different directions around the heliosphere would be much better for probing more deeply. But it is doubtful whether observations with a single wavelength (i.e. a single height in the atmosphere) can completely characterize the modes. In addition, recent work that indicates the inclination of the magnetic field plays a vital role in determining the energy transport of the acoustic waves has opened up a new line of enquiry that will require nearly-continual vector magnetograms in addition to helioseismic observations.

We identify the following key problems for understanding the mechanisms of the formation of active regions and their flaring and CME activity, and for developing space weather forecasts based on helioseismic diagnostics:

- Investigate the evolution of subsurface flows and seismic sound-speed perturbations during the life-cycle of active regions for large samples during different phases of the solar cycle
- Develop methods of detecting active regions in the interior prior their emergence
- Determine the relationship between the subsurface dynamics and growth and decay of active regions for predicting their evolution
- Determine the formation and subsurface dynamics of large long-living complexes of activity and methods of their early detection
- Investigate the mechanism of the formation of sunspots, and their subsurface magnetic and thermodynamics structure and dynamics;
- Determine characteristics of subsurface flows (e.g. helicity, shear velocity) related to the flare and CME activity of active regions.

The required capabilities include the following:

- Maintain the continuous helioseismic observations of active regions during the full solar cycle from SDO and GONG
- Multi-viewpoint helioseismic and magnetic field observations of active regions using space missions, Solar Orbiter, Solar-C (Plan-A), SPI
- Multi-wavelength observations of individual active regions and sunspots from SDO, GONG, Hinode/SOT, MOTH and Mt. Wilson
- Substantial improvement of local helioseismology methods for diagnostics of magnetic regions
- Large-scale radiative MHD simulations of sunspots and active regions for understanding effects of turbulent magnetic fields on wave excitation and propagation and to test methods of local helioseismology

Subsurface turbulent magneto-convection, wave excitation, and the coupling to the atmosphere — The immediate sub-photosphere is highly superadiabatic, and the hydrogen ionization zone is a source of convective driving. It is an extremely complex turbulent physical environment that can dominate many of the aspects of the waves. Thus lack of knowledge of the physics in this region greatly hampers our interpretation of some helioseismic observations. Spectacularly detailed simulations of turbulent convection in the upper layers of the convection zone seem to mimic very faithfully what we observe on the surface of the Sun: granulation, intense downflows, and the creation, transport, and destruction of small magnetic flux elements. However these models are limited in two important ways: Current observations of subsurface conditions are not good enough to validate many of the features seen in the models. The models, which need to span the huge range of spatial and temporal scales in order to adequately sample and calculate the scale heights, densities, temperatures, pressures, and magnetic fields, cannot be computed with today's computers. The importance of understanding how solar oscillations gain and lose energy provides important constraints on the turbulence generating the waves. Better oscillation physics in turn enables the proper interpretation of helioseismic data and understanding of subsurface dynamics.

The key requirement is to combine high-resolution measurements with theory of physical conditions where the solar waves are excited and propagate. Understanding how the modes couple to the photospheric plasma and how energy is transferred to the overlying magnetized atmosphere also

impacts understanding of the energetics and dynamics of the chromosphere and low corona. This will require simultaneous observations at multiple heights in the atmosphere under many conditions and over large regions of the solar surface. Multi-line measurements may also help reduce the background noise in the oscillation spectra.

The key problems that need to be addressed during the next decade are:

- Understanding of the wave excitation mechanism in the quiet Sun and magnetic regions
- Investigation of wave interaction with magnetized turbulence
- Determination of wave propagation and transformation properties in strong magnetic field regions of sunspots
- Investigation of wave leakage into the atmosphere and the role of waves in the energetic and dynamics of the chromosphere

These problems require the following capabilities:

- Coordinated multi-wavelength high-resolution Doppler and spectro-polarimetric observations of the solar photosphere from space missions and ground-based observatories (SDO, Hinode, Solar Orbiter, GONG, MOTH, NST, ATST, SOLIS, Mt. Wilson, BiSON)
- Coordinated high-resolution imaging and spectroscopic observations of the chromosphere (NST, ATST, IRIS, Gregor)
- Large-scale high-resolution radiative MHD simulations of the plasma dynamics from the subphotosphere to the chromosphere and corona
- Development of accurate and efficient sub-grid scale turbulence models of the magnetized plasma

Critical capabilities needed for future helioseismology studies of the solar interior

Helioseismic observations of the polar regions – Determination of the differential rotation and the meridional circulation in the polar regions will provide unprecedented insights into the dynamics of the convection zone and the operation of the solar dynamo. Helioseismic determinations of the solar internal rotation profile rely primarily on global acoustic oscillations. Current observations of global modes are insensitive to the polar regions because of the difficulty of viewing these regions from the ecliptic. Thus, the differential rotation of the polar regions is currently unknown (See Figure 1). Existing helioseismic measurements suggest sharp spatial and temporal variations, including the possibility of a polar jet, but these measurements can only be confirmed by observing the polar regions directly, from an out-of-ecliptic orbit, and by applying methods of local helioseismology. Observations of the polar regions is a key to understanding the basic mechanisms of solar magnetic activity and predicting sunspot cycles. Recent observations of the polar regions from Hinode, during the periods of the highest inclination of the solar axis, have provided evidence for concentration of the Sun's dipolar field into compact kilogauss structures and for alignment of the supergranulation pattern. However, detailed investigations of the magnetism and dynamics are only possible from out-of-ecliptic space missions. Solar Orbiter, SPI and Solar-C (Plan-A) are among the potential missions that may be developed during the next decade to provide this capability.

Multi-viewpoint helioseismology – Current helioseismology observations allow us to determine the axisymmetric structure and dynamics of the Sun by the methods of global helioseismology, and also to obtain maps of 3D flows and sound-speed perturbations in a relatively shallow subsurface layer on the

front side of the Sun. They also allow detecting large active regions on the far side. However, stereo-type multi-viewpoint observations are necessary for the reconstruction the global 3D structure and dynamics of the interior. These observations will enable stereo-helioseismology based on long-base time-distance cross-correlation analysis of solar oscillations observed from the Earth orbit and a remote spacecraft. These observations will provide more accurate diagnostics of the tachocline and the deeper interior, including estimates of the spatial large-scale inhomogeneity of the solar core, and measurements of the core rotation. They will allow us to investigate the lifecycle of active regions by following their evolution for longer periods, and also investigate the nature of long-living complexes of activity ("active longitudes").

Multi-wavelength observations of solar oscillations – The critical data missing from all existing programs to improve this are observations at multiple heights in the solar atmosphere. These are required for adequate understanding of mode physics and interpretation of the mode characteristics to reveal what is really happening beneath the surface. In addition, high spatial, spectral, and temporal resolution observations over small areas will also provide information on the unresolved small-scale processes that may be important. Finally, nearly continuous long-term observations of the surface vector magnetic field will be needed to fold the magnetic field inclination into the analysis. Limited modern multi-spectral helioseismic observations with the MOTH instrument in Antarctica have been used to study *p*-mode characteristics in the presence of magnetic fields, and led to a discovery that the regions of inclined magnetic field can serve as "magnetic portals" through which acoustic energy can leak into the chromosphere. In the future, GONG could be a unique tool for probing the solar interior with improved instrumentation, multi-wavelength diagnostics and automated data analysis. The solar interior changes continuously, from cycle to cycle; there is evidence for long-term trends in solar magnetism and irradiance. Advances in our understanding of stellar interiors require long-term observations of the conditions inside the Sun.

To meet these needs, substantial modification of the GONG network would be necessary. Preliminary discussions of an upgraded GONG network concept contain the following elements:

- Simultaneous Doppler and magnetic field observations in 4 to 16 different spectral lines. Possible instrument concepts include a Fabry-Perot (e.g. the Göttingen Fabry Perot Interferometer), fiber-fed multi-scanning spectrograph (e.g. FIRS), magneto-optical filter (e.g. MOTH), and Michelson interferometer (e.g. SIAMOIS).
- Images with a format of either 2k $\times$ 2k or 4k $\times$ 4k. This would require an entrance aperture of 20 or 40 cm to replace the current 10-cm diameter unit.
- The higher-resolution images would also require some type of image correction in order to achieve the full spatial resolution below the terrestrial atmosphere. Since the small aperture precludes adaptive optics, the image correction system might consist of a tip-tilt mirror and high-speed computing to perform differential destretching of the data as it is collected.

Synergy with realistic numerical simulations – Numerical simulations have become increasingly important for understanding the complex physics of the turbulent magneto-convection, formation and dynamics of sunspots, excitation and propagation of waves, and also provide artificial data for testing helioseismology methods. The helioseismic inversions are necessarily based on simplified wave models. Numerical "forward" modeling of more realistic conditions is absolutely essential for the verification and testing of the helioseismology methods and their improvements. Significant progress has been made for developing MHD simulations of wave propagation in a local 3D domain and in 3D models of the whole Sun , which are used for testing various aspects of helioseismic measurements and inversions, and for

understanding the differences and consolidating the inversion results obtained by different techniques. In the coming decade, a substantial role will be played by realistic radiative MHD simulations of solar turbulence and waves, from first principles. Several radiative MHD computer codes (Stein-Nordlund, MURaM, SOLAR_BOX) have been developed and used for modeling magneto-convection, magnetic structures, and oscillations. Currently, the simulations of this type are carried out for relatively small and shallow computational domains. There is a strong need of expanding these simulations for large and deep regions. This will require developing numerical methods, algorithms and codes, which can be used on massive parallel computers, 10^4-10^5 CPUs, which will be soon available (Teragrid and Blue Water systems). This is a substantial effort of critical importance for helioseismology requiring specific support.

Needs

The study of the physics of the solar interior, and its import for the solar atmosphere and corona, has been greatly influenced by the modern imaging helioseismic programs of GONG and SOHO/MDI, now extended by SDO/HMI. The next decade should bring online the ATST, a SOLIS Network, multi-spectral helioseismology, continued operation of GONG, and the beginning of routine multi-viewpoint helioseismology from space.

Top priority should be given to these initiatives:

- Adequate support for analysis of SDO, GONG, Mt. Wilson, BiSON and MOTH data and for development of new helioseismic diagnostics.
- Support for continued long-term operation of GONG to supply highly reliable, comprehensive ground-based synoptic observations of the changing Sun in cooperation with SDO. It is essential to have two systems to cross-calibrate the measurements and to provide long-term continuity of the observations.
- Development of out-of-ecliptic missions for studying the structure and dynamics of the polar regions, and the mechanism of the polar field reversals and their role in the activity cycles.
- Development of multi-view helioseismology missions (Solar Orbiter, SPI, Solar-C/Plan-A) for stereo-type observations of the solar interior, helioseismic reconstruction of the global 3D structure and dynamics of the Sun, and investigation of the lifecycle of active regions and complexes of activity.
- Multi-height/multi-spectral observations of waves propagating in and through the photosphere for understanding the physics of solar oscillations and waves.
- Studies of the links between the turbulent convection zone and energetic and dynamics of the solar chromosphere and corona along with further development of local helioseismlgy methods to forecast space weather.
- A design study for a new-generation ground-based network for multi-spectral helioseismology.
- Substantial computational resources to support realistic MHD simulations of the multi-scale convection, dynamics of emerging magnetic flux, formation of sunspots and active regions, coordinated modeling and helioseismology analysis efforts.
- Joint programs combining helioseismology and dynamo theory to advance understanding of the solar dynamo.

With these initiatives, helioseismology will provide greatly improved observations of the Sun's interior, the seat of all solar phenomena.

ACRONYMS

ATST: Advanced Technology Solar Telescope
BiSON:Birmingham Solar Oscillation Network
CME: Coronal Mass Ejection
CoRoT: **CO**nvection **RO**tation and planetary **T**ransits, European space mission
FIRS: Facility Infrared Spectropolarimeter, located at NSO/Sac Peak
GONG: Global Oscillaton Network Group, gorund-based six-site helioseismology network
MHD: Magnetohydrodynamics
HMI: Helioseismic and Magnetic Imager, an instrument on SDO
MDI: Michelson Doppler Imager, an instrument for helioseismology on SOHO
MOTH: Magneto-Optical Filter at Two Heights
MURaM: The **M**ax-Planck-Institute for Aeronomy/ **U**niversity of Chicago **Ra**diation **M**agneto-hydrodynamics code.
NST: New Solar Telescope, located at Big Bear Solar Observatory
SDO: Solar Dynamics Observator, NASA space mission
SIAMOIS: Seismic Interferometer to Measure Oscillations in the Interiors of Stars
SOHO: Solar and Heliospheric Observatory
SPI: Solar Polar Imager
SOLIS: Synoptic Optical Long-term Investigations of the Sun, an instrument at Kitt Peak to measure vector magnetic fields
SOT: Solar Optical Telescope

The need for synoptic optical solar observations from the ground

F. Hill[3], J. Burkepile[2], D.P. Choudhary[1], M. Giampapa[3], S. Keil[3], P. Goode[4], J. Kuhn[7], K.D. Leka[5], A. Pevtsov[3], E. Rhodes[8], M. Thompson[2], R. Ulrich[6]

[1]California State University - Northridge
[2]High Altitude Observatory
[3]National Solar Observatory
[4]New Jersey Institute of Technology
[5]Northwest Research Associates
[6]U. California – Los Angeles
[7]U. Hawaii
[8]U. Southern California

Summary
The Sun can impact society through long-term and short-term changes. We cannot understand or potentially predict such variations unless changes in the solar state are monitored with the best possible suite of measuring systems. A variety of approaches including space- and ground-based instruments are currently engaged in this task and we focus here on the ground-based optical wavelength systems. There are many advantages of ground-based observations, including reliability, longevity, cross-calibration, cost, flexibility and educational outreach. Funding agencies need to recognize the value of long-term observations and provide sufficient resources to obtain and improve the measurements. The most important augmentations are to support new instrument scientists and their involvement with ground-based long-term solar observations and to develop a facility for archiving and dissemination of solar observational results similar to the Heliospherics Events Knowledgebase.

Introduction and Goals
The purpose of this white paper is to describe the scientific motivation for long-term ground-based solar observations, discuss the many advantages of these types of systems, provide the status of existing US facilities, and discuss what is needed to keep these systems in the forefront of heliophysics.

Note that the scope of this white paper is limited to synoptic ground-based optical observations. Long-term space-based observations, such as total solar irradiance (TSI) measurements, ground-based radio observations (i.e. FASR) and high-resolution ground-based solar observations (i.e. ATST) will be covered in other white papers.

Scientific Motivation
The generation of cosmic magnetic fields is one of the great mysteries in physics. How is it that macroscopic magnetic fields can be created and continually maintained in galaxies, planets, and stars on time scales of billions of years? In the case of the Sun, this question has been asked in the context of the dynamo mechanism of the solar activity cycle during which the number of sunspots is seen to rise and fall over a period of 9 to 13 years while the polarity of the global solar magnetic field reverses with a roughly 22-year period. During times of maximum sunspot population, the activity of the sun in terms of the occurrence of flares and coronal mass ejections (CMEs) is also at a maximum. These events can affect daily life in modern human society, as they are able to disrupt the technology that has become a vital part of our daily lives. Telecommunications, air transport, space activities, navigation, and power grids have all been adversely affected by events on the Sun and there are mechanisms whereby solar activity may be able to influence climate.

Thus, understanding the basis of solar activity is both a major scientific goal and an important societal activity. Since the activity cycle takes place on decadal time scales, its understanding requires long-term observations on the order of hundreds of years. Indeed, the study of solar activity can be dated back to 364 BC when the Chinese astronomer Gan De recorded sunspots in a star catalog. This was followed by about 2000 years of sporadic naked-eye sunspot observations in China, the Middle East and Europe, culminating with the 1610 invention of the telescope and Galileo's subsequent realization that sunspots were not planetary transits. After this, sunspot observations became more regular and accurate, to the point that, in 1843, Schwabe was able to detect a cyclical change in the number of sunspots. This led Wolf to examine the historical record of observations in the 17th and 18th centuries, and to devise the sunspot number which is still used to characterize the level of solar activity today.

While synoptic observations of the Sun are frequently used in short-term projects, their true value is in long-term datasets - some tendencies may only become apparent from the perspective of long-term

variations. How many of us have thought in the past: "if I only had the data from the previous cycle to verify my theory"? Synoptic observations are the data providers for future studies.

For example, more than 100 years of observations of the Sun in Ca K spectral line taken by several observatories are still being used in studies of long-term variations in solar irradiance. The synoptic solar full disk longitudinal magnetic maps taken at Kitt Peak, the Wilcox Solar Observatory, and Mt. Wilson are the only data sets covering multiple activity cycles for studying the role of polar magnetic fields in setting the cycle characteristics and the open magnetic field affecting the TSI. These studies were not conceived of at the beginning of synoptic observations, but with the advances in understanding of solar phenomena, the data are available to all researchers investigating these questions.

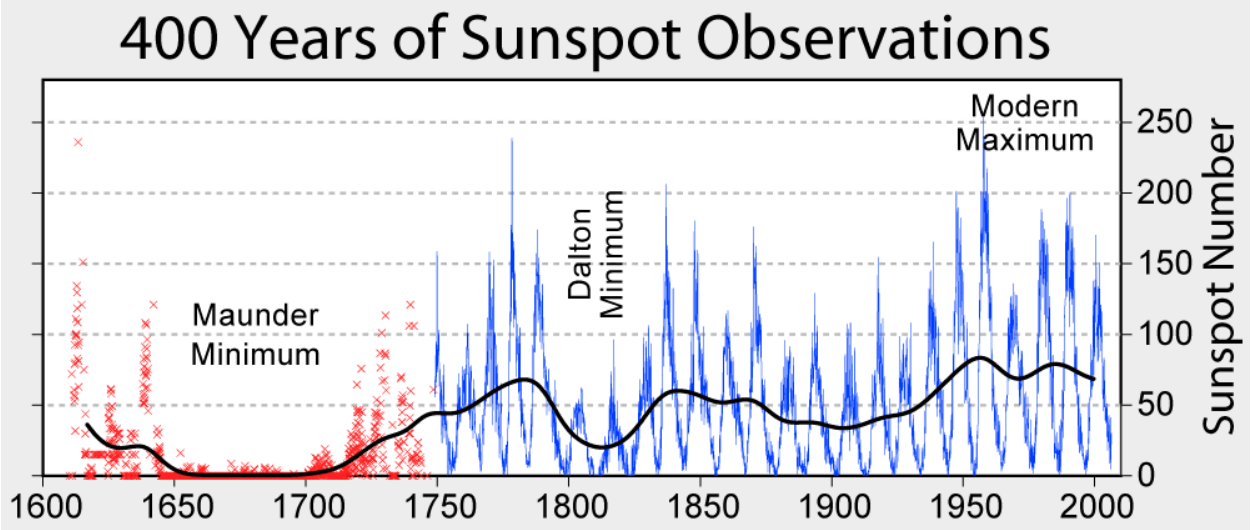


Figure 1: The sunspot number as a function of time over the last 400 years. The blue curve shows the monthly averages recorded by the Royal Observatory of Belgium. The red points are earlier, less reliable observations. The black curve shows the long-term variations. This figure was prepared by Robert A. Rohde and is part of the Global Warming Art project.

Figure 1 shows the number of sunspots as a function of time over the last 400 years. An uncertainty in this reconstruction comes from the need to adjust records from different observers to a common scale – a problem that continues to be present in a different but related form today. Figure 1 clearly shows the long-term variability of the level of solar activity. In particular, it shows the Maunder minimum when very few sunspots were visible. This period was also the time of the "Little Ice Age" in Europe, suggesting that there may be a link between solar activity and terrestrial climate. The large variability in amplitude of the sunspot number, along with the 40% range of the length of the cycle, suggests that the currently limited number of observed cycles (23 as of now, with cycle 24 just beginning) is insufficient to fully characterize the underlying dynamo mechanism and that we must continue to obtain observations over many more cycles.

This need is also indicated by the recent deep and peculiar solar minimum between cycle 23, which ended in 2008, and cycle 24, which only began with a significant number of spots in late 2009. During this minimum, which was 1.5 years longer than most, the Sun reached a very deep quiescent state with

- The most spotless days since cycle 15
- The lowest global solar wind pressure of the space age
- A solar magnetic field 36% weaker than last minimum

- Probably the lowest irradiance yet measured
- The lowest sustained 10.7-cm radio flux since 1947
- An unusually high tilt of dipole field
- No classical quiescent equatorial streamer belt
- Very high cosmic ray flux

Except for the first item in this list, none of these other aspects of the deep minimum were observable before WWII since the technology was not yet developed. In addition, there are considerable problems with using the sunspot number as an indicator of solar activity. Variations in observer bias, equipment, and terrestrial atmospheric conditions over 250 years make it extremely difficult to reliably compare important details of the activity. For example, there is now evidence that the strength of the magnetic field in sunspots has been declining over the last 20 years, and that the associated decrease in contrast will render sunspots invisible in the next 15 years or so. It is very possible that the same process occurred during the Maunder minimum, particularly since the Carbon 14 record indicates that the cosmic ray flux was being modulated with solar cycle periods during that time even though no sunspots were visible.

In addition to the space-age era measurements that date back three to five cycles, there are even newer measurements that are essential to unravel the solar cycle. One of these new techniques provides vector magnetic field observations to provide the direction of the surface magnetic field as well as an estimate of the true magnetic field strength and filling factor. Routine and consistent observations of the vector nature of the solar magnetic field are in their infancy, with sporadic ground-based observations dating back to the 1968 and synoptic programs starting up in the first decade of the 21st century. These observations have already detected a previously unknown and highly variable horizontal component of the solar magnetic field. The solar cycle dependence of the vector nature of the magnetic field is unknown at this point. We now know that the TSI is related to the distribution of magnetic field on the solar surface, but the relationship between TSI and the direction of the magnetic field has not yet been studied.

We now also have measurements of the velocity field inside and on the surface of the Sun. Velocity field measurements on the surface date back to the 1980s, and helioseismology has provided continual measurements of the flows inside the sun since 1995. These new internal flow measurements, particularly of the north-south meridional and the east-west zonal flows, are already being used by dynamo theorists to constrain their models.

A future measurement that is just now becoming possible is the direct measurement of the magnetic field in the solar corona. These observations will be invaluable for understanding the structure and dynamics of the corona over the solar cycle. Studies of this type are currently limited to using intensity images as proxies for the magnetic field. Interpreting this type of data has many challenges due to thermal and projection effects. Knowledge of the coronal magnetic field dynamics is also essential for studies of flare activity and CME evolution.

It is thus essential that long-term modern solar observations continue indefinitely in order to understand the solar cycle and the related activity that affects modern society. In the next section, we will discuss the advantages of using ground-based optical systems for this research.

Highlights of Science Results from Ground-Based Observations
A large number of fundamental solar characteristics have been discovered from ground-based measurements. Historically, some of the most significant are:

- Discovery of sunspots
- Discovery and measurement of solar surface differential rotation (the first astrophysical measurement)
- Discovery of the solar cycle
- Discovery of flares
- Discovery of the solar magnetic field
- Discovery of solar convection: granulation and supergranulation
- Discovery of the solar five-minute oscillations, development of helioseismology
- Discovery of the corona
- Discovery of the chromosphere
- Determination of the internal rotation rate

Some more recent discoveries and accomplishments are:

- Discovery of the horizontal magnetic field
- Discovery of flare-associated magnetic field changes
- Relationship between east-west zonal flows (torsional oscillation) and cycle timing
- Relationship between north-south meridional flows and cycle timing and amplitude
- Discovery of a secular decrease in sunspot magnetic field strength
- Relationship between subsurface vorticity and flares
- Discovery of coronal oscillations
- First measurements of coronal magnetic fields

Why ground-based?
In this era of significant solar scientific results from space-based solar missions such as SOHO, TRACE, Hinode, STEREO, SDO, etc. one might wonder why ground-based observations are needed. There are, in fact, several advantages of ground-based systems for long-term solar observations.

Only ground-based systems can reliably provide consistent long-term observations since instruments on the ground are always accessible and can be repaired. One can recall a loss of SOHO for an extended period of a few months. At that time, ground-based observations from NSO's Kitt Peak Vacuum Telescope were the only source of full disk solar magnetograms. Solar space missions are much less accessible, especially in the orbits to which they are now being launched (e.g. SOHO, STEREO, SDO). Satellites in near-earth orbit are accessible, but the transportation costs to the site are several orders of magnitude higher than for ground-based locations. Space missions are vulnerable to radiation exposure and micrometeorite collisions, and have limited lifetimes due to agency support and hydrazine consumption. In addition, space weather affects satellites reducing their utility as sources of space weather data at critical times.

Ground-based systems can be operated with much simpler and less expensive logistics than space missions. This is particularly true in terms of the return of the data from the instruments. Solar data rates are typically quite high, and are constantly increasing with advances in detectors and computing technology. Space missions have telemetry limits which restrict the amount of data return, and which require costly ground antennas to receive. Ground-based systems are limited only by disk-drive and tape capacities, which can be easily and inexpensively increased.

Ground-based observations are necessary to establish continuity between different space missions. An example is provided by the transition from SOHO/MDI to SDO/HMI. The calibration for MDI magnetograms was updated several times during the SOHO mission based on ground-based Mt. Wilson observatory measurements. A recent working group meeting has indicated that the new HMI vector magnetic field data may have unknown instrumental polarization that could be characterized using comparisons with ground-based observations. The cross-calibration of long-term observations over several decades is needed to provide a robust basis for understanding the Sun. In addition, synoptic ground-based solar irradiance measurements on a sustained basis with a cross-calibrated set of instruments provide a reliable reference for space measurements.

Table 1: US Ground-Based Solar Observatories

Observatory	Facilities	Data provided	Location	Founded
Big Bear Solar Observatory	New Solar Telescope	Magnetic fields, helioseismology, photometry	California, Global Hα network HQ	1969
High Altitude Observatory/NCAR	Mauna Loa Solar Observatory, 2 coronagraphs, 2 chromopheric imagers, photometer	Coronal magnetometry & photometry, chromospheric imaging, photospheric pho-tometry, helioseismology	Colorado (HQ), Hawaii (MLSO)	1940 (HQ) 1979 (MLSO)
Mees Solar Observatory	Polarimeter, spectrograph, coronagraph	Coronal magnetic fields and photometry, vector magnetic fields, photometry	Hawaii	1964
Mt. Wilson	150-ft and 60-ft Towers	Helioseismology, surface velocity fields, magnetic fields	California	1904
National Solar Observatory	Dunn ST, McMath-Pierce ST, Evans facility, GONG, SOLIS	Vector & LOS magnetic fields, helioseismology, photometry	New Mex-ico, Arizona, GONG network HQ	1947 (NM) 1959 (AZ)
San Fernando Observatory	Five telescopes, spectroheliograph, IR camera	Photometry	California	1976
Wilcox Solar Observatory	Magnetograph	Global magnetic fields	California	1975

Ground-based observations allow us to preserve and extend consistent data sequences. One of the big hindrances regarding any transition towards operational space weather products is that one needs large statistical databases on which to train, or develop a particular system, and then the same system needs to be available in order to provide the data for subsequent predictions. Long-term data provide essentially our only option for "control" data sets. As a remote observing science, our choices are limited in this really fundamental aspect of being able to perform physics experiments.

Ground-based observatories have great flexibility in testing innovative techniques and developing new capabilities. For example, chromospheric magnetic field measurements will provide a significant enhancement to photospheric magnetograms. It is likely that first routine full vector chromospheric magnetogram capability will be developed at a ground-based observatory, before a future space

mission. Ground-based observations typically employ state-of the art instrumentation using the most modern technology that does not have to withstand the rigors of the space environment. This often leads to important discoveries that subsequently guide the design of space-based instruments.

Ground-based observatories provide valuable opportunities for students to acquire hands-on experience in developing and operating instrumentation. This is very important for training and educating future generations of solar physicists and instrumentalists for both ground- and space-based programs.

Finally, ground-based observations impose many fewer restrictions on international collaboration as compared with space-based instrumentation. Limitations on advanced knowledge and technology transfer can restrict international collaborations in space research, while cooperation in ground-based observations can widen the collaboration and effectively share the costs of acquiring synoptic data.

Current Status

There are a number of ground-based observatories that are providing long-term optical solar measurements. In the US, these are listed in Table 1.

One of the most productive developments of ground-based solar observing systems has been the introduction of global networks, which are sets of instruments distributed geographically around the world. Networks are able to observe the sun nearly continuously for many years with only brief losses of data due to weather and equipment issues. Networks provide consistent observations spanning many years and are useful for capturing brief transient events. There are currently three major networks in operation: BiSON (Birmingham Solar Oscillations Network, non-imaging helioseismology, began operations in 1979), GONG (Global Oscillation Network Group, imaging helioseismology, began operations in 1995); and the Global High-Resolution Hα Network (Hα images, began operations in 2000). This observing strategy has proven to be very effective for general solar physics, as shown by the addition of magnetic field measurements and Hα images to GONG. One of the major goals of the next decade is to develop and install a network of instruments for vector magnetic field measurements.

There are several more ground-based solar observatories outside the US. Table 2 lists ones that are partners in global networks.

Table 2: Non-US observatories that are part of a global network

Observatory	Network	Location
BISON	Birmingham Solar Oscillation Network	United Kingdom (HQ)
Catania	Global Hα	Italy
Cerro Tololo	GONG	Chile
Hairou	Global Hα	China
Instituto Astrofisica de Canarias: La Palma and Teide Observatories	GONG	Spain
Kanzelhöhe	Global Hα	Austria
Learmonth	GONG	Australia
Meudon	Global Hα	France
Pic-du-Midi	Global Hα	France
Udaipur	GONG	India
Yunnan	Global Hα	China

Plans

All of the observatories listed in Table 1 have upgrades either in progress or planned. Typical projects range from improvements to data processing /storage facilities and new detectors to new instruments and even new telescopes. These projects range in cost from $10k to $300M. These projects will not be discussed in this white paper since the small ones are typically funded out of the observatories' operating budgets and the large ones will each have an associated white paper. These larger projects include a vector magnetic field network, an upgraded imaging helioseismology network, the Mauna Loa Solar Observatory coronagraph, the HAO COronal Solar Magnetism Observatory (COSMO), and the Advanced Technology Solar Telescope (ATST).

Needs

It is clear that long-term consistent observations are vital tools for understanding the Sun and its activity that impacts modern society. However, there seems to frequently be a preference in funding agencies for building new facilities, rather than the long-term operation of older systems. While the agencies certainly have many pressures from the research community and from Congress, the agencies should recognize the value of long-term continual consistent observations to understand the solar cycle and to provide relevant space weather data. We estimate that the current combined annual cost of running all of the observatories in Table 1 to be about $25-30M.

In addition to new instruments, it is important to maintain older instruments. The objective of establishing a reliable long-term record means that some instruments need to remain as nearly constant as possible. Many observable quantities depend on the details of how the solar spectrum is sampled so that changes in the record of the solar state may be from changes in the instruments rather than from changes in the Sun.

There are a number of specific areas where improvements could be made to increase the already substantial contribution of the ground-based observatories to heliophysics:

- Provision by the funding agencies to provide 3% annual budget increases to cover inflation.
- Quasi-periodic allocations for capital improvements in the form of instrumentation and detector upgrades and other improvements that reduce operations costs while enhancing scientific return on the investment in facilities.
- Establishment or augmentation of an associated theory program at each observatory to provide closer interaction between theory and observations.
- Establishment of a line of funding for research in statistical analysis specifically targeting space weather forecasting.
- Establishment of a national program to encourage and financially support new instrument scientists.
- Initiatives to establish new faculty positions in solar and space physics in order to increase student training in research and broaden the national research effort in this highly relevant area.
- Increased funding for the archiving, distribution, and analysis of ground-based data.
- Targeted funding for educational outreach programs at all levels combined with programs to broaden the participation of underrepresented groups.

With these improvements, the national investment in existing solar observatories would provide additional benefits to society in terms of increasing our ability to understand and forecast solar activity.

iDUST: Interstellar and Interplanetary Dust Near Earth
A mission concept for "dust tomography" of the heliosphere

Mihaly Horanyi, Eberhard Grün, Zoltan Sternovsky, Sascha Kempf, Ralf Srama, Siegfried Auer, Frank Postberg, Sean Hsu (contact: horanyi@colorado.edu)

Abstract: The observations of the inward transport of interstellar dust and the outflow of near-solar dust provide a unique opportunity to explore dusty plasma processes throughout the heliosphere. The flux, direction and size-distribution of interstellar dust can be used to test our models about the large-scale structure of the heliospheric magnetic field, and its temporal variability with solar cycle. The measurements of the speed, composition and size distribution of the recently discovered, solar wind-entrained nano-dust particles hold the key to understand their effects on the dynamics and composition of the solar wind plasma. Both the inflowing interstellar grains and the out-flowing nano-dust particles can be measured onboard a near Earth spacecraft at 1 AU, using recently developed dust detection techniques.

Transport of cosmic dust

Interstellar dust: After its fly-by of Jupiter, the dust detector onboard the Ulysses spacecraft detected impacts of particles in the mass range of 10^{-14} to 10^{-11} g, predominantly from a direction that was opposite to the expected impact direction of interplanetary dust grains. In addition, the impact velocities exceeded the local solar system escape velocity (Grün et al., 1993, 1994). Subsequent analysis showed that the motion of the interstellar grains through the solar system was parallel to the flow of neutral interstellar hydrogen and helium gas (**Figure 1**), both traveling at a speed of 26 km/s. The twenty year long journey of these small dust particles through the heliosphere is strongly coupled to the large-scale structure of the heliospheric magnetic fields (IMF) through almost two full solar cycles. Hence, monitoring the variability of the flux and velocity vector of ISD grains provides a unique test of our models of the IMF through solar cycles.

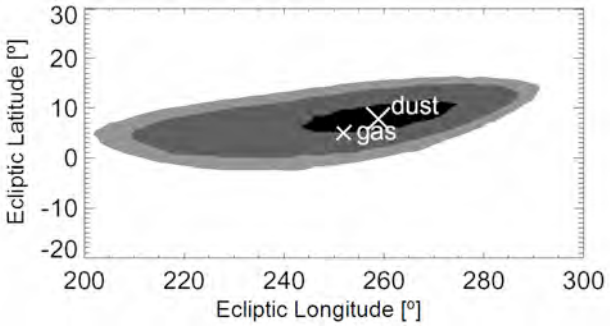


Figure 1. The upstream direction of the interstellar dust flux observed by Ulysses is λ=259±20° and θ=8±10° ecliptic longitude and latitude, respectively. The contour plot shows 1σ, 2σ, and 3σ confidence levels as black, dark grey, and light grey color. The helium upstream direction is λ=254°, θ=5.6° (Frisch et al., 1999).

Interplanetary dust: Early dust instruments onboard Pioneer 8 and 9 and Helios spacecraft detected a flow of submicron sized dust particles arriving from the solar direction. It was concluded that these particles originate in the inner solar system from mutual collisions among meteoroids and move on hyperbolic orbits that leave the Solar System under the prevailing radiation pressure force. As of to date little more has been learned since more sensitive dust could not face the solar direction because of severe interferences caused by solar UV radiation and solar wind entering the instruments. Only the dust instrument onboard the Ulysses spacecraft observed escaping dust particles high above the solar poles, which confirms the supposition that charged nanometer sized dust grains are carried to high heliographic latitudes by electromagnetic interactions with the IMF. Recently, the STEREO WAVES instruments recorded a large number of intense voltage signals in the ecliptic plane at 1 AU, which were interpreted as impact ionization signals of streams of nanometer sized particles striking the spacecraft at velocities of about solar wind speed **(Figure 2).** High fluxes of nanometer sized dust grains at low heliographic latitudes, as well as strong spatial and temporal fluctuations of dust streams uncorrelated with solar wind properties pose a mystery. The nano-dust, if real, represents a significant mass flux, which would require that the total collisional meteoroid debris inside 1 AU is cast in nanometer sized fragments.

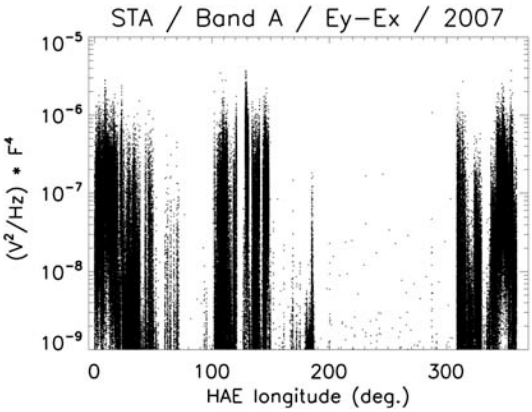


Figure 2. Average power observed by the STEREO/WAVES low frequency receiver on STEREO A as a function of ecliptic longitude during the whole 2007 orbit. The periods of high amplitudes have been suggested to be caused by high-speed nano-dust striking the spacecraft (Meyer-Vernet et al., 2009)

Dust dynamics:

a) ISD inflow. The interaction of interplanetary dust with the heliosphere depends on the grain's mass, charge, and material properties. The smallest interstellar grains (radii < 10 nm) remain tied to the interstellar magnetic fields and are diverted around the heliosphere (Frisch et al. 1999), only larger grains can penetrate the heliosphere. The Lorentz force created by the solar wind magnetic field sweeping past the grains both repels and focuses the grains, depending on the configuration of the field and thus on the phase of the solar cycle (Grün et al., 1994). In mid-1996, the flux of ISD grains was observed to decrease by a factor 3, and was interpreted to be the result of the electromagnetic interaction of the grains with the magnetic field of the solar wind (Landgraf et al., 2000). The polarity of this magnetic field changes with the 22-year solar cycle. During the solar maximum in 1991, the field polarity became North-pointing. The

azimuthal component of the expanding magnetic field deflects interstellar grains in the northern hemisphere to the north and grains in the southern hemisphere to the south. Thus, the net effect of the interplanetary magnetic field during the 1991 polarity cycle was to divert interstellar dust away from the ecliptic plane. During the 2000/2001 solar maximum the field polarity reversed again. However, interstellar grains need about 20 years (almost 2 solar cycles) to traverse the distance from the heliospheric boundary (at about 100 AU) to the Sun, and therefore, the changes in the magnetic field configuration does not immediately affect the dust flow. For example, trajectory calculations indicate a ~ 6 year delay between solar magnetic field-reversals and the shift in the flow of ISD grains from focusing toward to defocusing away from the ecliptic plan.

Ulysses measurements identified interstellar grains and determined their flow direction only by statistical means, and with large uncertainties. The details of the influence of the heliosphere on the trajectories of ISD grains remains poorly understood (Krüger and Grün, 2009)

b) IDP outflow. Our knowledge of the interplanetary dust cloud near the Sun is limited to in-situ measurements by Helios between 1 AU and 0.3 AU (Grün et al. 1980), remote-sensing observations, theoretical and modeling calculations. We have never done in-situ measurements of the size and spatial distributions, and chemical composition of dust, simultaneously with measurements of the solar wind flow, composition, and magnetic fields. In order to answer questions about the origin, and dynamics of the dust populations near the Sun and about their interactions with the solar wind. The open questions remain:
1) What is the dust environment in the inner heliosphere?
2) What is the nature of dust-plasma interactions close to the Sun?
3) What are the physical and chemical characteristics of dust-generated solar wind ions?
4) What is the effect of dust on the dynamics and composition of the solar wind flow?

The near-Sun dust population derives predominantly from comets and asteroids and to a lesser extent from the Kuiper belt and interstellar space (Leinert and Grün 1990). It comprises three particle families that differ in mass, size, origin, and orbital dynamics. Zodiacal dust (grain radius $a > 0.5$ µm; mass $m > 10^{-15}$ kg) originates beyond 1 AU, follows low-inclination, near-circular Kepler orbits, and is transported sunward by Poynting-Robertson (PR) drag. β-meteoroids ($0.1 < a < 0.5$ µm; $10^{-17} < m < 10^{-15}$ kg) are produced in the inner heliosphere from the collisional fragmentation, evaporation, and sputtering of the Zodiacal dust and are ejected on hyperbolic orbits from the near-Sun region by radiation pressure (Zook, 1975). Nanoparticles ($a < 0.1$ µm; $m < 10^{-17}$ kg) have been suggested to be produced by the collisional fragmentation of dust and by other processes of dust destruction (Mann et al. 2007). Owing to their high charge-to-mass ratio, nanoparticles are strongly affected by electromagnetic forces. Inside ~ 0.15 AU (30 RS) the combined effects of solar gravity and electromagnetic forces trap them on bound orbits; outside this region, whose spatial extent varies with solar wind conditions, the nanoparticles are picked up by the solar wind and accelerated to solar wind velocity (Czechowski and Mann

2010). This dust pickup process must have a yet unexplored effect on the expansion and composition of the solar wind.

The measurement of dust mass, composition, impact speed and direction at 1 AU provide critical constraints on models of the sinks, sources and transport processes of dust near the Sun. Coordinated dust, solar wind plasma, and magnetic field measurements can reveal the effect of electromagnetic forces on the complex dynamics and distribution of the smaller dust grains.

The effects of dust on the solar wind flow:

The ambient plasma and magnetic field environment both influences the near-Sun dust cloud and is influenced by it. Solar UV radiation causes the dust grains to acquire a positive surface charge. The charged grains are acted on by the Lorentz force, which perturbs the orbit of larger grains only slightly, but is the dominant force influencing the dynamics of grains ≤ 0.1 μm in size (Horanyi 1996). These grains can be deflected into high-inclination orbits from the near-ecliptic or can behave like heavy ions and be picked-up and entrained in the solar wind plasma flow (Mann et al. 2007).

Dust adds mass to the solar wind as it is a source of atomic and molecular ions and neutrals, which are rapidly ionized. The charged ions and molecules are picked up by the solar wind magnetic field (Lemaire 1990). However, the effects of mass loading by dust-generated species on the acceleration and dynamics of the solar wind near the Sun, including the position of the sonic point are not known and remain the subject of speculation. A slight increase in the average ion mass in the solar wind might, for example, generate shocks, at least in localized regions of high dust densities. These regions might also contribute to the energization of solar wind ions via two-stream plasma instabilities (Shapiro et al., 2005).

Mass loading by dust has been invoked to explain enhancements observed in the inter-planetary magnetic field (e.g. Russell et al. 1984; Jones et al. 2003). Recent results indicate that mass loading alone might not slow the solar wind sufficiently to account for these observations (Mann, et al. 2007). The exact nature, and the processes responsible for the generation of magnetic field enhancements remain an open issue.

Dust-generated neutrals and ions are thought to form an “inner source” of pickup ions (PUI) whose velocity distribution differs from that of interstellar (“outer source”) PUI. Both singly charged atomic ions and molecular ions (e.g., CH+, NH+, OH+, and H2O+) have been identified in the inner source population (Gloeckler et al. 2010). Processes proposed to explain the production of PUIs from the near-Sun dust include surface interactions, passage of solar wind particles through the dust grains, and gas production by sublimation, sputtering, and collisional vaporization. Modeling indicates that collisional vaporization can account for the heavy elements in the inner source PUI population as well as molecular species, while surface interactions produce significant amounts of pick-up protons (Mann and Czechowski 2005). PUI

production from dust sublimation or sputtering inside 0.1 AU is predicted to produce higher-charge-state PUIs; however, these processes have not been quantified (Mann et al. 2007).

Solar wind interaction with the near-Sun dust also contributes to the neutral solar wind and can generate energetic neutral atoms (ENAs) (Collier et al., 2001); recent results (Mewaldt et al. 2009) suggest that dust interaction with solar energetic particles may be an additional inner heliospheric ENA source.

Summary:

Dust is an important component of the heliospheric plasma environment. Dust carries significant mass, momentum and energy, yet the processes responsible for coupling it to the solar wind plasma flow remain largely unexplored. The nano-dust particles generated near the Sun and accelerated by the out-flowing solar wind, and the interstellar dust particles traversing the heliosphere can be measured at 1 AU, providing a 'tomographic image' of our heliosphere through dust measurements.

***iDUST* science goals**

#1 Explore the variability of the flow directions of ISD grains and verify our model of the large-scale magnetic fields and plasma distribution in the heliosphere, by comparing the accurately measured flux, mass, composition, and the velocity vector of individual ISD grains with the calculated trajectories, using global MHD models of the outer heliosphere.

#2. Explore the processes leading to the solar wind entrainment of small dust grains in the solar vicinity, and effects of dust loading on the dynamics and composition of the solar wind plasma flow, by comparing the flux, mass, and the velocity vector of nano-dust particles with our models of the inner heliosphere.

References

Collier, M. R., Moore, T. E., Ogilvie, K. W., Chornay, D., Keller, J. W., Boardsen, S., Burch, J., El Marji, B., Fok, M.-C., Fuselier, S. A., Ghielmetti, A. G., Giles, B. L., Hamilton, D. C., Peko, B. L., Quinn, J. M., Roelof, E. C., Stephen, T. M., Wilson, G. R., Wurz, P., Observations of neutral atoms from the solar wind, Journal of Geophysical Research, 106, 24893-24906, 2001

Czechowski, A.; Mann, I., Formation and Acceleration of Nano Dust in the Inner Heliosphere, The Astrophysical Journal, Volume 714, Issue 1, pp. 89-99 (2010).

Frisch, P. C. and 15 co-authors, Dust in the Local Interstellar Wind, Astrophysical Journal, 525, 492-516, 1999

Gloeckler, G., Fisk, L. A., Geiss, J., Schwadron, N. A., Zurbuchen, T. H., Elemental composition of the inner source pickup ions, Journal of Geophysical Research, 105, 7459-7464, 2000

Grün, E., Pailer, N., Fechtig, H., Kissel, J., Orbital and physical characteristics of micrometeoroids in the inner Solar System as observed by HELIOS 1. Planetary and Space Science 28, 333-349., 1980

Grün, E., Zook, H.A., Fechtig, H., Giese, R.H.: Collisional balance of the meteoritic complex. Icarus 62, 244-272., 1985b

Grün, E., H.A. Zook, H. Fechtig and R.H. Giese, 1985a, Mass Input into and Output from the Meteoritic Complex, In „Properties and Interactions of Interplanetary Dust“, eds. R.H. Giese and P. Lamy, Reidel, Doordrecht, S. 411-415

Leinert Ch., E. Grün, 1990, Interplanetary Dust, in Physics of the Inner Heliosphere I. eds. R. Schwenn, and E. Marsch, Springer Verlag,, Berlin, 207-275

M. Horanyi, Charged Dust Dynamics in the Solar System, Ann. Rev. Astron. Astrophys. 34, 383, 1996

Krüger, Harald; Grün, Eberhard, Interstellar Dust Inside and Outside the Heliosphere Space Science Reviews, Volume 143, Issue 1-4, pp. 347-356, 2009

Landgraf, M., Modeling the motion and distribution of interstellar dust inside the heliosphere, JGR 106, 10.303, 2000

Mann, I., Czechowski, A., Minato, T., Generation of pick-up ions and neutral atoms from cosmic dust, Physics of the Inner Heliosheath, 858, 366-371, 2006

Mann, I., Murad, E., Czechowski, A., Nanoparticles in the inner solar system, Planetary and Space Science, 55, 1000-1009, 2007

Meyer-Vernet, N.; Maksimovic, M.; Czechowski, A.; Mann, I.; Zouganelis, I.; Goetz, K.; Kaiser, M. L.; St. Cyr, O. C.; Bougeret, J.-L.; Bale, S. D. Dust Detection by the Wave Instrument on STEREO: Nanoparticles Picked up by the Solar Wind? Solar Physics, 256, 463-474, 2009

Shapiro, V. D.; Bingham, R.; Kellett, B. J.; Quest, K.; Mendis, D. A.; Bryans, P.; Torney, M.; Summers, H. P, X-Ray Emission from Comets and Nonmagnetic Planets. Theory and Comparison with CHANDRA Observations, Physica Scripta, Volume 2005, Issue T116, pp. 83-87. 2005

Zook, H.A., Berg, O.E.: A source for hyperbolic cosmic dust particles. Planetary and Space Science, 23 (1975) 183-203.

Zook, H.A.: Hyperbolic cosmic dust - Its origin and its astrophysical significance. Planetary and Space Science 23 (1975) 1391-1397.

Solar Flares and the Chromosphere

A white paper for the Decadal Survey

H. S. Hudson, L. Fletcher, R. Turkmani, S. L. Hawley, A. F. Kowalski, A. Berlicki, P. Heinzel

Background:
The solar chromosphere is at the center of many of the debates in solar physics. Highly complex, even in the quietest zones, the chromosphere is not an area that is easy to grasp. Across its few-Mm height it jumps from high plasma beta to low plasma beta, from neutral to ionized, from collisional to collisionless, becomes optically thin in the UV and mm-waves, and increases in temperature by almost three decades. It is no wonder that numerous instabilities occur. As observations improve, for example with the Hinode mission, these complexities become more apparent. Magnetic reconnection also occurs freely in this domain.

The chromosphere deserves attention in our attempts to find answers to the riddles of the corona, including flares. Understanding its magnetic complexity is a key to coronal modeling – particularly measurements of the vector field in the upper chromosphere are needed for a reliable reconstruction of coronal magnetic fields (e.g. DeRosa et al. 2009). The main provider of coronal heating may be in the transition region and upper chromosphere (Aschwanden et al 2007), and new evidence suggests that the chromosphere is the root of solar wind acceleration (De Pontieu et al. 2007).

Basic Characteristics

- The chromosphere and transition region anchor all coronal phenomena.
- Current systems that stress the coronal field must penetrate the chromosphere.
- A complex system of flows, shocks, jets, and waves originates in the chromosphere
- Chromospheric radiation contains most of the flare luminosity

The need for chromospheric observations of flares:
The solar flare spectrum peaks in the optical and ultraviolet, making the solar chromosphere the dominant source of radiation in a solar flare (Canfield et al 1986, Neidig 1989, Woods et al. 2004). In other words, solar flares radiate most of their luminous energy in the chromosphere. The chromosphere is where electrons accelerated in flares lose their energy, and may also play a role in their acceleration (Fletcher & Hudson 2008, Brown et al. 2009). Yet despite its evident importance, observational knowledge of this region is relatively sparse. From spectroscopic observations of much more powerful stellar flares, in which the optical flare output dominates that of the whole star, we know the powerful diagnostic potential of chromospheric flare radiation, and existing observations already point to the utility of spatially-resolved optical/UV radiation in decoding the evolving flare magnetic field. There is a pressing need to return to the chromosphere to progress on understanding solar flares.

Current Observations:
Current ground-based efforts utilize small custom instruments on mid-sized telecopes and concentrate on (1) patrol images (Hα, He 10830A, white light) and (2) some high resolution narrow-band spectroscopy of a few spectral lines (e.g. Hα, Ca 8542). There is little effort to conduct studies of chromospheric flares from space, though there has been notable success in targeted programs using TRACE, and serendipitous observations with Hinode, which have firmly established the relationship between the optical ('white light') footpoint, and the sites of strong non-thermal bremsstrahlung emission. It is to be expected that the Solar Dynamics Observatory will also reveal the detailed morphological evolution of ribbons, and possibly give some insight also into the vertical structure of the bright ribbon sources.

Information from imaging:
It is well established that solar flare footpoints are very compact and evolve rapidly in position and intensity, reflecting in some detail the progression of flare energy release. Flare chromospheric ribbons are much more extended. Interpreted as the chromospheric intersections of evolving magnetic (quasi-) separatrix surfaces, flare ribbons allow the progression of field restructuring to be followed (Masson et al. 2009). Chromospheric flares have of course long been beautifully observed in H alpha, but also now in other lines such as the 304nm image from Hinode SDO (Figure 1). These show abundant fine detail from 3D structures. Flare ribbons give insight into both energetics and magnetic restructuring, in the impulsive and the gradual phase.

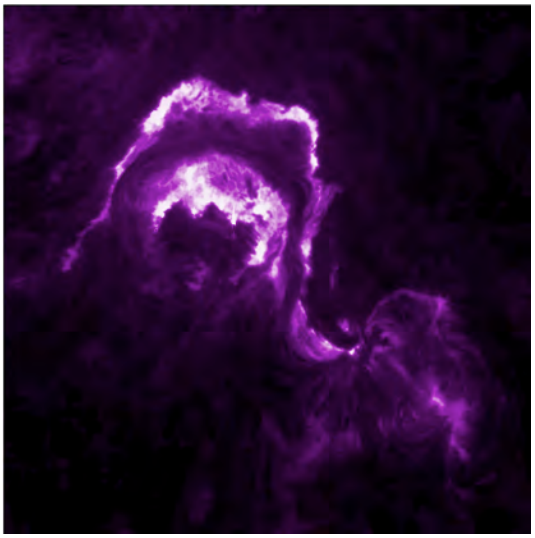


Figure 1 – chromospheric flare ribbons observed with the Hinode AIA in the 304nm channel (He II)

Compared to chromospheric flare ribbons, the optical flare footpoints are very little studied in the recent era. However, it is known that a very precise timing and spatial relationship to the hard X-rays (Fig 2) and thus the flare electrons – the major energy-carrying component (Metcalf et al. 2003, Hudson et al 2006, Fletcher et al. 2007, Watanabe et al 2009). Relatively easily observed compared to hard X-rays, they provide a view of the total energy of a flare complementary to that provided by HXR radiation. White-light emission had previously been thought to be a 'big flare' phenomenon, but has now been seen even in small C-class events (Hudson et al 2006, Jess et al. 2009).

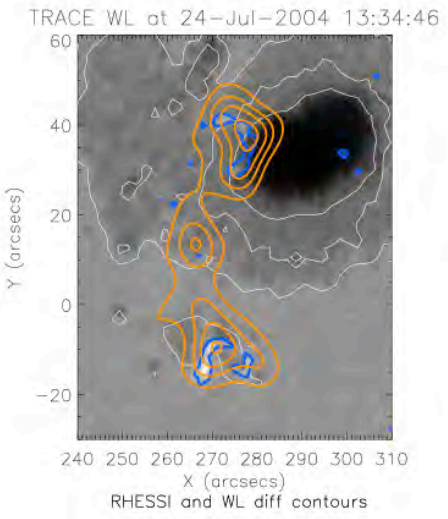


Figure 2: RHESSI 25-50 keV contours of a flare (orange) and regions of strongest WL contrast (blue). There is an excellent spatial and temporal correlation, but the WL readily shows detail at the arcsecond level which remains difficult to obtain with HXRs. Image from Fletcher et al. 2007

Information from spectroscopy:
During a flare, the chromosphere brightens in lines and continuum. At its most basic, continuum spectroscopic information is required for a proper assessment of the total radiative intensity in flares, which has so far been obtained only in a small number of events. Of particular interest is the identification (Neidig, 1989) of a Paschen jump (indicating the presence of the Paschen continuum from free-bound emission in the optical part of the spectrum) and flare-related changes to the Balmer and Paschen jumps in the flare spectrum, indicating variations in ionization fraction). Enhanced continuum shortwards of the Paschen jump may additional suggest increased Hα opacity, and photospheric excitation. This basic measurement has been made in very few flares (see e.g. Neidig 1983, Mauas 1990), and usually not for the strongest impulsive-phase optical flare kernels, but has major implications for the energy transport model (Canfield 1986, Neidig 1989, Fletcher et al 2007).

The chromospheric lines are potentially rich in diagnostic information. For example, Metcalf et al (1990a,b) have used Mg I lines to probe the temperature and density structure of the temperature minimum region. The Ca 8542 infrared triplet has been frequently used in stellar studies to diagnose the chromospheric temperature gradient, and the Balmer lines provide information on electron density (e.g. Svestka, 1976), and potentially also diagnostics of non-thermal electron distributions (Zharkova & Kobylinskii 1993, Kasparova & Heinzel 2002.) It has also been suggested that Lyα can also be used to detect flare proton fluxes in the 10-1000 keV energy range, with charge exchange on ambient neutrals resulting in shifted emission lines (Orrall & Zirker 1976). Never detected on the Sun, but known analogously in the terrestrial aurora, this would be a unique diagnostic of a major flare energy component as yet entirely invisible to us. The Balmer series line widths diagnose the electron density (via Stark broadening, e.g. Suemoto & Hiei 1959, Johns-Krull et al. 1997), and analysis implies lines imply that these flare lines are formed at electron densities in excess of 10^{13} cm^{-3} (e.g. Johns-Krull et al 1997), meaning within a transient condensation. From spectroscopy we can also image plasma flows, extended to 3D with modeling.

The poverty of solar flare chromospheric spectra:
It is an embarrassment that there is still no imaging spectroscopy of solar flares from space, although the IRIS instrument will soon be available. There is also precious little in the way of ground-based spectra – regular measurements have not been made since the 1980s and early 90s. HMI on SDO is now providing imaging spectroscopy of a single photospheric line, Fe I 6173 A, and the EVE instrument provides full EUV spectra but of the Sun as a star. In the meanwhile, Figure 3 suggests the power of what may emerge from better spectroscopy: this is a stellar flare spectrum dominated by the Balmer series, but with white-light continuum (and photospheric lines) in the background (Kowalski et al. 2010).

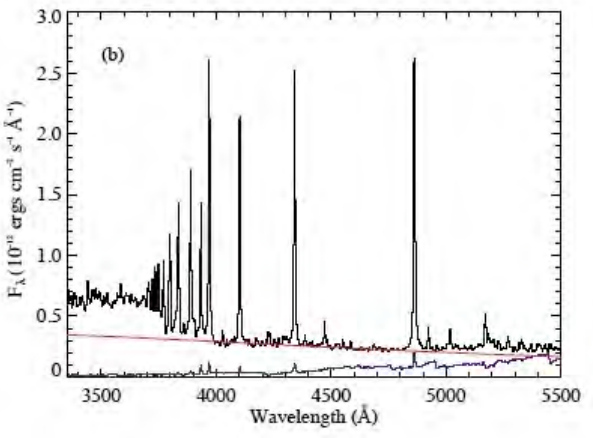


Fig. 3 The blue spectrum of the flare star YZ CMi (Kowalski et al. 2010). The red line indicates a hot blackbody component; most of the rest of the flare excess emission is the Balmer series plus the Balmer and Paschen continuum. The blue line is the quiescent star.

We do not have anywhere near compara-

comparable spectra for solar flares. Optical spectroscopic observations of a flare on the dM4.5e star YZ CMi have shown that two components consisting of a Balmer continuum and a hot (T~10,000 K) blackbody contribute to the total white light continuum in M dwarf flares (Kowalski et al 2010, Figure 3). These components may have similar solar counterparts, but unfortunately, very few broad-coverage spectra of the Sun during flares are available to compare with stellar spectra. The few that exist date back to 1970-80's (Machado & Rust 1974, Neidig 1983, Donati-Falchi et al 1984). None of these data can currently be used to determine what emission type(s) makes up white light on the Sun.

In lieu of taking broad wavelength-coverage solar flare spectra due to a lack of instrumentation, a common set of narrow-band continuum filters have been developed to directly compare solar and stellar white light. These filters are centered at 3500A and 4170A, thereby isolating continuum from line emission on both sides of the Balmer jump. The same set of filters is employed on the stellar instrument, ULTRACAM on the New Technology Telescope at La Silla, and on a solar instrument, ROSA at the Dunn Solar Telescope. The characteristics of the flare emission in these filters can be used to compare the spectral components in the white light continuum. A critical comparison of the flare continuum in the optical and near-UV are needed to determine if the same disagreement exists between solar flare chromospheric models (e.g., Allred et al 2005) and observations.

Implications for flare transport and acceleration:
Line spectra give information about the physical conditions at the location where the line is excited, and can thus be used to understand how energy propagates through the chromosphere. At present, our understanding of solar flare particle acceleration rests on the collisional thick-target interpretation of a particle beam accelerated in the corona and propagating to the chromosphere where it loses energy monotonically with depth, due to Coulomb collisions. This makes firm predictions about the location of beam stopping, and chromospheric heating, ionization and line excitation. Systematic optical observations can confront this, and other, models. Line and continuum observations so far imply significant energy deposited in dense layers of the chromosphere, and while the total energy requirements can be met by the collisional thick target inferences (e.g. Watanabe et al 2010), in some cases the implication that the energy is concentrated at electron energies which cannot penetrate sufficiently deep in the chromosphere (Fletcher et al. 2007).

Plasma physics:
The physics of the chromosphere differs from that of the corona and photosphere because of its low level of ionization and the dominance of ion-neutral coupling at lower altitudes. The Earth's ionosphere has some points of analogy. The current systems that support coronal energy storage as non-potential structures must flow through the chromosphere, resulting in a much more complicated pattern of electrical conductivity than common modeling approaches (e.g. "line-tying") can comprehend. It would be a fair criticism that most flare modeling at present simply ignores this essential physics, as well as other factors likely to be important in chromospheric and transition-region dynamics.

Seismic waves and Moreton waves:
Large-scale wave structures originate in the lower solar atmosphere during flares. These include the global coronal waves responsible for radio type II bursts and Moreton waves. The seismic waves in the solar interior (Kosovichev and Zharkova 1998) now have been shown to be relatively commonplace (e.g. Donea & Lindsey 2005). Each of the global wave types offers powerful diagnostic information regarding the early development of the flare/CME process.

New observations needed:
Flare research in the fundamental domain of the chromosphere and transition region has only begun to advance beyond the classical Hα stage. The text box summarizes the key observational goals for a new era in flare research. As Hinode and SDO demonstrate, space-based observations offer tremendous advantages in image stability, image dynamic range, and seeing. For the chromosphere and transition region, space also offers access to the key flare radiations at short wavelengths: spectra such as those of Figure 1 suggest temperatures above 10^4 K for the recombination-radiation component, the dominant flare luminosity.

We also recognize that ground-based facilities such as ATST will exist and produce key observations, and that ALMA (Karlicky et al. 2010) will bring an entirely new perspective of potentially great power.

Technical Requirements

• Imaging at high spatial (0.1'') and temporal (1 s) resolution in key optical and UV lines and continua, and with sufficient image dynamic range for flares.

• Lyα wing observations of charge-exchange signatures.

• Imaging spectroscopy of the chromosphere and transition region with full spectral coverage in optical-UV

• Modeling outside the "hydro-magnetic box": e.g., two-fluid, kinetic non-equilibrium, 3D radiation hydrodynamics etc.

References:

Aschwanden, M.J., Winebarger, A., Tsiklauri, D., and Peter, H. 2007, ApJ 659, 1673
Allred, J.C. et al 2006, ApJ 644, 484
Allred, J.C. et al 2005 ApJ 630, 573
Brown, J. C., Turkmani, R., Kontar, E. P., MacKinnon, A. L, Vlahos, L., 2009 A&A 508, 993
Berlicki, A. et al., 2008 A&A, 490, 315
Canfield, R. C., Bely-Dubau, F., Brown, J. C. et al., 1986, NASA Conf. Publ., NASA CP-2439, Ch 2.
De Pontieu et al. 2007, Science 318, 1574
De Rosa, M.L et al. 2009, ApJ 696, 1780
Donati-Falchi, A. et al 1984 A&A 131, 256
Donea, A. and Lindsey, C. 2005, ApJ 630, 1168
Eason, E.L.E. et al 1992, AJ 104, 1161
Fletcher, L., Hannah, I. G., Hudson, H. S., Metcalf, T. R., 2007, ApJ 656, 1187.
Fletcher, L., Hudson, H. S., 2008, ApJ 675, 1645
Fuhrmeister, B. et al 2008, A&A 487, 293
Hawley, S.L. & Pettersen, B.R. 1991, ApJ 378, 725
Hawley, S.L. et al 2003, ApJ 597, 535
Hudson, H. S., Wolfson, C. J., Metcalf, T. R. 2006, Sol. Phys. 234, 79
Jess, D. B., Mathioudakis, M., Crockett, P. J., Keenan, F.P. 2008, AoJ 688, L119
Johns-Krull, C. M., Hawley, S. L., Basri, G., Valenti, J. A., 1997, ApJS 112,221
Karlicky, M., Dabrowski, B., Barta, M., and Heinzel, P. 2010, Solar Phys. (in press)
Kasparova, J., Varady, M., Heinzel, P., Karlicky, M., Moravec, F. 2009, A& 499, 923

Kosovichev, A., and Zharkova, V. 1998, Nature 393, 317
Kowalski, A.F., Hawley, S.L., Holtzman, J.A., Wisniewski, J.P., and Hilton, E.J., 2010, ApJL 714, 98.
Machado, M.E. & Rust, D.M. 1974 Sol. Phys. 38, 499
Masson, S., Pariat, E., Aulanier, G., Schrijver, C. J. 2009, ApJ 700, 559
Mauas, P.,1990. ApJS 74, 609
Metcalf, T. R., Canfield, R. C., Avrett, E. H., Metcalf, F. T., 1990a, ApJ 350, 463
Metcalf, T. R., Canfield, R. C., Saba, J. R., 1990b, ApJ 365, 391
Metcalf, T. R., Alexander, D., Hudson, H. S., Longcope, D. W. 2003, ApJ 595, 483
Neidig, D. F., 1983, Sol. Phys. 85, 285
Neidig, D. F., 1989, Sol. Phys. 121, 261
Orrall. F. Q., Zirker, J. B.,1976 ApJ 208, 618.
Suemoto, Z. Hiei, E.,1959 PASJ 11, 185
Svestka, Z. F., 1976, *Solar Flares*
Watanabe, K., Krucker, S., Hudson, H. et al, 2010 ApJ 715,651
Woods, T. N., Eparvier, F. G., Fontenla, J., 2004, GRL 3110802
Zharkova, V.V., Kobylnskii, V. A., 1993, Sol Phys 143, 259.
Zirin, H., and Tanaka, K. 1973, Sol Phys 32, 173

A Science White Paper in response to the 2013-2022 Heliophysics Decadal Survey

Fundamental heliophysics processes

Unsteady wandering magnetic field lines, turbulence, magnetic reconnection, and flux ropes

T. Intrator (intrator@lanl.gov), J. Sears, T. Weber (P-24), J. Birn (ISR-1), Los Alamos National Laboratory
W. Gekelman, Department of Physics and Astronomy, UCLA
R.B. Dahlburg, Naval Research Laboratory
G. Lapenta, Katholieke Universiteit Leuven, Belgium

1. abstract

Magnetic flux tubes or flux ropes are the building blocks of magnetohydrodynamics (MHD). The interactions of flux ropes are key to understanding the unsteady dynamics of energy conversion, both magnetic to kinetic (reconnection) and kinetic to magnetic (dynamo). In nature, much of the basic MHD physics of reconnection, turbulence, dynamo, and magnetic topology evolution is irreducibly three dimensional (3D). Very little is known about the ramifications of this fact. We argue that much can be learned from *experimental* work in MHD relevant 3D systems. It is possible to scale dimensionless parameters to distill the essential physics, and for example assist data interpretation from existing and future spacecraft, as well as unravel the fundamental processes. There exist two linear (i.e. not toroidal) experimental devices, with fully 3D freedoms to evolve the plasma physics, that can tackle this problem: the Reconnection Scaling Experiment (RSX) at Los Alamos and BaPSF at Univ. California – Los Angeles.

2. Science objective

Explore fundamental 3D interactions of the flux rope building blocks of MHD

3. Introduction

3.1. Magnetic structure in 3D

Magnetic field is a unifying feature of heliophysics. These field lines occupy cross sectional areas, and fill up three dimensional (3D) volumes as flux tubes, and obey Newtonian mechanics and magneto-hydro-dynamics (MHD) equations of motion. If there is a vector component of the current parallel to the magnetic field, then the flux tube is a twisted flux rope. The rope geometry can be ubiquitous even in laminar sheets, since multiple reconnections in the inherently unstable current sheet create flux ropes (magnetic islands in 2D), and are often observed on the sun and the rest of the heliosphere. There are four 3D, topologically possible, classes of flux tube interactions [Linton2001]: (1) bounce (no appreciable reconnection), (2) merge, (3) slingshot (the most efficient reconnection), and (4) tunnel (a double reconnection).

Understanding the unsteady dynamics of flux ropes and their mutual interactions offers the key to many important solar phenomena including magnetic reconnection and turbulence, which cross the traditional heliophysics subdomains of the Sun, heliosphere, and magnetosphere. Observations [Frazier1972] show that the magnetic field of the Sun occurs in flux-tube bundles in the solar atmosphere, so flare models commonly invoke flux-tube collision as the trigger for flares. Tangled, unsteady magnetic field lines and flux ropes are ubiquitous, for example on the

sun from differential solar rotation, in the corona, solar wind [Borovsky2010], and the magnetotail. 3D isolated X-lines, kink dynamics and bursty driven dissipative reconnection processes may be related to reconnecting bursty bulk flows in the plasma sheet, solar corona and magnetotail turbulence. Astrophysical turbulence and 3D statistical mechanics of wandering flux ropes [Lazarian1999] are likely related.

3.2. Reconnection in 3D

For more than half a century it has been realized that processes associated with magnetic reconnection [Biskamp2000] are key to global changes magnetic topology in astrophysical [Kulsrud2005], solar [Priest2000], magnetosphere [Mozer2003] and laboratory [Yamada1999] plasma environments. Energy stored in stressed magnetic fields can produce large scale explosive events, that spontaneously evolve due to unsteady and impulsive [Bhattacharjee2004] local dissipative processes to energize particles. Explanations of satellite observed fast reconnection rates require patchy [Galeev1986] [Yin2008] yet volume filling processes, that can only exist in three dimensions (3D) [Titov1999]. At a fundamental level, reconnection generated current sheets are inherently prone to 3D instabilities, which creates a hierarchy of successively smaller spatial scales, that are generally not accounted for with 2D theory or simulations.

3.3. Questions cf reconnection in the heliosphere-magnetosphere

Does magnetic topology have a crucial role, and how does it affect the physics of reconnection? 2D models have been important to guide our physical picture of magnetic reconnection. On the other hand extension of these concepts to 3D is far from straightforward. Indeed, 3D magnetic field structure can be quite complex.

- How does 3D reconnection work: unsteady onset, evolution, and termination?

- How, when, and where is reconnection initiated?
- Can we identify typical situations for the onset of reconnection? Non driven examples?
- Is there a usually a dominant mechanism such as increased current density, decreased magnetic field, or increased magnetic shear, 3D localized resisitivity?
- How does reconnection terminate?
- Does back reaction of piled up magnetic and kinetic pressure stagnate the processes?
- Can we quantify the role of flow fields?
- Is there a topologically plausible continuum of possibilities spanning anti-parallel reconnection as observed at low latitudes in the presence of northward Interplanetary Magnetic Field (IMF) to component reconnection as observed at higher latitudes with southward IMF?

4. Value to satellite missions

4.1. Multi spacecraft data

Unraveling these major unresolved heliosphere and magnetosphere reconnection physics puzzles is a primary scientific goal of the 4 spacecraft Magnetospheric Multiscale (MMS) mission [MMS2009], a Solar-Terrestrial Probe. MMS has been designed with diagnostic capabilities to investigate transient, unsteady and 3D magnetic, electrostatic, and flow structures. At any given time, four data points in space are collected. This is sufficient to infer many features, but important spatial derivative-like quantities such as curl, divergence, vortices, or simultaneous small and large scale correlations are difficult to evaluate unambiguously. Distinguishing

between temporal and spatial variations has been made possible with coarse grained resolution with clusters of similarly instrumented spacecraft such as CLUSTER and THEMIS, and soon MMS. However in the laboratory this distinction can be measured with high accuracy in both time and space.

What types of 3D magnetic structures will the MMS mission likely observe? A 3D magnetic null configuration has been statistically inferred from CLUSTER data [Xiao2006], but the evidence is not detailed in 3D, and is far from solid. What happens when there is a guide field, e.g. in many typical cases including smaller spatial scales and turbulence, and a null cannot exist? What would the spacecraft see when there are stagnation flows, compressible effects, or local gradients? Collaborating spacecraft observational scientists will facilitate our interpretation of laboratory experiments, which in concert with magnetosphere data and simulations will provide our *only opportunity to nail down the answers* to these questions, and will increase the credibility of MMS (and other) data interpretation.

4.2. Reconcile spacecraft observations to experiment & simulation

By matching or scaling dimensionless parameters, with appropriate initial and boundary conditions, some natural phenomena can be studied on laboratory scales. Table 1 shows considerable laboratory overlap with MMS capabilities:

Diagnostic & Observation	**MMS comment**	**RSX resolution**	**RSX**	**VPIC resolution**
Electric field	yes	$< c/\omega_{pe}$	Triple probe, E=-dA/dt	$\approx \lambda_{De}$
Magnetic field	Yes	$< c/\omega_{pe}$	Magnetic probes	$\approx \lambda_{De}$
ion distributions	excellent	$< 4c/\omega_{pe}$	Energy analyzer	$\approx \lambda_{De}$
Ion speed	excellent	$< 4c/\omega_{pe}$	Tomography of chord averaged doppler shift H-beta	$\approx \lambda_{De}$
Ion temperature	excellent	$< 4c/\omega_{pe}$	Tomography of chord averaged thermal width H-beta	$\approx \lambda_{De}$
Electron currents	yes		J-env_{ion}	$\approx \lambda_{De}$
Wave disturbances	E, B, n	E, B, n $< c/\omega_{pe}$, $>\omega_{ce}$	Experimental probes, hodograms, data acquisition & bandwidth can be fast or slow, conditional sampling	$\approx \lambda_{De}$
Global angular momentum	Not obvious	$< c/\omega_{pi}$	magnetics, fast camera images, flows	Yes
Global 3D structure:	Not obvious	$< c/\omega_{pi}$	magnetics, fast camera images	yes
Number of spatially distinct data pts	4	$< c/\omega_{pe}$	$2\text{-}10x10^3$	grid cells > 10^9
Anti-parallel vs component reconnection	Depends on spacecaft location		adjustable	adjustable

Table 1. Survey of plasma characteristics accessible via spacecraft measurements, with their analogues in the RSX experiment and VPIC simulations. Sme features are tough to measure in space, but accessible in the laboratory and simulations.

We assume approximate magnetospheric environment parameters of B≈10 Gammas = 10^{-4} Gauss, density n≈1cm^{-3}, ion mass ≈ 1amu, where MMS cluster satellite separation is designed to be 10km or greater. The important scale lengths that must be considered include the Debye length which can be in the range λ_{De} ≈ 100m (possibly boom extension length of a single spacecraft), electron skin depth c/ω_{pe} ≈ 500km (resolvable with multiple spacecraft), electron gyroradius r_{Ge} ≈ 2-10km (possibly resolvable with multiple spacecraft). MMS burst time resolution can be better than 1msec, which corresponds to better than the electron cyclotron frequency f_{ce} ≈ 300Hz. Time resolution is better in the laboratory, even in dimensionless units.

5. Usefulness of experiment – theory – computational collaborations

5.1. Sparse experimental work on 3D flux rope dynamics, reconnection, turbulence

To our knowledge there is little or no laboratory work on possible links between reconnection and turbulence, and very little on the dynamics of intermittent, unsteady reconnection.

Many relevant helio, astro, and space plasmas are in the MHD regime, and therefore "frozen in" magnetic field lines serve as a proxy for flow. Magnetic field line locations and movements are much easier to measure in 3D in the laboratory than flow, and one can take advantage of experimental magnetic measurements to study magnetic field line, using flux ropes as proxies … to unravel macro and micro flow patterns and compare these with computational simulations. 3D magnetic field data allows us to find the B field line paths, and also to integrate along field lines global invariants such as entropy. The obvious exception is during magnetic reconnection inside the small diffusion region, where field lines lose their identity. A straightforward method to experimentally in RSX [Intrator2009] and LAPD [Lawrence2009] quantify these 3D quasi separatrix layers [Milano1999] is to measure magnetic field lines up to the boundaries of the diffusion region, and then calculate the slip-squash characteristics [Titov2009].

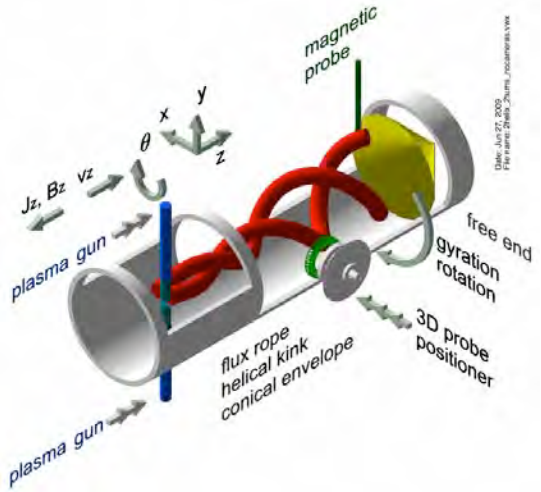


Fig. 1 RSX experiment schematic (2m long x 40cm radius), showing coordinate system, geometry, background axial magnetic field B_z (single arrows). Two plasma guns (double arrows) insert radially but create axial flux rope currents and net helical magnetic field. Two kinking flux ropes terminate at the external conical anode (triple arrow) which allows adjustable axial boundary conditions. Magnetic probes are inserted through 3D probe positioners. Radial excursion from the reference fiducial **z** axis exaggerated for clarity.

The RSX (Fig. 1) experiment has yielded 3D data that show the onset, stagnation, and cessation of magnetic reconnection [Sun2010] [Intrator2009] and also unexpected behavior of flux ropes [Furno2006] [Sun2008] [Sun2010]. The LAPD facility has demonstrated that 2D concepts such as the reconnection current sheet are better explained by 3D signatures such as the quasi separatrix layer [Lawrence2009]. This requires large 3D data sets, which complement spacecraft data with limited numbers of spatially separated data points for each time step.

6. Experiments + simulations connect micro to macro physics

Computational models must couple kinetic micro-physics of particles, waves and momentum transfer to macroscopic characteristics. Our collaborating magnetosphere observers and theorists can advise on reconciling laboratory and spacecraft data. By matching or scaling dimensionless parameters, computer codes can bridge the gap between nature and laboratory, while enabling benchmarks plus verification & validation for code predictions.

6.1. Experimental and computational resources

LANL has world class and unique capabilities for investigating 3D reconnection that is spontaneous or self driven by instabilities, using VPIC, a particle in cell code running at LANL on Roadrunner, using unprecedentedly large numbers of grid points, particles, and petaflop speed. The kinetic micro-physics of particles, waves and momentum transfer to macroscopic characteristics can be coupled via fluid simulations with the bulk properties. Also G. Lapenta has modeled RSX at Katholieke Universiteit, Belgium, using 3D particle code CELESTE-3D and variants and 3D fluid codes that he refined while at LANL. Massage and 3D imaging of huge 3D datasets has been carried out with LAPD data.

The unique Reconnection Scaling Experiment (RSX) [Furno2003] at Los Alamos National Laboratory showed for the first time [Intrator 2009] a plasma instability that initiated reconnection, with subsequent saturation or stagnation related to the fluid-like forces, mass, flows, and dynamics of the participating plasma and its current system. The electron diffusion region was experimentally resolved to electron skin depth scales. The Basic Plasma Science Facility performs frontier-level research on fundamental properties of the plasma state of matter. Usage of the facility is available to qualified scientists from national and international institutions, and industry.

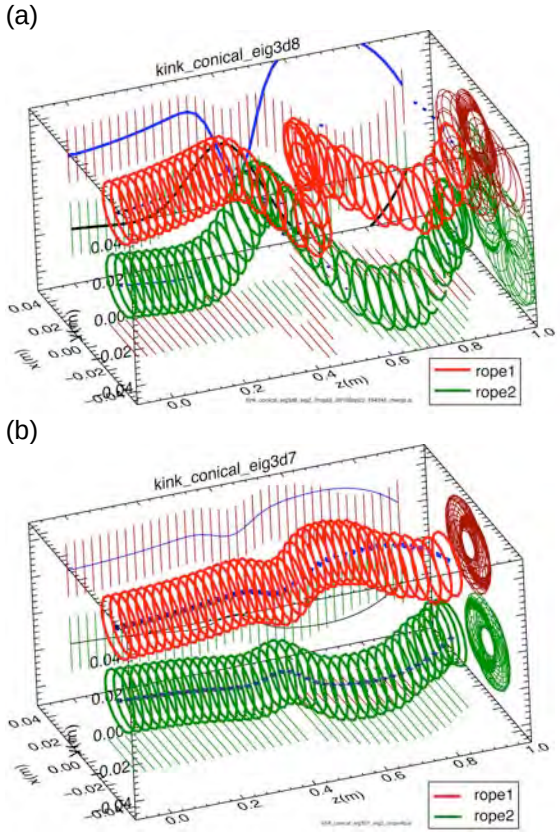


Fig. 2. RSX 3D magnetic data with model reconstruction. Flux rope building blocks of MHD dynamics include a variety of topologies, statistical mechanics, and final states not usually accounted for in the usual MHD picture. For example (a) Kinked flux ropes gyrate, parallel currents attract which yields expected merging/reconnection [Intrator2009]. (b) B_z field line bending forces restrain merging and flux ropes bounce! [Sun2010].

7. Experiment descriptions

7.1. RSX

The RSX (wsx.lanl.gov) schematic in Fig. 1 includes a view of two plasma guns located at z=0 which are radially inserted into the center of the RSX cylindrical vacuum vessel. This forms two cylindrical plasma columns in the externally imposed constant, uniform, axial magnetic field of $B_{z0} \approx$ 10-1000 Gauss. A screw pinch is formed by driving a current in the plasma to an external anode [Furno2003]. When two flux ropes are created, their parallel currents (I_p mostly in the z

direction) attract each other. The plasma is generated by plasma gun arc discharge, independent of initial conditions or evolution, without recourse to any complicated startup scenario typical of tokamaks, spheromak like coaxial guns [Bellan2000], or flux cores [Yamada1997]. This allows a wide independently adjustable range [Furno2003] for density, current density (skin depth), embedded magnetic field (gyro orbit size), and electron temperature (collisionality).

A special vacuum vessel section containing six 3D probe positioners allows many independent insertable probes to explore a 3D volume (dx•dy•dz = 15cm•10cm•30cm) with spatial relative accuracy better than 0.5mm in an x-y-z Cartesian coordinate system [Intrator2008]. This allows experimental resolvution to spatial scales where reconnection breaks MHD assumptions, i.e. electron skin depth which is typically 1-2mm in size. Fig. 2 shows two examples of theoretical reconstructions constrained by experimental data that show (a) merging and reconnecting flux ropes and (b) unexpected bouncing flux ropes. Probe data was built up over many repeatable shots, where we measured plasma density, electron temperature, and electron pressure on sub mm spatial scales.

7.2. BaPSF-LAPD

The Basic Plasma Science Facility (http://plasma.physics.ucla.edu) performs frontier-level research of the plasma state of matter. Usage of the facility is available to qualified scientists from national and international institutions, and industry. Research on the fundamental properties of plasmas aids our understanding of applications ranging from fusion energy to space science. It is a user facility operated at the University of California, Los Angeles. The core of the facility is a modern, large plasma device (the LAPD, Fig. 3) constructed by Walter Gekelman (the facility director) and his staff of research scientists and technicians. The publication profile is impressive (6 PRLs so far in 2010). The facility provides an environment in which teams with complementary expertise (e.g., Laser Induced Fluorescence or High Power RF) can come together to attack problems that they would not pursue individually. The operational procedures foster the exchange of technical information across diverse areas of research (e.g., fusion studies, space investigations, laser-plasma interactions, plasma applications) in which the basic properties of plasmas play an essential role.



Fig. 3. LAPD experimental device at UCLA

The facility is available to scientists from all institutions, national and international, with access to state-of-the-art hardware and a broad range of plasma conditions in which to exert their creativity. Through cooperative research programs involving researchers from large and small institutions, the BaPSF aims to yield major advances on outstanding problems related to the behavior of plasmas, and contribute significantly to the training of the next generation of plasma researchers. A much larger facility (ET, Enormous Toroidal experiment) is also being explored.

8. Relevance to the Decadal survey and Estimated cost

There is no steward organization for earth based *experimental* basic plasma physics capability, which is almost extinct in the present NASA program, and is primarily supported by NSF and DOE. Both experimental efforts (RSX, BaPSF) have large and vibrant student programs.

Multi point measurements in space and time, indeed massive 3D data sets, complement the spacecraft multi point (e.g. 4 spatial points per time step) capabilities. We make the case that a small investment represents a logical next step for Helio and magnetosphere physics, with huge bang for the buck. Our objectives extend across several of the National Capabilities Working Groups; namely, 'Research to Operations', 'Theory & Modeling', 'Explorers', and 'Innovations'. The mission combines both operational and research objectives and therefore has both societal benefits and advances the state of knowledge for the Sun-Earth system.

The cost for fielding experimental work on earth is far less than a spacecraft mission, but more than an observer post doc with a computer. Expenses are typically post doc, student FTE, 15-20% allowance for materials and supplies, small fraction of PI and technician FTE, and some travel and publication expenses. Significant progress can be made for less than $300-500k/year per experiment.

9. References

Bellan, P.M. *Spheromak: A practical applicationof magnetohydrodynamic dynamos and plasma self organization*, (Imperial College Press, London, 2000).

Bhattacharjee, A. Impulsive magnetic reconnection in the Earth's magnetotail and the solar corona. *Annu. Rev. Astron. Astrophys* **42**, 365-384 (2004).

Biskamp, D. *Magnetic Reconnection in Plasmas*, (Cambridge University Press, Cambridge, 2000).

Borovsky, J.E. Contribution of Strong Discontinuities to the Power Spectrum of the Solar Wind. *Physical Review Letters* **105**, 111102 (2010).

Frazier, E. & Stenflo, J. ON THE SMALL-SCALE STRUCTURE OF SOLAR MAGNETIC FIELDS. *Solar Physics* **27**, 330-346 (1972).

Furno, I., *et al.* Reconnection scaling experiment: a new device for three-dimensional magnetic reconnection studies. *Review of Scientific Instruments* **74**, 2324-2331 (2003).

Furno, I., *et al.* Current-driven rotating-kink mode in a plasma column with a non-line-tied free end. *Physical Review Letters* **97**, 015002-015004 (2006).

Galeev, A.A., Kuznetsova, M.M. & Zeleny, L.M. Magnetopause stability threshold for patchy reconnection. *Space Science Reviews* **44**, 1-41 (1986).

Intrator, T.P., Sun, X., Dorf, L., Furno, I. & Lapenta, G. A three dimensional probe positioner. *Review of Scientific Instruments* **79**, 10F129 (2008).

Intrator, T.P., Sun, X., Dorf, L., Lapenta, G. & Furno, I. Experimental onset threshold and magnetic pressure pileup for 3D reconnection. *Nature-Physics* **5**, 521-526 (2009).

Kulsrud, R.M. *Plasma physics for astrophysics*, (Princeton Univ. Press, Princeton, NJ, 2005).

Lawrence, E.E. & Gekelman, W. Identification of a Quasiseparatrix Layer in a Reconnecting Laboratory Magnetoplasma. Physical Review Letters 103, 105002 (2009).

Lazarian, A. & Vishniac, E.T. Reconnection in a weakly stochastic field. *Astrophysical Journal* **517**, 700-718 (1999).

Linton, M.G., Dahlburg, R.B. & Antiochos, S.K. Reconnection of twisted flux tubes as a function of contact angle. *Astrophysical Journal* **553**, 905-921 (2001).

Milano, L.J., Dmitruk, P., Mandrini, C.H. & Gomez, D.O. Quasi separatrix layers in a reduced magnetohydrodynamic model of a coronal loop. *Astrophysical Journal* **521**, 889-897 (1999).

MMS. Magnetosphere Multiscale Science. in *http://stp.gsfc.nasa.gov/missions/mms/mms_science.htm* (2009).

Mozer, F.S., Phan, T.D. & Bale, S.D. The complex structure of the reconnecting magnetopause. *Physics of Plasmas* **10**, 2480-2485 (2003).

Priest, E.R. & Forbes, T. *Magnetic Reconnection*, (Cambridge University Press, New York, 2000).

Sun, X., Intrator, T. & Lapenta, G. Flux Rope Dynamics: Experimental Study of Bouncing and Merging. *Physical Review Letters* **to appear**(2010).

Sun, X., Intrator, T., Dorf, L., Furno, I. & Lapenta, G. Transition of MHD kink stability properties between line-tied and non line tied boundary conditions. *Physical Review Letters* **100**, 205004 (2008).

Titov, V.S., Demoulin, P. Basic topology of twisted magnetic configurations in solar flares. Astron. Astrophys 351, 707-720 (1999).

Titov, V.S., T.G., F., Priest, E.R., Mikic, Z., Linker, J.A. Slip-Squashing Factors as a Measure Of Three-Dimensional Magnetic Reconnection. *Astrophysical Journal* **693**, 1029-1044 (2009).

Xiao, C.J., *et al.* In situ evidence for the structure of a magnetic null in a 3D reconnection event in the Earth's magnetotail. *Nature-Physics* **2**, 478-483 (2006).

Yamada, M. Review of controlled laboratory experiments on physics of magnetic reconnection. *Journal of Geophysical Research* **104**, 14529-14541 (1999).

Yamada, M., *et al.* Study of driven magnetic reconnection in a laboratory plasma. *Physics of Plasmas* **4**, 1936-1944 (1997).

Yin, L., *et al.* Three Dimensional Dynamics of Collisionless Magnetic Reconnection in Large-Scale Pair Plasmas. *Physical Review Letters* **101**, 125001 (2008).

The Effect of the Heliosphere on Galactic and Anomalous Cosmic Rays

Principal Author: Israel, Martin H.
Washington University in St. Louis
mhi@wustl.edu
314-935-6263

Co-Authors:		
	Leske, Richard A.	California Institute of Technology
	Binns, W. Robert	Washington University in St. Louis
	Christian, Eric	NASA-Goddard Space Flight Center
	Cummings, Alan C.	California Institute of Technology
	Labrador, Allan W.	California Institute of Technology
	Lave, Kelly A.	Washington University in St. Louis
	de Nolfo, Georgia A.	NASA-Goddard Space Flight Center
	von Rosenvinge, Tycho T.	NASA-Goddard Space Flight Center
	Wiedenbeck, Mark E.	Jet Propulsion Laboratory

Brief White Paper Description:
The interplanetary magnetic field modulates both Galactic and anomalous cosmic rays in the inner solar system. GCR and ACR observed during the recent solar minimum show that the modulation is not well understood. Observations over at least the next 22 years are needed to resolve this puzzle.

This White Paper addresses the following relevant panels:
Solar and Heliospheric Physics (SHP)

The interplanetary magnetic field (IMF) of the heliosphere has a major effect on the intensity of the Galactic cosmic rays (GCRs) in the inner solar system. The GCR intensity reaches a maximum during solar minimum, when the sunspot number and other indicators of solar activity are at a minimum and declines substantially when solar activity increases. This modulation of the GCR intensity is most pronounced at energies below a few GeV/nucleon, at energies near and below the peak in the GCR energy spectrum. Since the GCRs are a major component of the radiation environment to which astronauts and spacecraft are subjected, these variations in intensity have implications for both human and robotic space exploration.

Although this solar modulation of GCRs has been studied extensively over several solar cycles, the most recent (2008 – 2010) solar minimum has been surprisingly different from those previously studied. This minimum of solar activity has lasted longer than previous minima, and the intensity of low-energy GCRs in 2010 reached a level more than 20% higher than has ever been recorded since the beginning of the space age sixty years ago (Mewaldt, et al., 2010).

Figure 1 shows the intensity of GCR Oxygen over the past thirteen years. Also shown is the expected intensity predicted in 2007 extrapolated from the neutron monitor data to that date, based on the observations from previous solar minima. The difference between the projection and the actual data for the past three years demonstrates that the solar cycle and its effect on GCRs is still imperfectly understood. This figure also shows several parameters describing the interplanetary medium, each of which is likely to have some effect on the local GCR intensity. The relative effect of each of these parameters on the GCRs is still not well understood.

Since the sunspot number on the Sun reaches a maximum approximately every eleven years, the solar cycle is often thought of as being approximately eleven years in duration. In fact the cycle of solar modulation of GCRs is about twenty-two years. Every eleven years the polarity of the solar magnetic field reverses, and the reversed polarity, combined with the spiral shape of the interplanetary magnetic field, leads to a marked difference in GCR modulation between the two halves of this twenty-two year cycle. (Jokipii, Levy, & Hubbard., 1977; Jokipii & Thomas, 1981). If we are to untangle the various effects on local GCR intensity and understand the present “anomalous” solar minimum, it will be necessary to monitor the GCR intensity and the various interplanetary parameters that affect it for at least the next twenty-two years. In such studies, it is particularly useful to have long-term continuous data sets from the same instrument. The low-energy oxygen data shown in both figures of this white paper come from thirteen years of data from instruments on the ACE spacecraft, and we note that the consumables on ACE are adequate to continue the mission to ~2024. At higher energies long-term data sets from ground-level neutron monitors are also valuable.

A vital tool in developing an understanding of solar modulation of GCRs is comparison between measurements of GCRs near Earth at 1 AU with measurements in the outer heliosphere and beyond. When Voyager-1 crosses the heliopause there will be a once-in-a-lifetime opportunity to compare directly-measured GCR spectra in interstellar space

with spectra in the outer heliosphere measured by Voyager-2 and with spectra at 1 AU, measured by instruments on spacecraft like the Advanced Composition Explorer (ACE). In addition to tying down the major unknown in the modulation problem – the interstellar spectrum – we will also finally know the maximum GCR intensity to which Earth can be subjected in times like the Maunder minimum. Only then will it be possible to decode accurately the ^{10}Be cosmic-ray record in terrestrial samples.

In addition to GCRs, anomalous cosmic rays (ACRs) provide an important probe of solar modulation in the heliosphere. ACRs are produced when neutral atoms from the local interstellar medium drift into the heliosphere, where they are ionized and convected to the outer heliosphere where they are accelerated, possibly by diffusive shock acceleration at the termination shock. Since they are accelerated at the outermost parts of the heliosphere, one would expect that they would suffer the same modulation as they drift into 1 AU as do the GCRs.

As shown in Figure 2, although GCRs were at record-high intensities this solar minimum, ACRs at 1 AU were not (Leske et al., 2010). In fact they were quite typical, with peak intensities similar to those during the last solar minimum with A<0 polarity in the mid-1980's and below those of the A>0 minimum of ~11 years ago. Voyager measurements in the outer heliosphere show that the intensity of high-energy ACRs was the same this solar minimum as it was in the mid-1980's, suggesting that the source intensity of ACRs was unaffected by the changes in the interplanetary parameters shown in Figure 1.

During A<0 cycles ACRs drift into the inner heliosphere along the heliospheric current sheet (HCS) (Jokipii and Thomas, 1981), and as Figure 2 indicates there was a striking correlation between the HCS tilt angle, ACR intensities, and GCR (neutron monitor) intensities during the 1980's A<0 epoch. Prior to 2000, the major, prolonged deviations between the tilt angle and cosmic ray intensities occur during the approach to solar maximum from the two A>0 minima, when the neutron monitor and ACR rates declined only slowly as the tilt angle rapidly increased. This is as expected: since particles drift in from the polar regions of the heliosphere during A>0 periods, they are virtually unaffected by changes in the largely near-equatorial HCS, at least until the HCS reaches high latitudes. However, after the last field reversal in ~2000, the neutron monitor rate, ACR intensity, and HCS tilt angle no longer scale as they did for the previous three solar minima. For a given tilt angle, both GCR and ACR intensities are elevated compared with the last A<0 solar minimum, but GCRs are relatively more enhanced than ACRs. Although the long, deep solar minimum allowed for less cosmic-ray modulation, the relatively high HCS tilt angle partially compensated for this. If the ACR source distribution is concentrated at low latitudes during A<0 periods (Jokipii and Giacalone, 1998), the ACRs may be more sensitive to the high tilt angles than the GCRs were.

Discovering the puzzle that solar modulation can yield record high GCR intensities without affecting the ACRs required observations in both the inner and outer heliosphere across a broad range of particle energies throughout more than one full 22-year solar magnetic cycle. Resolving the puzzle, and testing any proposed resolution, will require that such observations continue for future solar cycles, and will likely better elucidate the

underlying physics of solar modulation, the origin of ACRs, and the interaction of our Sun and heliosphere with the local interstellar medium.

Observations required

As indicated above, an understanding of these heliospheric effects on the interplanetary radiation environment will require continuous observations of both the galactic and the anomalous cosmic rays over at least the next twenty-two-year solar cycle. Ground-level neutron monitors are helpful in this regard, but their mean response is to particles with energy ~10 GeV, while the bulk of the ionizing radiation is at lower energies. Thus particle detectors near 1 AU outside the Earth's magnetosphere (like ACE at L1) or in low-earth polar orbit or in highly eccentric earth orbit (like IMP-8) will be needed. Detectors on the ACE spacecraft have been serving this function superbly for the past thirteen years, and that spacecraft has consumables capable of extending those observations until ~2024. Continuation of the ACE mission as long as the spacecraft and its cosmic-ray instruments last will be the least expensive way of assuring continuation of this capability. Beyond the ACE mission, a relatively simple cosmic-ray telescope with capability of identifying individual elements could be placed on a polar orbiting spacecraft to provide a low-cost means of extending this capability to twenty-two years from now, and beyond. Such continuous monitoring of the cosmic-ray intensity will be vital for planning of future human and robotic exploration of our Solar System.

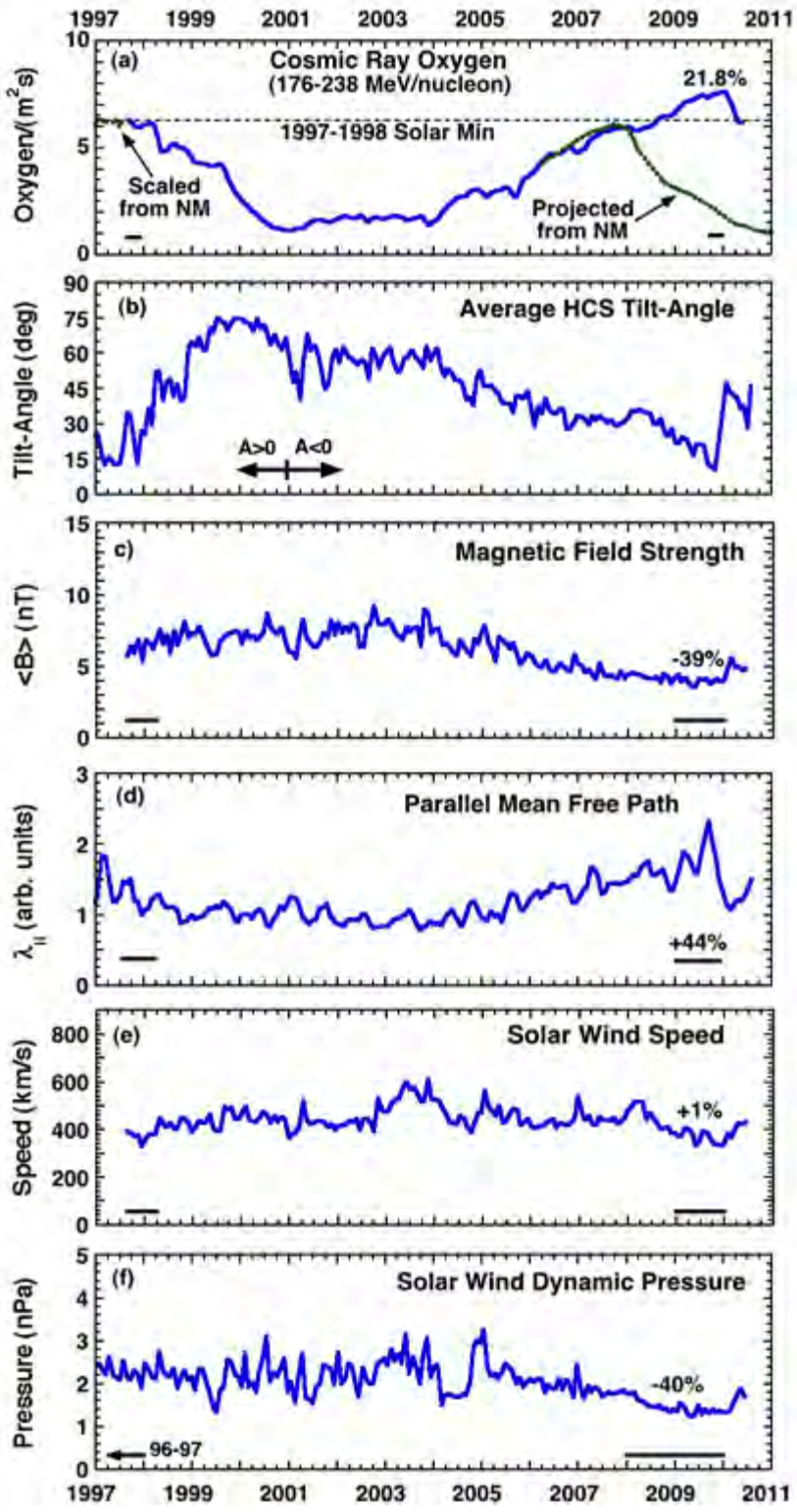


Figure 1

a) GCR oxygen intensity

b) Average tilt angle of the heliospheric current sheet

c) Mean strength of the interplanetary magnetic field.

d) Mean free path of particles parallel to the local magnetic field

e) Solar wind speed

f) Dynamic pressure of the solar wind

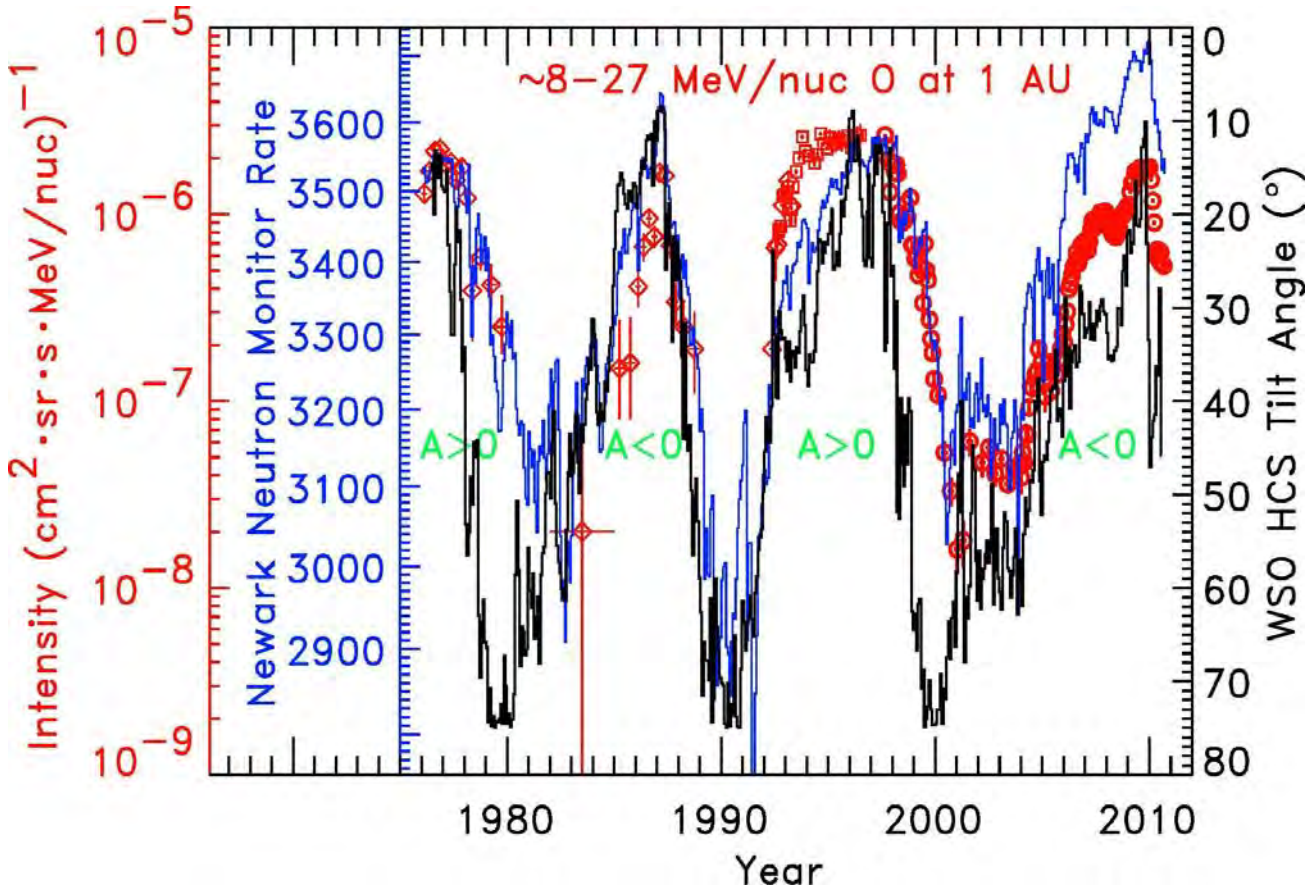


Figure 2: Comparison of the ACR intensity (red data points; first left axis) and the Newark neutron monitor count rate (blue curve; second left axis) with the tilt angle of the heliospheric current sheet from the Wilcox Solar Observatory "classic" line-of-sight model (http://wso.stanford.edu; thick black curve; right axis -- note inverted scale) over the past several solar cycles, from Leske et al. (2010).

References:

J.R. Jokipii, E.H. Levy, & W.B. Hubbard (1977) Effects of Particle Drift on Cosmic-ray Transport. I. General Properties, Application to Solar Modulation, *Astrophys. J.* **213**, 861.

J.R. Jokipii & B. Thomas (1981) Effects of Drift on the Transport of Cosmic Rays IV. Modulation by a Wavy Interplanetary Current Sheet", *Astrophys. J.* **243**, 1115.

J.R. Jokipii and J. Giacalone, (1998) The Theory of Anomalous Cosmic Rays, *Space Sci. Rev*. **83**, 123.

R.A. Leske, A.C. Cummings, R.A. Mewaldt, and E.C. Stone (2010) Anomalous and Galactic Cosmic Rays at 1 AU During the Present Solar Minimum, *Space Sci. Rev*., submitted.

R.A. Mewaldt, A.J. Davis, K.A. Lave, R.A. Leske, E.C. Stone, M.E. Wiedenbeck, W.R. Binns, E.R. Christian, A.C. Cummings, G.A. deNolfo, M.H. Israel, A.W. Labrador, & T.T. vonRosenvinge, (2010) Record-Setting Cosmic-Ray Intensities in 2009 and 2010. *Astrophys. J. Lett.* **723**, L1-L6.

“PERSEUS”, a Pegasus Explorer for remote sensing and in-situ space science

Bernard V. Jackson, Andrew Buffington, John M. Clover & P. Paul Hick
Center for Astrophysics and Space Sciences,
University of California, San Diego
La Jolla, CA, 92093-0424

Abstract

PERSEUS instruments work together to provide all sky coverage that enables mapping and 3D reconstruction of the global heliosphere. A coronagraph (CRG) observes rapidly-changing material flow and events (CME’s) near the solar surface while all-sky visible-light imagers (ASIs) observe light from these same solar wind flows and heliospheric structures as they pass through the inner solar system. The CRG fills the inner portion of the FOV 360º around the Sun, and the four-degree separation between ASI’s two camera’s views are placed so that the ecliptic plane is viewed completely out to the near-Earth vicinity. The views join near the ecliptic poles and overlap at elongations greater than 90º. The ASIs view heliospheric structures that pass Earth most efficiently near 90º elongation and greater allowing these structures to be reconstructed and compared with measurements from two in-situ instruments that observe plasma and magnetic fields. This concept was presented in 2007 as a Small Explorer Program, a joint collaboration between UCSD, NASA Ames, NASA Goddard, Lockheed Martin, Palo Alto, and Air Force Research Lab.

I. Introduction

NASA's "Vision for Space Exploration" charts a bold new course into the cosmos that will take humans back to the Moon and eventually to Mars and beyond. The present white paper describes instrumentation that will map this plasma environment globally, and will fill the gap in observations between the solar corona and throughout interplanetary space. Furthermore, the mission will maximize the scientific return from existing remote-sensing and in-situ measurements by adding crucial physics information about the evolution of material in the heliosphere between the Sun and the Earth and between current Earth-based space instrumentation.

In support of the above, PERSEUS (Pegasus Explorer for Remote SEnsing and in-sitU Space-science) places a four-sensor satellite in orbit about the L_1 Lagrange point. During its mission, it provides a continuous, all-sky observation and modeling of the heliosphere's fundamental plasma parameters of velocity and density. The remote-viewing sensors provide an order of magnitude greater signal-to-noise over current measurements. The sensor suite consists of:

1. An **All-Sky Imager (ASI)** consisting of two 180-degree field of view (FOV) cameras and optics to measure solar wind flow and brightness including Coronal Mass Ejections (CMEs) well away from the Sun (see Figure D-1);
2. One Solar **CoRonaGraph (CRG)** to measure these in the corona;
3. One **Solar Wind Ion Flux Telescope (SWIFT)** to measure the velocity, density, and temperature of positive ions as they stream over PERSEUS; and
4. One **MAGnetometer (MAG)** to measure solar wind magnetic properties.

The PERSEUS mission addresses the following NASA science objectives (from "*Sun-Earth Connection 2003 Roadmap to Understand the Sun, Heliosphere, and Planetary Environments as a Single Connected System"* and the decadal survey challenges in *"The Sun to the Earth and Beyond"):*

1. *"Understand the changing flow of energy, matter and magnetic field from the Sun and throughout the heliosphere and planetary environments."*
2. *"Explore the fundamental physical processes of solar and space plasma systems, and their interaction with solar system bodies."*
3. *"Define the origins and societal impacts of variability in the Sun-Earth connection."*

PERSEUS is an important and timely resource to the study of the solar corona and heliosphere. It is a natural successor to the Solar Mass Ejection Imager (SMEI), which is currently the only space borne photometric instrument viewing nearly the whole sky in white light. Current additional satellites making similar observations include STEREO A & B, each of which have a 70-degree FOV heliospheric imagers viewing in the direction between Sun and Earth (specifically aimed at observing the Sun-Earth line). Since these spacecraft are moving in opposite directions close to the Earth's orbital path, they will drift out of their optimal viewing positions in the next few years. Other solar-study spacecraft include SOHO, Ulysses, and Wind; these are already past their life expectancies and are into well-extended missions.
PERSEUS data enables high resolution 3D computer modeling (see Fig. D-2) of CME dynamics throughout the heliosphere. These models significantly improve our understanding

and prediction of solar wind structure, transport, and its impact on planetary systems and space assets. Moreover, PERSEUS's forecasting ability provides a critical early warning for sensitive space-borne assets.

PERSEUS in-situ observations must be made from deep space near Earth, and well outside the magnetosphere. Imager observations from L_1 are also well removed from near-Earth interferences such as the Earth shadow and its radiation bands and high-altitude aurora. The L_1 location thus provides a stable, benign environment to achieve PERSEUS objectives, and is essential for its in-situ measurements. Added advantages are that it is upwind of Earth (allowing up to an hour of advance warning) and proximity to other in-situ monitors near L_1 (allowing multipoint studies).

II. Science Data and Other Products

PERSEUS data consist of remote-view images from the ASI and CRG and in-situ plasma and magnetic-field measurements from SWIFT and MAG. The images consist of various header information, then a sequence of photometric readings over the camera's FOVs, cosmic rays removed and suitably rebinned onboard for telemetry to Earth. The latter are a time series of the two instrument's individual measurements. The plasma and magnetic-field measurements provide a "ground truth" measurement for the 3D reconstructions described below, and augment other such measurements from other existing spacecraft at L_1 and elsewhere.

PERSEUS ASI images and in-situ measurements will be unprecedented for completeness, clarity and measurement of two-dimensional flows. Solar wind physics above a few solar radii is dominated by plasma flow, and since both flow and brightness corresponding to dense and less dense material can be viewed directly many questions will be answered from analysis of these data. More importantly, these data will be used to provide input to 3D models as never before possible, and will allow an order of magnitude better resolution of the global heliosphere.

The 3D reconstructions from PERSEUS observations derive from the fact that the spacecraft (SC) is within the solar wind "cloud" that passes by the SC (Figure 1). Optically thin, this always-present material passes the SC registered as line of sight variations that undergo changes following very specific physics and brightness change as they move outward. This has allowed several groups to reconstruct the 3D heliosphere from perspective views of the passing structure. Much of this work is described in Jackson et al. (2008), and the numerous references therein.

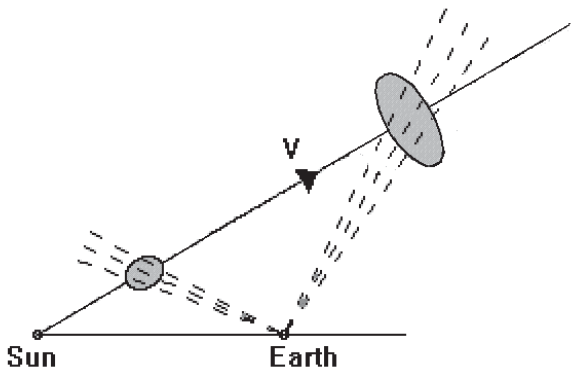


Figure 1. – Perspective views of outward-moving structure provide heliospheric structure line-of-sight differentiation.

There is no better place to observe Thomson-scattered signals that relate to the SC than at the SC itself, because all lines of sight emanate from here. Beyond elongations (angular distances from the Sun) of 90º the average maximum response from electrons is situated near the SC (see Figure 2).

The ASI consists of a novel baffle-and-optics system (see Figures 3 & 4). Laboratory measurements have shown that such a baffle provides many decades of stray-light rejection for light sources which lie below the baffles 180° "field of regard" (Figure 5).

ASI data are dominated by the bright night sky: the Milky Way and stars together with the zodiacal light; these are 100× brighter than the transient CMEs and other solar wind variations that are the focus of PERSEUS. This bright background will be removed, leaving the residue of white light emission that we are interested in: Thomson-scattered sunlight from free electrons, which renders visible the ionized solar wind. Analysis of SMEI observations has shown how to remove these background sources of light effectively. The residue provides Thomson-scattered light above a long-term mean average. The brightness peak for the CME observed remotely by SMEI (Figure 2) has a corresponding peak in proton density measured by plasma monitors from the ACE and other near-Earth spacecraft. The perspective views of outward-flowing plasma allow 3D reconstruction of this structure density and a modeled density peak observed in situ that is more sharply defined than the brightness measurements. These large density variations are modeled to provide densities relative to the ambient mean. For PERSEUS, the long-term mean, a minor component of the large structure variations, will be provided by the in-situ densities obtained by the PERSEUS plasma monitor.

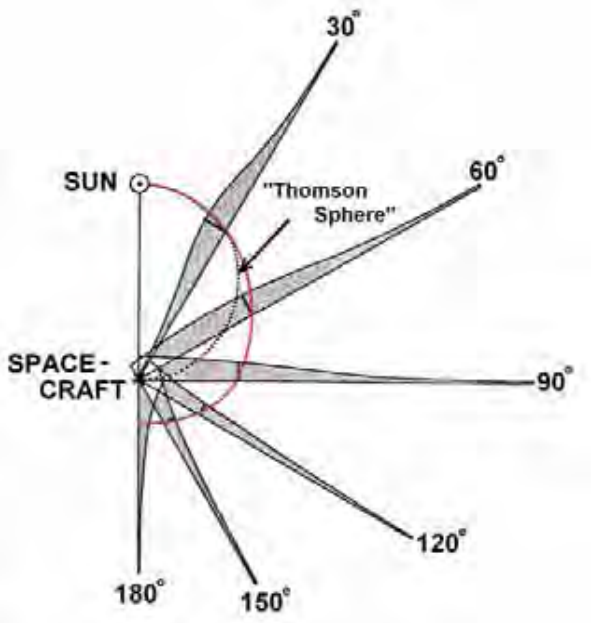


Figure 2. – Line-of-sight distribution of the heliospheric brightness response for an r^{-2} spherically symmetric heliosphere. The red line shows the location of the mean response.

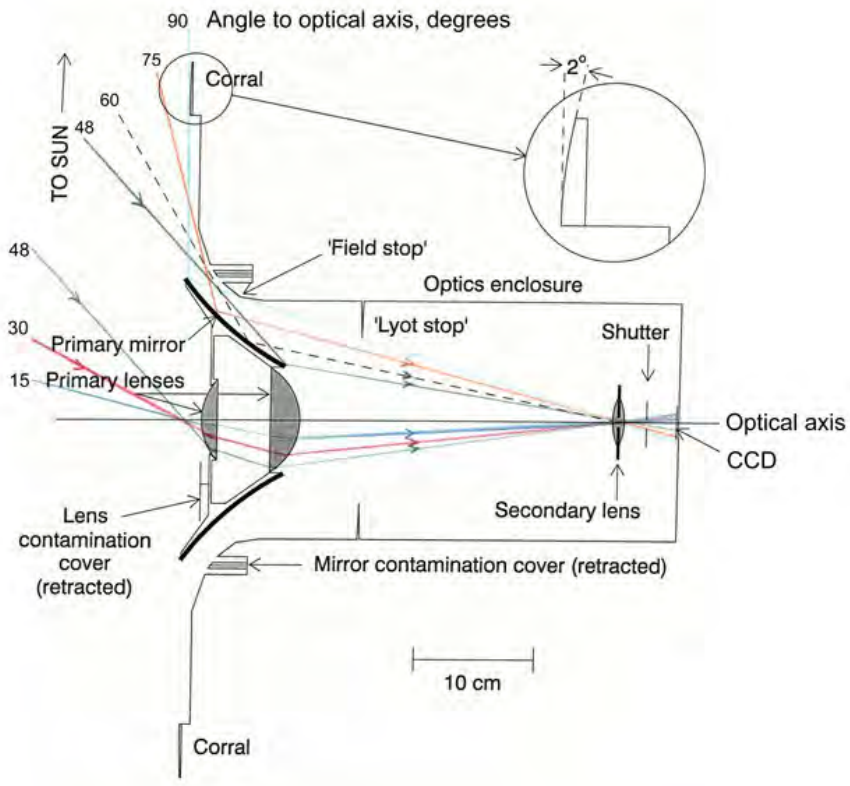


Figure 3. – Schematic Optics Cross-section for an ASI imager.



Figure 4. Cutaway diagram of ASI imager (left) and toroidal mirror (right). For more description, see Buffington 2000, and references therein.

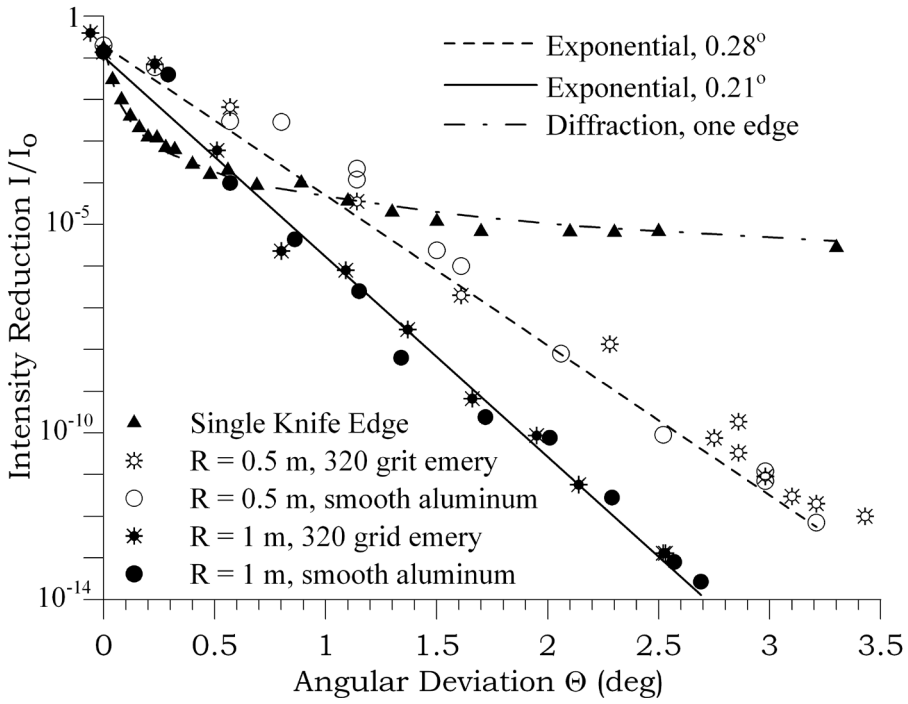


Figure 5. Laboratory measurements of stray-light reduction for a smooth corral (from Buffington, 2000).

CRG data are similar to ASI data except that the FOV is close to the Sun and much smaller, and the fractional contribution of background from Milky Way and stars is reduced. In addition, our base mission will use polarization filters for CRG to determine long-term coronal density using polarization brightness (pB) signals. These are sensitive primarily to Thomson scattering and not zodiacal light and can therefore be separated. PB also gives a more peaked coronagraph brightness response near the solar limb than does the unpolarized brightness response. This has

been used to separate solar wind flow both toward and away from Earth. To the extent that zodiacal light brightness remains stable during a particular observation, this means that pB provides better contrast for viewing solar limb events, while unpolarized light is better for viewing halo events headed toward Earth.

Both CRG and ASI will provide measurements of flow in the transient and ambient solar wind. Subtraction of a sidereal map as in SMEI, will remove stellar signals from both instruments, and long and/or combined shorter-image exposures to provide sufficient throughput to measure large (one to several degrees in size) ambient structures at statistically-significant signal-to-noise levels. Image subtractions then enable measurement of outward flow over small sections of each set of difference images, similar to analyses using LASCO C3 coronagraph image sequences.

In addition to the heliospheric science objectives described above, the continuous, full-sky coverage of PERSEUS enables other measurements of astrophysical interest. White light observations enable studies of the zodiacal dust cloud, and the extended geocorona. PERSEUS will also obtain long-term photometric brightness timeseries for thousands of bright stars across the sky. The analysis of brightness variations of these stellar timeseries will support a number of astrophysical studies, such as the search for extrasolar planets by stellar occultation; the study of eclipsing binaries, and intrinsic variable stars; the search for optical counterparts of gamma-ray bursts, and detections of novae. PERSEUS will record zodiacal dust brightness variability and allow observations of comet tails extending many tens of degrees from the comet core, which can be used to derive estimates for the solar wind speed near the comet. The full sky coverage of PERSEUS provides an opportunity for asteroid studies, including the search for Near-Earth Objects (NEOs) and possible detections of dust bursts from colliding asteroids.

III. References

Jackson, B.V., Hick, P.P., Buffington, A., Bisi, M.M., Clover, J.M., and Tokumaru, M., 2008, "Solar Mass Ejection Imager (SMEI) and Interplanetary Scintillation (IPS) 3D-Reconstructions of the Inner Heliosphere", *Adv. in Geosciences*, **21**, 339-366

Buffington, A., "Improved design for stray-light reduction with a hemispherical imager", Applied Optics 39, 2683-2686

Campaign Observations of the Heliosphere During the STEREO Superior Conjunction

E.A. Jensen, M.M. Bisi, A.R. Breen, A. Buffington, J. Clover, R.A. Fallows, A. Kraus, D. Oberoi, F. Vilas, A. Vourlidas

Abstract

In 2009 during the longest solar minimum in a century, we conducted a comprehensive observation of the heliosphere from the solar limb to 1 AU measuring the radio signal from natural sources and the MESSENGER (MSGR) spacecraft in superior conjunction using radio antennas in the USA, Germany, India, and Japan providing a snapshot of the magnetic, velocity, and density fields in this unusual heliosphere. While the integration of these datasets continues, the effort has shown that important parameters of the heliosphere can be simultaneously measured during these opportunities. We recommend to the committee to consider the scientific advances that can be achieved during the upcoming superior conjunction of the two STEREO spacecraft following solar maximum. Enhancing the capabilities of the Deep Space Network (DSN) maximizes the return on this opportunity; we argue that the CME advance warning component to this experiment justifies the cost.

Introduction

The plasma of the solar corona streams away from the Sun accelerating to become the supersonic solar wind, and carrying with it the interplanetary magnetic field. At the orbit of the Earth, the solar wind encounters the obstacle of the Earth's magnetosphere, where it decelerates with the aid of a standing bow shock, behind which the plasma and interplanetary magnetic field interacts with the geomagnetic field at the magnetopause. Reconnection of the interplanetary magnetic field with the geomagnetic field couples the momentum flux of the solar wind to the magnetosphere. As a result, the Earth's geomagnetic environment is very sensitive to changes in the solar wind's properties. The solar cycle variability of geomagnetic activity is caused by the changes in the occurrence rate and strength of coronal mass ejections, the changes of the interplanetary magnetic field and solar wind flow, and other solar phenomena such as, solar energetic particle events that evolve in the course of the Sun's 11 and 22 year cycles.

From 2007 to 2009, the Sun experienced its longest solar minimum in a century. Around November 2009, observations were made of this unusually generated solar wind from 1 to 215 solar radii using the Greenbank Telescope, Effelsberg Telescope, the Ootacamund Radio Telescope, BSA-LPI, and the Solar-Terrestrial Environment Laboratory radio arrays. We are beginning to mesh these datasets with other solar observations and models to investigate these unusual heliospheric conditions.

Much of the plasma of the heliosphere can be studied as a fluid using magnetohydrodynamics (MHD). The critical parameters to understanding the solar wind consists of the magnetic vector field, the velocity vector field, the density scalar field, and the overall structure (i.e. Stream Interaction Regions). We are looking to estimate these parameters for November 2009, and perform a similar analysis with other coordinated observing campaigns, such as the unique opportunity of the superior conjunction of the STEREO spacecraft, as we move closer to solar maximum estimated to be early-to-mid 2013. The two STEREO spacecraft during superior conjunction will provide a several months-long opportunity to measure the Sun's "invisible" magnetic field within 10 solar radii; this has not been done before.

This research addresses the survey's goals to prioritize scientific advances with NASAs and NSFs programmatic goals including the DSN portion of a previous NRC survey. There are two scientific problems which fall under NASA's Strategic Goals Sub goal 3B: understanding the basic principles of the origin of the solar wind in the corona and using this information to predict the Earth's "space weather."

Table 1 illustrates how this proposed research contributes to the SR&T research thrust. Table 2 lists the NSF/SHINE program priorities under which these observations advance solar physics and provide a needed operational service to the space weather community. Table 3 lists the NRC challenges this research addresses.

***Table 1**: Relevant SR&T research thrust topics.*

SR&T Investigations	Relevant Proposal Sections
Ground- and Space-based Observations	Faraday rotation, IPS, and Dispersion Measure (DM) observations of spacecraft and natural sources giving magnetic field, velocity, and density measurements
Theory and/or Modeling	3-D tomography applications and CME sheath observations
Data Analysis	Intrinsic coronal magnetic, velocity and density fields, CMEs, and MHD waves

***Table 2**: NSF/SHINE topics this research covers.*

Shine Topics	Task
"variations in the solar wind, interplanetary magnetic field, and plasma structure"	Investigate the solar wind structure in Faraday rotation, IPS, and DM datasets collected in observing campaigns
4. "Quasi-steady structure of the corona and interplanetary medium, as it relates to geoeffective phenomena in general (this includes understanding or characterizing the background solar wind and how it evolves, and its relationship to solar structures such as coronal holes, streamers, and filaments)"	Processing and analyzing previous FR datasets has provided valuable insight into (a) predicting the geoeffectiveness of CMEs, (b) the net efflux energy into the solar wind from MHD waves, and (c) the first order investigations in the magnetic contribution to force balance in the solar corona (Jensen, 2007)
5. "Solar cycle dependence of the above phenomena"	Data collected during the declining phase following the 2013 solar max; data collected during the 2007-2009 solar minimum, a once in a century event

***Table 3**: NRC Challenge Summary and STEREO FR Application; the 5 challenges are:*
(1) Understanding the structure and dynamics of the solar corona
(2) "How do energetic solar events propagate through the heliosphere?"
(3) "How does Earth's global space environment respond to solar variations?"
(4) Understand the basic physical principles and processes in solar plasma
(5) Develop near real-time predictive capability

STEREO FR Application	Challenge
Observe MHD waves crossing the line of sight to study the transfer of wave energy in the corona	1 & 4
Use FR observations to compare to the results from heliospheric models	1, 4, & 5
Observe magnetic flux rope structure in CMEs crossing the line of sight and study changes in orientation, size, velocity, and structure upon reaching 1AU	2, 3, & 5

Coronal Mass Ejections

Coronal Mass Ejections (CMEs) consist of 10^{15} to 10^{16} grams of mass (Subramanian and Vourlidas, 2007) and "frozen-in" magnetic field erupting from the corona of the Sun achieving speeds between 250

and >2000 km/s. At speeds exceeding those of the downstream solar wind, a fast shock develops at the boundary between the CME and the downstream solar-wind plasma. The Solar Energetic Particles (SEPs) accelerated by a shock can adversely affect spacecraft and initiate ionospheric disturbances (Kane, 2007), while the shock itself correlates with the onset of more intense geomagnetic storms (Badruddin and Singh, 2009). Intense geomagnetic storms generate significant radiation and cause the plasmasphere and ionosphere to expand (ie Bisi et al, 2010). While spacecraft components can usually tolerate 100 krads of radiation over their lifetime, a geomagnetic storm can generate 1 krad thus reducing the lifetime of spacecraft; as a result, shielding is commonly implemented increasing the overall cost. For instance Telstar 401, an AT&T television relay satellite, failed suddenly on 21 January 1997 following an Interplanetary CME (ICME) triggered geomagnetic storm in the middle of a Star Trek broadcast (Anselmo, 1997; Livesey, 1998). The expansion of the ionosphere increases drag on satellites and debris in Low Earth Orbit (LEO); this led to the early termination of the Solar Maximum Mission in 1990 (Smith, 1988), Skylab in 1979, and the U.S. Space Command had to recompute orbits for more than 1,300 objects in 1989. The ionospheric currents generated during the storms induce electrical currents in electricity lines: 2.6 kV surges were recorded in the Atlantic cable between Scotland and Newfoundland in March 1940, and a geomagnetic storm caused a power blackout in Quebec on 13 March 1989. Ionospheric scintillations can disturb the amplitude and phase of traversing radio signals; this can cause navigation errors and signal loss (Nordwall, 1996). In 1998, satellite insurance companies paid $0.9 billion in claims for satellite failures. By 2005, $80 billion worth of telecommunications services were carried by satellites. The Committee On The Societal and Economic Impacts Of Severe Space Weather Events (2008) estimated a “severe geomagnetic storm scenario” could incur $1 trillion in societal and economics damage.

The NOAA Space Weather Center issues warnings regarding CME eruptions from the Sun; however, the geoeffectiveness of the ICME cannot be accurately quantified until the spacecraft in the L_1 libration point sunward of the Earth provide *in-situ* magnetic-field measurements. This gives an advance warning time of approximately an hour depending on ICME speed; in the case of the Halloween event in 2003, this was just 15 minutes. The only remote-sensing observing technique that can determine the magnetic-field orientation earlier and extend this advance warning time is that of Faraday rotation (FR).

Jensen *et al*. (2010) demonstrated with spacecraft FR observations of CME/ICMEs that the critical parameters of the orientation and radius of a flux cylinder can be determined; however, the orientation measurement from a single line of site (LOS) is degenerate. Using the capabilities of a FR imager such as the Murchison Widefield Array (MWA), we find that the magnetic-field configuration in a CME/ICME can be uniquely determined. We also find that FR images pose an unparalleled opportunity to investigate the configuration of the magnetic field within the shock/sheath region on the leading edge of fast CME/ICMEs. These shock/sheath regions can potentially produce a FR response of similar strength as the flux cylinder; therefore, we recommend measuring FR at the greatest frequency possible thus observing the CME at the smallest heliospheric distance possible. Using the Halloween 2003 October 29 event to study the use of FR images in increasing the advance warning time of accurate geoeffectiveness from magnetic-field measurements, we find an order of magnitude improvement from the average 1 hour to ten hours.

A FR experiment using the STEREO spacecraft could accurately test the feasibility of the advance warning capabilities of this technique.

The 2009 November Observing Campaign

During early November 2009, radio observations were made of the heliosphere in the USA, Germany, India, Russia, and Japan of different regions of the sky. The Greenbank (GBT, USA) and Effelsberg (Eff, Germany) Telescopes observed the radio frequency signal of the MESSENGER (MSGR) spacecraft in

superior conjunction while other radio antennas observed natural sources. GBT and Eff observed the FR and scintillation of the MSGR RF signal by the solar corona as the spacecraft passed through superior conjunction. While FR provides a measure of the magnetic field, radio scintillation measurements allow the measurement of velocity and *g*-value. If two antennas observe the same source simultaneously, then, provided the geometry is correct, the time lag of a cross-correlation of the two signals can be used to provide an estimate of the solar outflow speed.

Magnetic Vector Field Datasets

Outside of active regions, the only available technique to measure the magnetic field is FR. This is the measurement of the rotation of the plane of polarization of an electromagnetic wave from either a spacecraft or a natural radio source as it traverses the birefringent medium of the plasma of the solar corona. The amount of rotation is an integrated quantity proportional to the component of the magnetic field along the LOS weighted by the electron density. This technique is also used with a density model (theoretical, empirical, or inversion) to estimate the magnetic field of the region along the LOS.

*From 2009 Nov 8-11, observations of FR due to the solar corona were collected using the MSGR spacecraft as it passed through superior conjunction. GBT measured the raw RF samples; these have been processed to obtain plane of polarization (PP) data. FR measurements are obtained from PP data following calibration for antenna effects on the signal. Figure 1 shows an unusual MHD wave (11-min fluctuation) observed in this dataset from the 8th; this type of analysis has been used to study the net wave efflux energy of the solar corona (Jensen and Russell, 2009), a suspected source of coronal heating.

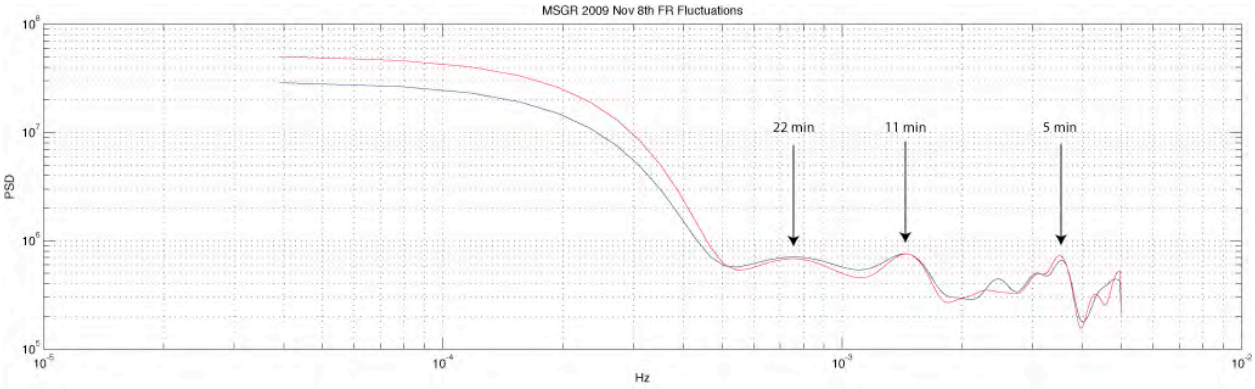


***Figure 1**: The PSD spectrum of the MSGR FR fluctuations observed on 2009 Nov 8th. The 5-min fluctuation has been observed previously (i.e. Andreev et al, 1997); however, the 11-min fluctuation is unusual. Jensen (2007) demonstrated that magnetosonic waves will generate a secondary harmonic at twice the primary frequency; a suspected source of the 22-min fluctuation*

Within active regions of the Sun, the protons orbiting the magnetic-field lines generate Zeeman-splitting lines and the Hanle effect giving information on the configuration of the magnetic field in the area along the LOS (Zeeman) and in the plane of the sky (Hanle). This technique is used as one of the boundary conditions in the Potential Field Source Surface (PFSS) magnetic field model.

*WSO and MWSO generate PFSS magnetic field solutions on a continuous basis.

Radio frequency bursts release electromagnetic waves along the magnetic-field lines connected to the active region; these are detected by spacecraft/instruments that happen to intersect with the magnetic-field line at the time the particles flow past. This technique is used with a density model (theoretical, empirical, or inversion) to estimate the magnetic field of the active region.

*Multiple spacecraft carry instruments that can measure solar radio bursts: *i.e.* Cassini Radio and Plasma Wave Science Instrument, STEREO|SWAVES instrument, and WIND|WAVES. Additionally, radio bursts are regularly measured with ground-based instruments such as the Green Bank Solar Radio Burst Spectrometer.

Velocity Vector Field

Interplanetary scintillation (IPS) arises from the small-angle scattering of radio waves by density inhomogeneities in the solar wind flowing approximately radially outward from the Sun. If variations in an apparent source intensity are recorded simultaneously by two (or more) telescopes with adequate geometry, the cross correlation of the signals can be used to estimate the solar wind speed (or speeds when there are multiple streams crossing the LOS).

*This is done on a continuous basis by STELab when weather is favorable (around April-December each year).

*University of California, San Diego (UCSD) 3-D tomographic codes use single antenna IPS observations (*g*-level) of several hundred sources to generate a density scalar reconstructions and STELab cross-correlation velocities for a velocity vector field of the heliosphere.

*During the MSGR superior conjunction on Nov 11th, GBT and Eff measured the scintillating signal.

Higher-resolution Density Scalar Field

White-light observations of the solar corona from Thomson scattering of photons off of electrons are regularly taken with a variety of instruments. These are weighted brightness observations along the LOS for each pixel in the images. Three-dimensional (3-D) tomographic techniques can invert these observations to produce a density scalar field.

*SOHO|LASCO, STEREO|SECCHI HIs, and SMEI all provide sufficient data for the UCSD tomographic codes to generate density scalar fields of the heliosphere.

*UCSD tomographic codes use single antenna IPS observations (*g*-level) of several hundred sources to also generate a density scalar field of the heliosphere.

The STEREO Extended Mission Opportunity

From January 18 to July 28, 2015, the two STEREO spacecraft will be in superior conjunction (within 20 solar radii of the Sun). The unique conjunction offsets are shown in Figure 2. The STEREO-A spacecraft will observe the same region of space over a fraction of a solar rotation allowing limited time varying observations of the region; its offset distance changes roughly 1 solar radii over the period of 7 days. From February to April, the two spacecraft will be on the same side of the Sun observing at different solar offsets allowing time and space varying observations. The majority of observations from STEREO-B will be spatially varying. Because the conjunction time period lasts for 10 months, there is a significant probability of CMEs or MHD waves crossing the lines of sight. We note that there will not be a time that the spacecraft are on opposite sides of the Sun simultaneously providing FR data due to the signal-to-noise requirements.

With 2 antennas constituting a large spatial baseline observing the same spacecraft signal, the solar wind velocity can be measured in the plane of the sky. If an MHD wave is observed to cross the LOS, the 2 antennas can also simultaneously obtain its wave speed (Andreev *et al*, 1997). Performing this experiment to maximize the data return (as estimated in the budget) requires 2 antennas at each complex (1 per spacecraft) from February to May providing 24-hour coverage.

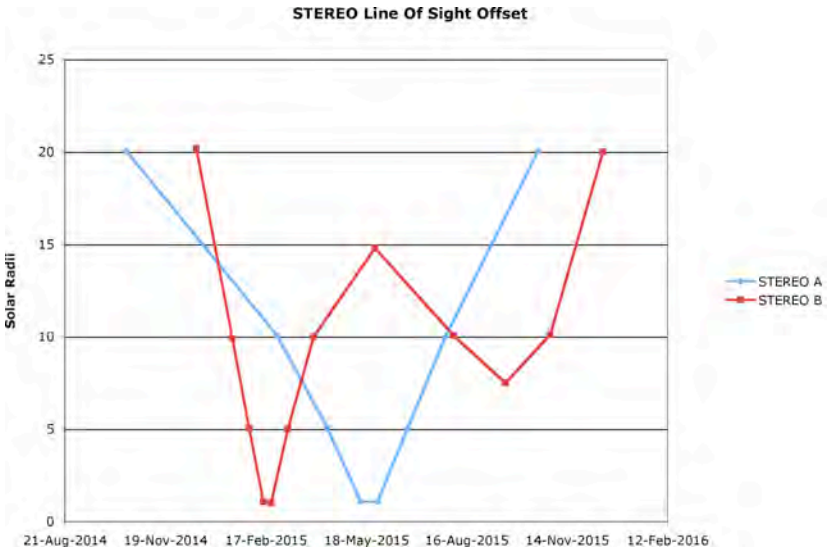


***Figure 2**: Offsets of the STEREO lines of sight during the 2015 superior conjunction.*

The Deep Space Network

JPL's Deep Space Network (DSN) is the best antenna system for the FR portion of the observing campaign. It has antenna arrays at roughly every 120 degrees longitude allowing a few hours of simultaneous observations between arrays, and the antennas are designed specifically for the transmission of data on the narrow bandwidth carrier frequencies with minimum signal loss. The Future of Solar System Exploration 2003-2013 decadal survey (Asmar *et al*., 2002) lists the benefits of the DSN to solar studies including the study of the large-scale solar magnetic structure and electron density distribution. The recommendation in the study to enable dual-polarization capability in the 34-meter antennas will need to be implemented to pursue the STEREO FR experiment. This enhancement of the DSN's infrastructure will further support NASA's research objectives in Earth Science, Planetary Science, and Astrophysics enabling studies that incorporate polarization observations.

Timeline

We suggest the following timeline for implementing this research opportunity:

Jan 1, 2012	Latest date for decision on the STEREO FR opportunity
Jun 1, 2012	Proposal submission deadline for funding to use the FR data
Jan 1, 2013	Proposal Approval and start to configure the data products and formats based on the studies for which they will be used and archiving in the Planetary Data System
Jan 15, 2014	Begin to build/update the radio receivers that will be used in the 34-meter antennas and design the computer system that will be implemented for real-time data release on the internet
Jun 15, 2014	Place radio receivers in the antennas and begin calibrating the antennas and ground system
Sep 16, 2014	Perform last set of calibration observations using the STEREO radio frequency carrier and have real time data system enabled

Jan 18, 2015	Begin experiment
Jul 28, 2015	End experiment
Sep 28, 2015	Perform last set of calibration observations using the STEREO radio frequency carrier and have data archive reviewed

Budget

We estimate that the overall cost to perform this study, and maximize its data return, to be between $28.7 and $29.3 million; the table below gives the estimated costs for the major components of the project. We estimate that the STEREO operations can be performed at half-capacity over the two years of calibration observations and experiment time frame. Research projects in this study will probably vary in scope requiring 1-to-3 years and individual-to-small science teams based on the topic of study, and roughly 10 proposals would cover the range of research areas that would benefit the most from the observations. The radio receiver upgrades will cost approximately $100K per receiver to add the opposite polarization capability. Because technology in RF amplifiers and other waveguide components has significantly improved, new construction of the receivers is preferred (at $200K/receiver). Differences between the components in the two polarizations can thus be minimized.

STEREO Operations (half capacity)	$21 million
Proposal Funding	$5 million
DSN and Radio Science Personnel	$2 million
Radio receiver upgrades or construction	$0.6 or $1.2 million
Computer hardware and resources	$0.1 million

References

Andreev V.E., Efimov A.I., Samoznaev L.N., Chashei I.V., Bird M.K., 1997, SoPh, 176, 387
Anselmo J.C., 1997, AW&ST, 146, 61
Asmar S.W., et al., 2002, ASPC, 272, 355
Badruddin, Singh Y.P., 2009, P&SS, 57, 318
Bisi M.M., et al., 2010, SoPh, 265, 49
Committee On The Societal, Economic Impacts Of Severe Space Weather Events, 2008, sswe.rept
Jensen E.A., 2007, PhDT
Jensen E.A., Hick P.P., Bisi M.M., Jackson B.V., Clover J., Mulligan T., 2010, SoPh, 265, 31
Jensen E.A., Russell C.T., 2009, GeoRL, 36, 5104
Jensen E.A., Russell C.T., 2007, SPIE, 6689
Kane R.P., 2007, AdSpR, 39, 1890
Livesey R.J., 1998, JBAA, 108, 207
Nordwall B.D., 1996, AW&ST, 145, 60
Smith E.A., 1988, fmet.symp, 571
Subramanian P., Vourlidas A., 2007, A&A, 467, 685

Next Generation Experiments for Laboratory Investigations of Magnetic Reconnection Relevant to Heliophysics

H. Ji[1], M. Yamada[1], S. Bale[2], A. Bhattacharjee[3], T. Carter[4], W. Daughton[5], J. Egedal[6], T. Phan[2], S. Prager[1], R. Rosner[7], V. Roytershteyn[5], K. Shibata[8], S. Tsuneta[9], E. Zweibel[10]

[1]Princeton Plasma Physics Laboratory, Princeton University, [2]University of California, Berkeley, [3]University of New Hampshire, [4]University of California, Los Angels, [5]Los Alamos National Laboratory, [6]Massachusetts Institute of Technology, [7]University of Chicago, [8]Kyoto University, Japan, [9]National Astronomical Observatory, Japan, [10]University of Wisconsin - Madison

Abstract

This concept paper describes the scientific opportunity for next generation laboratory experiments to study magnetic reconnection in regimes directly relevant to space and solar plasmas. In parallel to theory and simulations, dedicated laboratory experiments have contributed significantly to the understanding of fast reconnection in the past decade, such as validation of two-fluid effects and detection of microturbulence in the reconnecting current sheet. However, further critical contributions to space and solar plasmas in the coming decade are limited by the achievable values of controlling parameters in these ongoing experiments. The new theme of magnetic reconnection with multiple X-lines can possibly provide solutions for fast reconnection in large systems and for efficient particle acceleration as suggested by recent numerical and theoretical studies. To meet these challenges, new ideas for next generation experiments are emerging to provide the required experimental capabilities to study detailed reconnection process, efficient particle acceleration and spontaneous global self-organization in the parameter regimes directly relevant to space and solar plasmas.

1. Introduction

Magnetic reconnection, the efficient release of magnetic energy by topological rearrangement of field lines, is one of the most important and fundamental plasma processes in heliophysics. It plays key roles in a wide range of heliospheric phenomena from solar flares, coronal mass ejections, solar wind propagation and dissipation, interaction of interplanetary plasma with Earth and other planets' magnetosphere, dynamical response of planetary magnetopheres to solar wind such as magnetic substorms, and interaction of interstellar media with the heliosphere. Magnetic reconnection is considered to be a necessary part of dynamo processes in generating and maintaining magnetic field in the sun and in planetary cores. The importance of magnetic reconnection has been also recognized in more distant astrophysical plasmas, such as star formation and explosive phenomena from strongly magnetized neutron stars. Back on Earth, laboratory fusion plasmas often suffer from uncontrolled reconnection events that can adversely affect confinement. Therefore, understanding physics of magnetic reconnection is not only of fundamental importance for heliophysics but also with broader impacts across other disciplines.

2. Roles of Dedicated Laboratory Experiments in Understanding Reconnection

Despite the long history of magnetic reconnection research, the most important progress has been achieved only recently, especially during the past decade (see recent reviews [1,2]). Many parts of this progress were accomplished with contributions from dedicated laboratory experiments, which have become increasingly well-controlled and well-diagnosed. Below we list a few examples of these accomplishments.

- Quantitative tests of Sweet-Parker model [3]. This classical reconnection model proposed in 1950s was not tested experimentally until late 1990s. Significant modifications were required for quantitative agreements with experiment.
- Two-fluid effects for fast reconnection [4,5]. Existence of a quadrupolar out-of-plane magnetic field was predicted numerically [6], consistent with space data [e.g., 7], but definite confirmation by 2D measurements was done in the laboratory with theoretically predicted dependence on collisionality.
- Detection of electron diffusion region [8,9]. Its thickness was measured accurately and compared directly with state-of-the-art 2D PIC simulations. The comparisons challenged modeling and provided physical insights that 3D physics might be important in explaining the remaining disagreements with data. Detecting and studying the electron diffusion region is a goal of the MMS mission.
- Microturbulence in current sheets [10-12]. Predicted theoretically, electrostatic fluctuations were detected for lower-hybrid drift waves or electron phase space holes, but electromagnetic fluctuations were surprises to theory [13].
- Potential well and electron kinetics [14]. With measurements and physics understanding of electrostatic potential around the X-line, one can quantitatively explain the measured electron distribution function near the Earth's magnetotail.
- Ion heating by reconnection [15,16]. It was found in the laboratory that increments of ion energy are proportional to square root of ion mass or ion temperature is proportional to the charge-to-mass ratio. These findings may help solve the ion heating problem in the solar wind.
- Flux rope dynamics [17-21]. In attempts to be relevant to the dynamics of coronal loops, several experiments were performed either in linear or curved geometries with line-tied boundary conditions. Equilibrium and stability of the line-tied current loops is being studied quantitatively and compared with theories [e.g., 22].

It is clear from these examples that well-controlled and well-diagnosed experiments play increasingly important roles in understanding reconnection phenomena in heliophysical plasmas.

3. Outstanding Issues for Magnetic Reconnection in Heliophysical Plasmas

Despite significant progress in reconnection research, there exist a number of outstanding issues to be addressed before a better understanding of reconnection phenomena in heliophysical plasmas can be achieved.

- How is the reconnection rate determined? *(The rate problem)* Although two-fluid effects (e.g. the Hall term) have been shown to facilitate fast reconnection, the required efficient dissipation within the electron diffusion region is still elusive, and micro-turbulence in the current sheet can contribute to the solution of this problem. How does magnetic reconnection take place in partially ionized plasmas, such as in solar chromospheres where neutral particles dominate? Recent observations by HINODE

show that reconnection-like energetic activity is abundant in these partially ionized plasmas.

- How does reconnection take place in 3D? *(The 3D problem)* In addition to the 3D microturbulence within the reconnecting current sheets, there are situations where the magnetic field vanishes in isolated three-dimensional points, or so-called 3D null points. Magnetic reconnection around these 3D null points or null-null lines [23] is important for understanding complex dynamics of the general magnetic topology, such as flux transfer events in the Earth's magnetosphere [24].
- How does reconnection start? *(The onset problem)* Large-scale reconnection phenomena, such as solar flares, coronal mass ejections, or magnetic substorms, often manifest themselves as impulsive relaxation events. Laboratory experiments [25, 26] and some numerical models [e.g. 27, 28] also exhibit spontaneous reconnection. However, it is unclear under what conditions these spontaneous reconnection takes place or whether these conditions are universal.
- How are particles energized? *(The energy problem)* Reconnection mainly manifests itself as an energetic process of releasing magnetic energy to plasma particles, sometime in the form of a high-energy tail of distribution functions [29-31]. Understanding how this process occurs is a major challenge. In addition to the acceleration due to single X-line reconnection, interactive multiple X-lines reconnection may provide a solution to this problem [32,33]. Additional possibilities, such as shock-based reconnection models, may be important to generate energetic particles [34].
- How do boundary conditions affect the reconnection process? *(The boundary problem)* It is not self-evident that reconnection physics in periodic systems can be straightforwardly applied to space and solar plasmas which are always non-periodic, and often line-tied at their ends [35]. Determining effects of boundary conditions is important to understand reconnection in heliophysical plasmas.
- How to apply local reconnection physics to a large system? *(The scaling problem)* Even all above issues are completely understood, there still exists an important issue on how to translate physics in small systems to larger heliophysical systems. When the plasma size

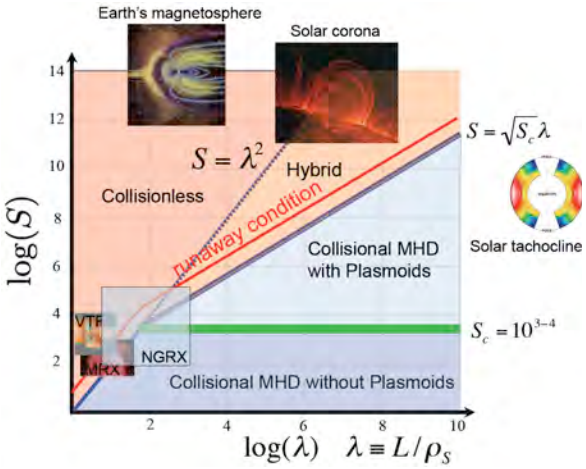


Figure 1. *A phase diagram for magnetic reconnection in 2D. If either S or the normalized size, λ, is small, reconnection takes place in collisional MHD (without plasmoids) regime or in collisionless regime. When both S and λ are sufficiently large, two new regimes appear: a regime for collisional MHD with plasmoids and a regime for collisional MHD plasmoids with kinetic current sheets in between. The regimes for reconnection in Earth's magnetosphere, solar corona, and solar tachocline are also shown. The existing experiments, such as MRX and VTF, do not have accesses to these new regimes while the next generation reconnection experiments (NGRX) should provide access to these relevant regimes for heliophysical plasmas.*

normalized by ion sound gyroradius, λ, and the Lundquist number, S, are sufficiently large, the reconnection physics derivates significantly from that either in small S or in small λ. A new phase diagram [36,37,28] emerged from recent large-scale numerical simulations [e.g., 38,39] as shown in Fig.1 where four reconnection phases are illustrated: collisional MHD without plasmoids, collisional MHD with plasmoids (corresponding to reconnection in solar tachocline), collisional MHD plasmoids with kinetic current sheets in between (or hybrid, corresponding to solar corona) and collisionless (corresponding to Earth's magnetosphere). In order to be directly relevant to reconnection phenomena observed in heliophysical plasmas, it is crucial for laboratory experiments to access all of these reconnection regimes. We note that reconnection in these new regimes is associated with multiple X-lines, and thus can possibly provide solutions to the onset problem and the energy problem mentioned above.

4. Existing laboratory experiments and their capabilities

Table I summarizes the capabilities by existing experiments for each issue listed above. With diversity in geometries and focus areas, these experiments are expected to continue to help address many of these outstanding issues into the next decades. Advancing in plasma diagnostics and interacting closely with observations, theory and modeling will continue to be a key. However, the critical issue of applying local reconnection physics leaned from these experiments to a large system, such as solar coronal plasmas, is difficult to be addressed due to the limited size and Lundquist number available in these experiments.

Issue	Existing Experiments	Opportunities for new Experimental Study
Rate	TS-3/4, MRX	Partially ionized plasma
3D	LAPD, MRX, SSX, VTF, Caltech	3D-null point
Onset	VTF	Global self-organization
Energy	3D-CS, LAPD, TS-3/4, MRX, SSX	Particle acceleration
Boundary	RSX, RWM, MRX, LAPD, Caltech	Non-periodic
Scaling	MRX	Large parameter space

Table I. *Outstanding reconnection issues addressed by the existing experiments and opportunities for new experimental studies.*

5. Opportunities and new ideas for next-generation reconnection experiments

Also listed in Table I are opportunities for new experimental studies in the coming decades for each outstanding issue. Among these, a major opportunity identified is to access the new reconnection regimes described in Sec. 3. Next generation experiments would therefore seek to achieve plasmas of normalized size on the order of 10^3 and Lundquist number on the order of 10^5 (see Fig.1). The physics of reconnection in these new regimes has more direct relevance with heliophysical plasmas than those already accessible by the existing experiments. With the capabilities to access these new regimes with multiple X-lines reconnection, the critical issues on particle acceleration, global impulsive relaxation, and non-periodic boundary conditions in a large system can

be studied. Reconnection in partially ionized plasmas would be facilitated through larger plasma volumes. Shock-based reconnection models could also be tested more conclusively in larger plasmas.

New ideas for next generation experiments are emerging to meet these challenges. One example concept for such a next generation experiment is based on the Magnetic Reconnection Experiment (MRX) [40]. The MRX is a small experimental device dedicated for laboratory study of magnetic reconnection, using a set of specially designed internal coil assembles. Many significant results [3,4, 8,9,10,12,21,41,42] have been documented on a wide range of reconnection processes from collisionless to collisional, from periodic to line-tied, from fully ionized plasmas to partially ionized plasmas, from local electron dynamics to global MHD physics. A next generation experiment with a device diameter of 3 m and length of 4 m, with enhanced field strength and heating power, can increase the normalized size (λ) by a factor of 3-6 and Lundquist number (S) by a factor of 8-50. Preliminary estimates show that, for anti-parallel reconnection, $\lambda \sim 100$ and $S \sim 5000$, and for guide field reconnection, $\lambda \sim 1{,}000$ and $S \sim 33{,}000$ (using reconnecting field) can be achieved. The new reconnection regimes shown in Fig.1 are within the reach by these numbers. In addition, the number of internal coils and electrodes can be increased to control connection better and to provide necessary flexibilities for reconnection studies in new geometries more relevant to heliophysics, such as interacting, line-tied plasma arcs. Measurements with high spatial resolution and wider coverage (to allow the capturing of reconnection at multiple X-lines) would be required. Additionally, non-perturbative diagnostics are desirable. The estimated construction costs are $5-10M over 2 years and operational costs are $2.5-3M per year. The main goals of this concept are

1. To study new regimes of multiple X-line reconnection directly relevant to heliophysical plasmas
2. To study particle acceleration in multiple X-line reconnection regimes
3. To study global impulsive relaxation, and the effect of non-periodic boundary conditions

Another example idea for a next generation experiment is based on Versatile Toroidal Facility (VTF) [43]. VTF is a small experiment dedicated to study fast reconnection in collisionless plasmas with a strong and variable guide field. Many important results [11,14,26,44] have been obtained on particle dynamics, micro-turbulence, and reconnection 3D onset, and have been successfully applied to understanding satellite observations. The parallel electric field required for the electron pressure anisotropy can provide effective dissipation energizing super thermal electrons. This particular mechanism of dissipation and associated particle acceleration can be studied by a next generation reconnection experiment with the estimated device diameter of 5 m and length of 2 m, with a series of internal coils which can be controlled to drive reconnection. In addition, multiple island interactions can be also studied as another source of super thermal electrons. The main goals of this concept are

1. To study electron pressure anisotropy near the X-line and associated particle acceleration in collisionless plasmas
2. To study the reconnection onset as a consequence of 3D dynamics

In addition to these new ideas for the next generation reconnection experiments, experimental campaigns can be formulated to study specific reconnection issues in laboratory plasma experiments designed for other purposes, such as fusion experiments for the ion heating [e.g. 15,16] and for the global spontaneous relaxation [e.g. 25]. Basic plasma experiments, such as the Basic Plasma Science Facility (BaPSF) at UCLA, can also be used for the reconnection studies. The BaPSF has been utilized to study fundamental physics with relevance to reconnection, for example generation of solitary structures by electron beams [45] and the interaction of magnetic flux ropes [20]. The former Electric Tokamak at UCLA [46] has been modified with a novel plasma source to generate large volume, high beta, magnetized plasmas which could be used for investigation of a wide range of basic plasma phenomena as part of the BaPSF [47]. Given the potential for large plasma size and demonstrated plasma parameters (temperature, plasma beta, collisionality), this facility would be highly suitable for staging future campaigns on the physics of reconnection.

6. Relation with theory and modeling

As have been demonstrated in the past decade, theory and modeling has been playing increasingly important roles not only in understanding laboratory results but also in linking them to heliophysical observations. This will be even more true for the next generation reconnection experiments which should be guided by scoping simulation studies from their beginning. Therefore, a healthy theory and modeling program is also critical for the success of the next generation reconnection experiments.

References

[1] E. Zweibel and M. Yamada, Ann. Rev. Astron. Astrophys. **47**, 291 (2009).

[2] M. Yamada, R. Kulsrud, and H. Ji, Rev. Mod. Phy. **82**, 603 (2010).

[3] H. Ji et al., Phys. Rev. Lett. **80**, 3256 (1998).

[4] Y. Ren et al., Phys. Rev. Lett. 95, 055003 (2005).

[5] M. Brown et al., Phys. Plasmas **13**, 056503 (2006).

[6] J. Birn et al., J. Geophys. Res. **106**, 3715 (2001).

[7] F. Mozer, S. Bale, and T. Phan, Phys. Rev. Lett. **89**, 015002 (2002).

[8] Y. Ren et al., Phys. Rev. Lett. **101**, 085003 (2008).

[9] H. Ji et al., Geophys. Res. Lett. **35**, L13106 (2008).

[10] T. Carter et al., Phys. Rev. Lett. **88**, 015001 (2001).

[11] W. Fox et al., Phys. Rev. Lett. **101**, 255003 (2008).

[12] H. Ji et al., Phys. Rev. Lett. **92**, 115001 (2004).

[13] W. Daughton, Phys. Plasmas **10**, 3103 (2003).

[14] J. Egedal et al., Phys. Rev. Lett. **94**, 025006 (2005).

[15] G. Fiksel et al., Phys. Rev. Lett. **103**, 145002 (2009).

[16] M. Brown, US-Japan Workshop on Magnetic Reconnection (2009).

[17] F. Hansen and P. Bellan, Astrophys. J. **563**, L183 (2001).

[18] W. Bergerson et al., Phys. Rev. Lett. **96**, 15004 (2006).

[19] T. Intrator et al., Nature Phys. **5**, 521 (2009).

[20] E. Lawrence and W. Gekelman, Phys. Rev. Lett. **103**, 105002 (2009).

[21] E. Oz et al., to be submitted to Phy. Rev. Lett. (2010).

[22] D. Ryutov et al., Phys. Plasmas **13**, 032105 (2006).
[23] Y.-M. Lau and J. Finn, Astrophys. J. **350**, 672 (1990).
[24] J. C. Dorelli et al., J. Geophys. Res. **112**, A02202 (2007).
[25] J. Sarff et al., in The Magnetized Plasma in Galaxy Evolution, edited by K. T. Chyzy, K. Otmianowska-Mazur, M. Soida, and R.-J. Dettmar (Observatorium Astronomiczne, Uniwersytet Jagiellonski), pp. 48–55. (2005).
[26] N. Katz et al., Phys. Rev. Lett. **104**, 255004 (2010).
[27] P. Cassak, M. Shay, and J. Drake, Phys. Rev. Lett. **95**, 235002 (2005).
[28] Y.-M. Huang et al., arXiv:1010.5284 (2010).
[29] M. Øieroset et al., Phys. Rev. Lett., **89**, 195001 (2002).
[30] S. Krucker et al., Astrophys. J. **714,** 1108 (2010).
[31] A. Shih, R. Lin and D. Smith, Astrophys. J. Lett. **698**, L152 (2009).
[32] J. Drake et al., Nature **443**, 553 (2006).
[33] M. Oka et al., J. Geophys. Res. **115**, A08223 (2010).
[34] S. Tsuneta and T. Naito, Astrophys. J. Lett. **495**, L67 (1998).
[35] K. Shibata et al., Astrophys. J. Lett. **451**, L83 (1995).
[36] W. Daughton and V. Roytershteyn, submitted to Proceedings of ISSI meeting (2010).
[37] H. Ji and W. Daughton, to be submitted to Physics of Plasmas (2010).
[38] W. Daughton et al., Phys. Rev. Lett. **103**, 065004 (2009).
[39] A. Bhattacharjee et al., Phys. Plasmas **16**, 112102 (2009).
[40] M. Yamada et al., Phys. Plasmas **4**, 1936 (1997).
[41] M. Yamada, Phys. Plasmas **13**, 052119 (2006).
[42] M. Yamada, Phys. Plasmas **14**, 058102 (2007).
[43] J. Egedal et al., Rev. Sci. Instrum. **71**, 3351 (2000).
[44] A. Le et al., Geophys. Res. Lett. **37**, L03106 (2010).
[45] B. Levebvre at al., Phys. Rev. Lett. **105**, 115001 (2010).
[46] R.J. Taylor, et al., Nucl. Fusion **42**, 46 (2002).
[47] C.M. Cooper et al., Rev. Sci. Instrum. **81**, 083503 (2010).

Strengthening Heliophysics Through Coordinated Plasma Astrophysics Programs With Laboratory Plasma Physics and Astrophysics

H. Ji and S. Prager

Princeton Plasma Physics Laboratory, Princeton University

This white paper introduces scientific opportunities articulated by a grass-roots community activity, called Workshop on Opportunities in Plasma Astrophysics or WOPA. The WOPA activity was initiated in summer 2009, and the workshop was held in January 2010, the workshop report was released in October 2010. The WOPA participants were organized into 10 working groups, each focusing on a particular plasma process important to heliophysics and astrophysics. The membership of each working group is roughly equally divided between three communities (laboratory plasma physics. heliophysics, and astrophysics), between theory/modeler and experimenter/observer, and between institutions. 10 major plasma astrophysics questions emerged from this activity, and have been highlighted in the report. We note that 8 out of 10 plasma processes are directly relevant to heliophysics:

- Magnetic reconnection
- Collisionless shocks and particle acceleration
- Waves and turbulence
- Magnetic dynamos
- Interfacial and shear instabilities
- Angular momentum transport
- Magnetized dusty plasmas
- Radiative hydrodynamics,

and the opportunities on each of these processes are described in a chapter of the WOPA report. It was recommendation that the plasma astrophysics programs in the U.S. be strengthened in structure and coordination across DOE , NSF, and NASA, to embrace the unity, coherence, and opportunities of the field. A strengthened program of plasma astrophysics should greatly aid the missions of these agencies, including NASA Heliophysics Division. The full report can be downloaded from the WOPA website:

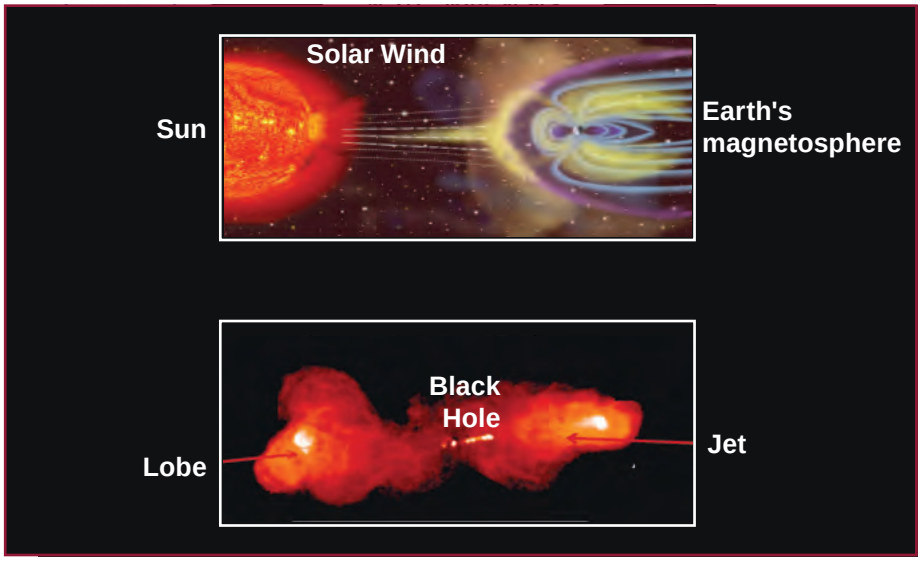
http://www.pppl.gov/conferences/2010/WOPA

and the Executive Summary is attached to this white paper.

EXECUTIVE SUMMARY

INTRODUCTION

Plasma pervades the universe at all measurable scales. At the very small scale, coupled processes in plasmas determine the behavior of the solar system. The Sun rotates, generates magnetic fields, and ejects mass in part because of plasma processes. The ejected plasma expands as the solar wind toward the Earth, becoming turbulent and hot as it travels. It then encounters and becomes trapped in the Earth's magnetic field, causes shocks, and produces magnetic substorms, aurora, and other plasma phenomena. This Sun-Earth system spans the short distance of 10^{-4} light years. Jumping ten orders of magnitude in size, extra-galactic jet systems are among the largest plasma structures in the universe. They begin with a rotating, accretion disk surrounding a supermassive black hole. Plasma transport processes determine the rate of accretion of matter onto the black hole, while producing the most luminous source of energy in the universe. The rotating plasma also launches a collimated jet that travels distances in the range of one million light years, ending in confined lobes of plasma. Dynamics of astrophysical systems at all scales between the solar system and jets are similarly regulated by plasma physics.



Plasma structures span all spatial scales in the universe, from the small scale of the solar system (shown in the top figure as an artist's sketch of solar wind spanning ~10^{-4} light years) to the large scale of extra-galactic jets, ~10^{6} light years (shown in the bottom figure from observation).

The study of plasmas beyond the Earth's atmosphere is here denoted as *plasma astrophysics*. This definition encompasses the usual realm of astrophysics (beyond the solar system), but also the domain of space physics (the Sun, the Heliosphere, and the magnetospheres of the Earth and the planets). The power of plasma astrophysics is that the same fundamental plasma processes appear in many different venues. For example, magnetic reconnection can drive not only magnetic substorms in the Earth's magnetosphere, but also flares on the surfaces of distant stars. The usual distinction between astrophysics and space physics disappears when viewed through the unifying lens of plasma physics.

Plasma astrophysics is positioned for rapid advance resulting from huge strides in astronomical observations, plasma technology, diagnostics, and plasma computation, combined with the maturity of plasma physics. Satellite and ground-based observations are set to measure plasma processes long invisible to us, from the solar interior to accretion disks. In-situ measurements of local plasma properties, including both field and particle information, have been expanded into every corner of our solar system: Earth's magnetosphere, the solar wind, other planets' magnetospheres, and the boundaries of our solar system. Multiple, coordinated satellites have greatly improved spatial and temporal resolution of magnetospheric measurements. Remote-sensing observations from both space and ground have moved beyond the traditional visible wavelengths to almost every wavelength band from far infrared emission from cold, partially ionized plasmas during star formation, to hard X-ray emission from extremely hot relativistic plasmas around supermassive black holes. High-power lasers now produce new plasma regimes with high energy density. These laser-produced, warm or hot, dense plasmas are similar to the interiors of giant planet cores and to the plasma that surrounds compact objects.

Plasma diagnostics can now measure, often remotely and non-perturbatively, a huge range of key particle and field quantities in the laboratory, both at the large scale and the small scale characteristic of turbulence. Modern techniques include laser scattering, laser-induced fluorescence, laser Faraday rotation, active spectroscopy using injected neutral atoms, miniaturized insertable probes, and electron cyclotron imaging techniques, to name a few. These provide new windows to detailed properties of magnetic fields, electric fields, electron and ion densities, plasma flow, and aspects of particle distribution functions. Advances in computation are revolutionizing how we study the complex behavior of plasmas. The surge in available computational power is being coupled with expansion of physics captured in computational models. Many plasma phenomena are governed by coupling between the large scale of the plasma system and the small scale characterized by microscopic plasma quantities (such as the particle gyroradius). New computational models can now treat this coupling, whether in multi-fluid treatments or new approaches that solve kinetic equations.

The opportunities represented by these technical advances can only be fully realized through a coordinated effort that brings together the communities of astrophysicists and laboratory plasma physicists. These communities involve observers, laboratory experimentalists, theorists, and computational physicists. Recent years have seen very significant beginnings of such coordinated efforts.

These beginnings indicate the large potential for accelerated progress and the need for an articulation of the major scientific challenges and opportunities in plasma astrophysics. Such an articulation would also express the unity and coherence of plasma astrophysics as a scientific discipline. Since it merges multiple areas of expertise, its unity can be overlooked, as reflected in the absence of a clear funding home for plasma astrophysics in the U.S.

To express the challenges, opportunities, and coherence of plasma astrophysics, more than 100 scientists were involved in preparing and participating in the Workshop on Opportunities in Plasma Astrophysics held in January 2010. The workshop, preceded by preparatory efforts of ten topical working groups, was a grass-roots effort organized by the plasma astrophysics community.

This effort brought together observers, experimentalists, computational plasma physicists, and theorists from universities, national laboratories, government research institutions, and private industry, including several scientists from outside the U.S. It also encompassed physicists studying magnetized plasmas and those studying high energy density plasmas. The breadth of participation uncovered cross-cutting opportunities previously unappreciated. This document reports the results from the workshop.

MAJOR QUESTIONS AND TOPICS

There are two approaches to articulating the challenges and opportunities: through plasma processes or through astrophysical systems. Individual plasma processes affect multiple systems, and individual systems encompass multiple processes. We have taken both approaches. Each of the ten working groups focused on one of the following processes that express the physics challenges (listed here in random order):

- Magnetic reconnection
- Collisionless shocks and particle acceleration
- Waves and turbulence
- Magnetic dynamos
- Interface and shear instabilities
- Angular momentum transport
- Dusty plasmas
- Radiative hydrodynamics
- Relativistic, pair-dominated and strongly magnetized plasmas
- Jets and outflows

Discussion of these topics, and their links to astrophysics, constitutes the bulk of this report. From these studies, we then extracted ten major, system-based questions for plasma astrophysics (listed here in random order):

How do magnetic explosions work?
Astrophysical plasmas exhibit spontaneous "explosions," such as solar flares, stellar flares, and substorms in planetary magnetospheres. These events accelerate particles to high energy, and affect radio communications on Earth. The explosions are driven by magnetic reconnection, the physics of which must be unraveled to understand why energy contained in magnetic fields of stars and planets is released explosively, and how this energy is so efficiently converted to particle energy.

How are cosmic rays accelerated to ultra-high energies?
Energetic particles bombard the Earth from space with energies up to 10^{20} electron volts, enormously more energetic than those achieved in the most powerful accelerator in the laboratory.

The energy spectrum of the particles fits a power law, yet the source of the strong acceleration remains a mystery. Possible sources include plasma shocks initiated by supernova explosions, as well as magnetic reconnection and plasma turbulence.

What is the origin of coronae and winds in virtually all stars, including the Sun?
Most stars have hot coronae with temperatures exceeding a million degrees. A hot wind blows from the Sun into the interstellar medium. A major challenge is how these nearly collisionless plasmas can be heated by waves, turbulence, shocks, and magnetic reconnection. This challenge requires advances in understanding the nature of the anisotropic turbulence in these plasmas.

How are magnetic fields generated in stars, galaxies, and clusters?
The universe appears to be magnetized at nearly all observable spatial scales. Stars, galaxies, galaxy clusters, and generally accretion disks, all contain magnetic fields. The fields often vary in time, sometimes in cycles. Understanding the origin and dynamics of these fields is a major puzzle, and is key to explaining accretion, stellar evolution, and galaxy evolution. The behavior of the fields depends on the plasma physics of dynamos, turbulence, reconnection, and flows.

What powers the most luminous sources in the universe?
The accretion of matter onto supermassive black holes generates prodigious fluxes of radiation, so intense that radiation pressure can dominate gravity. Understanding this process, which underlies the most luminous sources that light up the universe, is a challenge in plasma physics. Accretion is sufficiently rapid that it must involve plasma processes that enhance the accretion rate, such as plasma instabilities, turbulence, and transport.

How is star and planet formation impacted by plasma dynamics?
Despite progress in understanding star formation, the role of magnetic fields, especially in angular momentum transport, are still not understood. How rotating gas, plasma, and charged dust lose their angular momentum to collapse to form stars and planets remains an unsolved problem. Answers to this problem require advances in the physics of dusty plasmas, magnetic reconnection, and magnetized turbulence.

How do magnetic fields, radiation, and turbulence impact supernova explosions?
Supernovae and gamma-ray bursts are nature's grandest explosions that originate from the collapse of massive stars. Recent observations suggest that these explosions may be highly anisotropic and may be turbulent. Investigating the roles of magnetic fields, radiation, and turbulence on the explosions is a fundamental challenge. Magnetic fields are particularly important in the so-called magnetars.

How are jets launched and collimated?
Powerful jets of plasma are observed in a variety of astrophysical systems. They emanate from compact objects ranging from protostars to supermassive black holes, and can deposit large amounts of energy in the surrounding medium over large distances. Magnetic fields are believed to govern the process of jet formation and collimation. But it is not yet known how the jets are launched and why they survive stably across large distances (up to extragalactic scales). The behavior of jets encompasses a broad range of plasma phenomena, including instabilities, transport, dynamo, and reconnection.

How is the plasma state altered by strong magnetic fields?
When magnetic energy density is higher than the rest mass energy density of plasmas, the dynamics are altered significantly, as found in environments such as pulsars and relativistic jets. The ultra-strong magnetic field can produce electron-positron pairs, and alter plasma wind propagation and dissipation of plasma energy. Fundamental plasma processes, such as magnetic reconnection, are changed in these exotic environments.

Can magnetic fields affect cosmic structure formation?
Magnetic fields are observed in galaxy clusters, raising the question of whether the fields play a fundamental role in determining the topology and properties of these very large-scale structures. Magnetic fields may also be important in understanding the baryon dynamics in cosmic structure formation. These questions extend the applications of many of the basic plasma processes investigated at smaller scale to these new environments.

These questions are manifestly of major significance in astrophysics. What might not have been so obvious is that their answers largely reside in plasma physics. This report is dedicated to describing the challenges, opportunities, and impact of the following plasma research areas (listed here in random order):

Magnetic reconnection — a change in the topology of magnetic fields in plasmas — can drastically alter plasma transport and plasma structures, as well as convert magnetic energy to particle energy. Reconnection is thought to be everywhere in the universe, particularly underlying processes such as stellar flares.

Collisionless shocks accompany explosive events, such as supernova, and can accelerate particles to high energies. The behavior of shocks, and dissipation of their energy, in astrophysical plasmas in which collisions between particles are rare, is quite different and more complex than for the better-understood case with strong collisions.

Plasma turbulence is ubiquitous since most astrophysical plasmas have substantial free energy — such as nonuniformities in density, temperature, fields, and velocity space distributions — to excite turbulence. Turbulence affects the macroscopic behavior of plasmas in many ways, and has profound influences on essentially all the astrophysical questions articulated above.

The *dynamo* problem has long been a major astrophysical puzzle: how does the mechanical energy associated with a flowing plasma give rise to large scale magnetic field growth from a small seed field? This question is motivated by the prevalence of magnetic field in the universe. The dynamo problem also encompasses the questions of how magnetic field is sustained in the presence of dissipation, and why it varies in time in various astrophysical venues.

Interface and shear instabilities arise in both high energy density plasmas and magnetized plasmas in the presence of sheared flow. Their onset, behavior, and nonlinear development underlie a wide range of astrophysical phenomena including stellar wind flow around planetary and pulsar magnetospheres, the transitional region from solar wind to the interstellar medium, photoevaporated molecular clouds, supernova explosions, and blast waves in supernova remnants.

Momentum transport in plasmas is often observed to occur at a rate faster than can be explained from viscosity due to collisions between particles. Believed to be driven by plasma instabilities or turbulence, an explanation of momentum transport is required to understand how rotating matter collapses into gravitational potential wells. This question is key to understanding the accretion of matter onto compact objects (from protostars to black holes), the formation of stars, and the formation of planets.

Dusty plasmas consist of relatively massive charged dust particles, and thereby have properties altered from the conventional plasma of lighter ions and electrons. As breeding grounds for planets and stars, dusty plasmas have influence at both the solar system and galactic scales.

Radiation hydrodynamics — the interaction of radiation with matter — introduces an additional player (radiation) to the plasma environment. It is important for the formation of the largest stars (through photo-ionization and heating of the gas surrounding young stars), the behavior of accretion disks (when radiation pressure is substantial), the structure of molecular clouds (due to photoevaporation), the atmospheres of extrasolar planets (influenced by radiation from stars), and supernova explosions of massive stars (where radiation pressure exceeds material pressure).

Relativistic, pair-dominated and strongly magnetized plasmas are somewhat exotic systems that are ubiquitous in the high energy universe. When particles are moving at relativistic speeds, when electron-positron pairs dominate the plasma (rather than electrons and ions of unequal masses) and when the magnetic field is ultra-strong (altering collisions, ionization, radiation, and atomic structure), the behavior of the plasma changes significantly. These conditions can occur in violent astrophysical phenomena such as pulsar winds, gamma-ray bursts, jets in active galactic nuclei, microquasars, neutron star atmospheres, and the first few seconds of the early universe.

Jets and outflows represent a type of plasma structure whose formation and behavior challenges many aspects of plasma physics. They are observed to emanate from a wide range of compact objects, and might have a significant impact on cosmic structures.

These ten topics in plasma physics are separately described in the chapters that follow. The opportunities articulated in this report both advance research in plasma topical areas and lead toward resolution of the major astrophysical questions cited earlier.

IMPACTS AND CONCLUSIONS

These research opportunities will additionally have impact in three areas beyond these targeted major questions. First, much of the physics overlaps with central challenges for fusion energy, both magnetic and inertial (the purview of the DOE). For example, in magnetically confined plasmas, magnetic reconnection drives sawtooth oscillations and disruptions in tokamaks, momentum transport determines the rotation of toroidal plasmas, plasma turbulence regulates transport in nearly all magnetic fusion plasmas, and particle acceleration and heating by waves and instabilities occurs both deliberately and spontaneously. In inertially confined plasmas, Rayleigh-Taylor instabilities influence the symmetry of implosion and radiative hydrodynamics influences the burn phase. Many of the advances in plasma astrophysics have evolved directly from plasma physics developed for fusion energy, and plasma physics learned through astrophysical applica-

tions is increasingly influencing aspects of fusion plasma physics. Second, development of plasma physics targeted to astrophysics advances basic plasma physics (the purview of the NSF and DOE). The wide range, sometimes extreme, of scales, particle energies, field strengths, and overall plasma parameters encountered in astrophysics extend greatly the scope and depth of plasma physics. Third, plasma astrophysics is crucial to the guidance and interpretation of observational missions (the purview of NASA). From satellite observations of the magnetosphere to the heliosphere and other galaxies, guidance from plasma astrophysics is needed to determine what quantities to measure and to make scientific sense of those measurements. And for some missions, the primary goal is to directly study a particular plasma process. That is, plasma astrophysics is needed in the preparation stage of missions, to optimize its value, and in the operational stage, to reap the benefits. For example, understanding aspects of magnetic reconnection, waves, and turbulence is essential to the success of the Magnetospheric Multiscale Mission, the Solar Dynamics Observatory, the Solar Orbiter, and the Solar Probe Plus. Understanding aspects of accretion and particle acceleration is key to reap the full benefit of the Nuclear Spectroscopic Telescope, the James Webb Space Telescope, Large Synoptic Survey Telescope, Atmospheric Čerenkov Telescope Array, and the International X-ray Observatory.

The report articulates many general and specific research opportunities. The content of each of the following chapters was written by a topical working group (see Appendix D for a list of the membership of each working group). Being a community report, no effort is made to rank the many opportunities that for each area were developed by the specific topical experts. That exercise is beyond the province of this activity. Thus, this report produces no recommendations except one. We recommend that the plasma astrophysics program in the U.S. be strengthened in structure and coordination across DOE, NSF, and NASA, to embrace the unity, coherence, and opportunities of the field. A strengthened program of plasma astrophysics greatly aids the missions of these agencies. One intention of this report, in addition to the immediate scientific value of the effort, is to provide motivation and justification for deeper consideration of the funding strategy for plasma astrophysics.

SOLAR SAIL PROPULSION: ENABLING NEW CAPABILITIES FOR HELIOPHYSICS

L. Johnson[1], R. Young[1], D. Alhorn[1], A. Heaton[1], T. Vansant[2], B. Campbell[2], R. Pappa[3], W. Keats[3], P. C. Liewer[4], D. Alexander[5], J. Ayon[4], G. Wawrzyniak[6], R. Burton[7], D. Carroll[7], G. Matloff[8], and R. Ya. Kezerashvili[8]

[1] NASA Marshall Space Flight Center, Huntsville, AL
[2] NASA Goddard Space Flight Center, Greenbelt, MD
[3] NASA Langley Research Center, Hampton, VA
[4] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA
[5] Rice University, Houston, TX
[6] Purdue University, West Lafayette, IN
[7] CU Aerospace, Champaign, IL
8 New York City College of Technology, CUNY, New York, NY

ABSTRACT

Solar sails can play a critical role in enabling solar and heliophysics missions. Solar sail technology within NASA is currently at 80% of TRL-6, suitable for an in-flight technology demonstration. It is conceivable that an initial demonstration could carry scientific payloads that, depending on the type of mission, are commensurate with the goals of the three study panels of the 2010 Heliophysics Survey. Follow-on solar sail missions, leveraging advances in solar sail technology to support Heliophysics Survey goals, would then be feasible. This white paper reports on a sampling of missions enabled by solar sails, the current state of the technology, and what funding is required to advance the current state of technology such that solar sails can enable these missions.

INTRODUCTION

Solar sail propulsion uses sunlight to propel vehicles through space by reflecting solar photons from a large (> 100 meters per side), mirror-like sail made of a lightweight, highly reflective material. The continuous photonic pressure provides propellantless thrust to perform a wide range of advanced maneuvers, such as to hover indefinitely at points in space, or conduct orbital plane changes more efficiently than conventional chemical propulsion. Eventually, a solar sail propulsion system could propel a space vehicle to tremendous speeds—theoretically much faster than any present-day propulsion system. Since the Sun supplies the necessary propulsive energy solar sails require no onboard propellant, thereby significantly increasing useful payload mass.

Practical concepts for solar sailing have existed for approximately 100 years, beginning with Tsiolkovsky and Tsander in the 1920s. A team at JPL completed the first serious mission study in the late 1970s for a rendezvous with Halley's Comet (Friedman, 1978). An effort by McInnes in the 1990s and the publication of his PhD dissertation as a textbook on solar sailing helped re-invigorate interest in solar sailing as a research topic (McInnes, 1999). In the early-to-mid 2000's, NASA's In-Space Propulsion Technology Project made substantial progress in the development of solar sail propulsion systems. Two different 20-m solar sail systems were produced and successfully completed functional vacuum testing in the Glenn Research Center's (GRC's) Space Power Facility at Plum Brook Station, Ohio (Johnson, 2007). Solar sail propulsion was then selected as a candidate for flight validation by NASA's New Millennium Program on their proposed Space Technology 9 (ST-9) flight. Unfortunately, funding for both the In-Space Propulsion Technology Project and the New Millennium Program were terminated shortly thereafter due to NASA's changing priorities toward Project Constellation.

Outside of NASA, solar sailing has been tested in space. In the summer of 2010, the Japanese Aerospace Exploration Agency, JAXA, launched a solar sail spacecraft named IKAROS in tandem with another mission to Venus. The sailcraft IKAROS (14 m by 14 m) is the first in-flight demonstration of solar sailing (Tsuda, 2010). While the effects of solar radiation pressure (SRP) are smaller on this sailcraft as compared to other concepts for solar sails, numerous program objectives have been achieved, including verifying solar radiation pressure (SRP) effects on the sail and performing in-flight guidance and navigation techniques using the solar sail.

Two programs based in the United States aim to deploy their own solar sails in space before the end 2010.

NASA's *NanoSail-D*, slated to launch November 2010, has two mission requirements: (1) to successfully stow and deploy the sail and (2) to demonstrate deorbit functionality. The Planetary Society's *Lightsail-1* (5.65 m by 5.65 m), is of similar configuration to *NanoSail-D*. *LightSail-1* is a combination of three 10 cm cubesats intended to demonstrate deployment in low Earth orbit. Funded by NASA's Small Business Innovative Research, CU Aerospace and The University of Illinois are developing the *CubeSail* (Burton et al, 2010), which consists of two nearly-identical cubesat satellites to deploy a 250-m long, $20m^2$ sail. The design can be scaled to build a heliogyro solar sail that can potentially achieve square kilometer sail areas.. The *CubeSail* is not yet scheduled for flight.

Because of the continuous force provided by solar radiation pressure on a solar sail, solar sail spacecraft can fly in non-Keplerian orbits and can continually maneuver throughout flight without the use of finite propellant. The first section of this white paper identifies key solar and heliospheric missions proposed by in recent years that are enabled by solar sail propulsion. The next section provides more details on the current state of solar sailing technology and what improvements are required to advance solar sail technology to levels necessary to accomplish sail-enabled solar and heliospheric missions. The paper concludes with funding recommendations to achieve those necessary advances.

SCIENTIFIC MISSIONS ENABLED BY SOLAR SAILS

A number of science mission concepts have been identified that make optimum use of solar sail technology as the next phase in the development of solar sail propulsion as the go-to technology for high C3 missions. We briefly highlight four concept missions leaving the main details to the cited references. These are Solar Polar Imager (see also white paper by Liewer et al.), GeoSail (McInnes et al. 2001), Heliostorm, and Interstellar Probe (Mewaldt et al 2001). Each of these mission concepts have been envisaged assuming the successful development of the relevant technological capability and have significantly detailed science goals of relevance to heliophysics and that are not generally achievable without access to reliable solar sail propulsion and control. Table 1 summarizes the key parameters of these missions.

Table 1. Solar sail mission concept designs.

Mission	Sail dimension	Sail areal density	Orbital/Trajectory parameters	Primary Science Objective
Solar Polar Imager	150 – 180 m on a side	8.5-14 gm^{-2}	Heliocentric with semi-major axis of 0.48 AU at inclination of 75 degrees	Convective flows in polar regions and the deep solar interior
Heliostorm	100m x 100m	~15 gm^{-2}	Sun-centered orbit at artificial Lagrange point (0.967 AU from Sun)	Advanced Early Warning of Geo-effective solar events
GeoSail	41.2m x 41.2m	~33 gm^{-2}	Earth Orbit: 11R_E Perigee, 23 R_E Apogee; Period of 4.5 days	Extended presence in Earth's geotail using sail to precess inertial orbit
Interstellar Probe	~400m diameter	~1 gm^{-2}	Solar swing by (0.25 AU) then radial trajectory at ~15AU/yr (sail jettisoned at 5 AU	Measurements of the nearby interstellar medium

A concept that uses a hybrid low-thrust and solar sail propulsion system has been proposed to orbit above one of Earth's poles, providing a continual view of that pole. This Earth Polesitter mission could be used to continually monitor polar weather, observe the interaction between solar wind and the Earth's magnetic field in the Earth's atmosphere, and perform geological surveys at high latitudes (Ceriotti and McInnes, 2010).

Additionally, the observation of the trajectory of a solar sail by itself may provide a test of the fundamental

principles of general relativity. The curvature of spacetime, in conjunction with solar radiation pressure, affects the bound orbital motion of solar sails and leads to deviations from Kepler's third law for heliocentric and non-Keplerian orbits, as well as to the new phenomenon for non-Keplerian orbits when the orbital plane precesses around the sun that is an analog of the Lense-Thirring effect (Kezerashvili and Vazquaz-Poritz, 2009; 2010).

CURRENT TECHNOLOGY AND NEAR-FUTURE IMPROVEMENTS

NASA's In-Space Propulsion Technology Project funded the development of two prototype solar sail systems for ground testing. These prototypes were used to validate design concepts for sail manufacturing, packaging, launch and deployment in space; attitude control subsystem function; and characterization of the structural mechanics and dynamics of the deployed sail in a simulated space environment. Two square sail architectures were developed. They consisted of 4-quadrant, reflective sail membranes, a deployable sail support structure, an attitude control subsystem, and hardware necessary to stow the sail for launch. Additionally, these systems were required to meet the characteristics given in Table 2. A sub-L1 solar monitoring mission concept was provided as a reference mission for guidance in design and scalability issues.

NASA awarded two ground demonstration contracts. ATK Space Systems' approach incorporated their rigid coilable boom, a sliding mass attitude control system (ACS) subsystem, and partner SRS' CP1™ sail membrane (Murphy, 2005). L'Garde Inc. used their inflatable and cold-temperature rigidizable boom, a control vane based ACS, and Mylar for the sail membrane (Lichodziejewski, 2006). The parallel testing and development of these two system-level demonstrations that have varied technologies in the three major components removed the risk to this technology development if one provider encountered an unrecoverable failure. The system-level ground demonstration work was divided into three phases. A 6-month concept refinement phase was completed in May 2003. During this phase, the two teams provided analysis of their system's performance when scaled to the Design Reference Mission, and a preliminary test plan for the following two 12-month phases. The first 12-month hardware development phase began in June 2003. In this phase both teams built and tested components and subsystems, with ATK concentrating on a single 10-m quadrant, and L'Garde developing a 10-m square sail. The most comprehensive of these tests occurred in mid-2004 when the respective teams deployed their integrated subsystem in the NASA Langley Research Center (LaRC) 14-m vacuum facility (ATK), and the 30-m vacuum chamber at GRC's Plum Brook Space Power Facility (L'Garde). Following this phase, the teams culminated their work in a 12-month system verification phase. In this phase both teams built and tested fully integrated 20-m sail systems that included a launch packaging container and operational ACS subsystems. In mid-2005, the respective teams tested their systems in the Plum Brook Facility under a high vacuum and appropriate thermal environment, and subjected their systems to launch vibration and ascent vent tests. Figure 1 shows the 20-m deployed systems at the Plum Brook Facility. Since the sails represent the largest systems that could be tested in a vacuum chamber on the ground, a significant effort was made to collect static and dynamic data on the sails and booms with ~400 Gb of data collected, primarily raw photogrammetric data.

Table 2. Characteristics of system-level solar sail ground demonstrations

Metric	RFP	ATK	L'Garde
Dimensions	20 meters × 20 m or greater	• 20-m system with flight like central structure • 4 sails scaled from 80 m • Truncated 80-m masts • Central structure scaled from 40 m	• 19.5 m due to Plumbrook • 1 subscale TVCAD vane • No flight central structure scaled for 100-m system • Sails and mast truncated 100-m system
Sail Subsystem Areal Density	< 20 g/m^2 (scalability to 12 g/m^2 for 104 m^2)	• 112 g/m^2 — includes spacecraft bus structure, ACS, power, instrument boom • Scaled to 11.3 g/m^2 for 100-m design and no payload	• 30 g/m^2 – includes ACS (4 vanes calculated), central structure dropped • Scaled to 14.1 $g/m^2$with 50-kg payload and 41.4-kg bus

Stowed Volume	< 0.5 m^3 (scalability to 1.5 m^3 for 104 m^2)	• 0.9 m^3 scaled to1.5 m^3 for 100-m design	• 2.14 m^3 scaled to 1.04 m^3 for 100-m design
Thrust Vector Turning Rate About Roll Axis	> 1.5°/hr	• > 35° maneuver in 2 hrs	• 63°/hr (.0175°/sec)
Effective Sail Reflectance	> 0.75	• 92% over solar spectrum	• 85.9
Anti-sunward Emissivity	> 0.30	• 0.30 for 3-μ film	• 0.40
Membrane Characteristics	Space-durable, tear-resistant, designed for 1 yr in the near-GEO environment	• ~2-μm CP1™ with 1,000 Å of aluminum on front, uncoated CP1™ on back of sail. All materials have space flight heritage	• 2-μm Mylar with 1,000 Å of aluminum on front and 200 A black chromium on back
System Flatness	Effective for propulsion	• 3-point quadrant support with shear compliant border to insure a flat sail, with a proper stress level to obtain local flatness	• Striped net loss ~ 2 %
ACS	3-axis, minimize propellant usage	• Primary: Sliding trim control masses on truss for pitch and yaw control and rotating tip bars for pinwheel control of roll. Secondary: Micro PPT thrusters at boom tips.	• Primary: propellant-less using four tip vanes

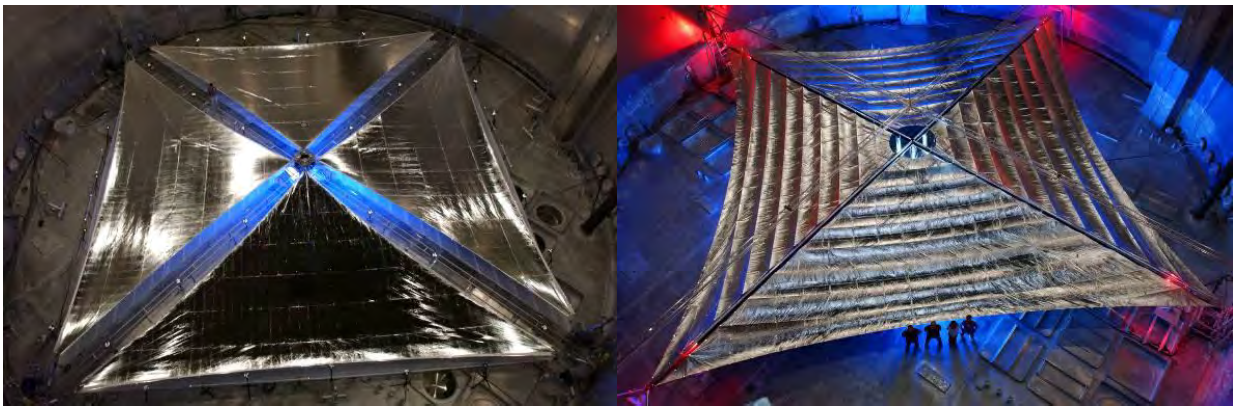


Fig. 1. Two 20m x 20m solar sail prototypes are shown — ATK (left) and L'Garde (right).

Integrated Software Tools: The In-Space Propulsion Technology Project also funded the development of two sets of integrated simulation tools to predict the trajectory, maneuvers, and propulsive performance of a solar sail during a representative flight profile, one by Princeton Satellite Systems (Thomas et al., 2004) and another by the Jet Propulsion Laboratory (Ellis et al., 2004). The tools were designed to be integrated into an optimal guidance and navigational control (GNC) subsystem on a future flight mission. The tools incorporated the following analytical models: 1) Solar radiation pressure acting on the sail as a function of sail orientation and distance from the Sun, 2) Disturbance forces acting on the sail such as gravitational torques and thermal deformation of the support structure, 3) Orbital mechanics, 4)Sail structural dynamics, 5) Attitude control system dynamics, and 6) Navigational sensors.

As part of the Ground System Demonstrations, LaRC, ATK, and L'Garde created Finite Element Models (FEMs) of the 10-m systems using both commercial off-the-shelf (COTS) and custom software. The purpose of

this task was to create methods/algorithms/techniques that improved the conventional FEM of the membranes, booms, and other subsystems. Several methods were identified, but the option phases to complete and validate them were not funded, due to the success of the prime contractors in modeling the 20-m systems with conventional techniques, thus lowering the priority of this task below the level of the available funding.

Optical Diagnostic System: The requirements for this task included the development of methods for observation of the sail deployment, and monitoring of the health and integrity of the sail during and after deployment. During operation, the ODS would be available to provide shape and vibration measurements adequate enough to infer the stress state of the solar sail by aid of computational structural models, which could then feed real time into a closed loop spacecraft Guidance, Navigation and Control (GNC) system. The difficulty of photogrammetry was not insurmountable, but it did indicate that the challenges identified in the conceptual design phase were real. Combined with a better understanding of the lack of a need for on-orbit photogrammetry, the further development of a flight ODS was not pursued.

Structural Analysis and Synthesis Tools: This task also addressed the structural analysis issue by developing from the ground-up new and unconventional modeling techniques. Techniques considered included, Direct Transfer Function Modeling (DTFM) and Parameter Variation Processing (PVP). The overall objectives for this task were; completion of DTFM modeling/analysis methods for long booms, completion of capability to evaluate effects of imperfections, completion of PVP method for analyzing wrinkled membranes, and completion of test/analysis correlation by using existing test data. As in the above task, this effort was terminated prior to completion due to success using conventional modeling methods with the 20-m systems and a lack of funding.

Lightweight Attitude Control System: The objectives of this task were; (1) to design, integrate, and test a sail attitude control system (SACS) employing a two-axis gimbaled control boom, and (2) to develop a high-fidelity, multi-flexible body model of ATK's solar sail for the purpose of validating a thrust vector control (TVC) concept employing a two-axis gimbaled control boom. One of the major findings from this study was that the two-axis gimbaled control boom was not a mass efficient method of controlling a sail. A more efficient method was derived based on an offset mass moved along the booms by a clothesline-like apparatus to control pitch and yaw, and rotating stabilizer bars at the sail tips to pinwheel the sail quadrants for roll control. This finding led to a major redesign of the ATK 20-m hardware to accommodate the new TVC concept. An attitude determination and control block diagram were derived to present the application/integration of the inertial stellar compass with a range of ACS options from cp/cm offset to pulsed plasma micro thrusters (Thomas et al., 2005).

Characterization of Candidate Solar Sail Materials: The purpose of this task was to conduct laboratory characterization of several candidate solar sail materials. The space radiation and micrometeoroid environments for 1.0 and 0.5 astronomical units missions were defined and candidate materials were tested against these radiation and meteoroid environments. Through a series of learning tests, the sample hold-down design was optimized and a flexure test developed. Several samples of the SRS and L'Garde membrane materials were tested by subjecting them to gigarad levels of radiation in simulations of long duration solar wind mission type (Edwards et al., 2004). While some of the samples showed significant levels of degradation in mechanical strength, solar sail loading is so low, very little strength is needed.

Advanced Manufacturing Technologies: The purpose of this task was to investigate and develop an integrated approach to ultrathin film solar sail manufacturing. The focus was on; improving coating processes and technologies, developing sail seaming technologies for large monolithic sails, providing an integrated approach to membrane coating, acceptance, assembly and integration, and integrating future improvements into the process such as electrospun nanofibers for ripstop enhancement without added mass, and the addition of carbon black nanotubes to the sail backside to increase emissivity. The final results of this 2-year effort are the development of a scroll coating system, the development of coating capabilities of less than 2.5 μ, and the development of a membrane seaming system able to form monolithic sails with coatings as thin as 2.5 μ.

Smart Adaptive Structures: In order to mature the TRL of solar sail propulsion, advancements must be made in the pointing and dynamical control of these large space structures. This tasks' objective was to develop and verify structural analytical models, develop structural scaling laws, and develop adaptive control laws for solar sails to be verified on a >30-m vertically supported boom.

Sail Charging: Due to two extreme characteristics of future solar sail missions, the large surface area of the

sail, and the long duration of potential missions, typical spacecraft charging issues will be exacerbated. The purpose of this task was to characterize charged particle environments for analyzing solar sail charging in the solar wind and at geostationary orbit, and to model surface and internal electric fields and potentials for solar sails using existing spacecraft charging models. Solar sail materials were tested in simulated charging environments to determine permeability and charge retention properties. A significant finding was that there will be little charging of the sail surfaces, ~10 v as a worst case in sunlight. The study found that problems arise if the sail material backing is non-conductive or electrically decoupled from the front surface. In that case, the shadowed back surface can reach potentials of –30 to –40 v relative to the space plasma in the solar wind on the order of arcing onset potentials. The solution is to make sure the sail material is conductive front to back and end-to-end if the sail is to be in geosynchronous orbit or in the auroral zone, and be careful with electrically isolated objects in the shadow of the sail.

Long-Term Space Environmental Effects: Critical to the development of solar sails is an investigation of space environmental effects on these large thin film materials and the edge support technologies. This task was related to the "Characterization of Solar Sail Materials" task above that used accelerated dose levels over a shorter period of time to simulate the total dose of radiation received by a material for many years. The purpose of this task is to provide critical thermal, optical, mechanical, and surface data on large sails taking into account edge stresses and edge support technologies that can only be characterized using large size sails but not at accelerated levels. These resulting test data could be used to validate the accelerated dose test methodology regarding the durability of candidate sail material (embrittlement, optical, mechanical, surface, and thermal properties).

FUNDING RECOMMENDATIONS

Taking solar sail propulsion technology from ground testing to flight validation and mission implementation is one of the top ten priorities described in the NASA Office of Chief Technologist's draft In-Space Propulsion Systems Technology Roadmap. The next step in making this technology available for use in future missions is flight validation of a full-scale (> 1000m^2) solar sail propulsion system. Based on the NASA ST-9 mission concept study, the development cost for such a system would be less than $200M (not including the cost of launch or mission operations). In addition to funding for solar sail technology programs, we also advocate funding on the order of $250K to $500K per year (5 to 10 grants per year) to be distributed to NASA centers and universities for solar sail mission design. Solar sail trajectory analysis leads to new mission applications, uncovering new regimes for heliophysics scientific exploration. This grant program would stimulate new research into solar sail missions and encourage growth and development of new methods and mission analysis tools for solar sail trajectory, navigation, and attitude control analysis. Assuming 3-year grants, the total program cost for this additional research would be on the order of $0.75 to 1.5M.

CONCLUSIONS

The TRL of solar sailing technology continues to advance. It has been confirmed by flight testing in deep space that Solar Sail technology is viable for space flight operations. Many important applications of solar sails have been identified as useful to the international science community, especially missions significant to the goals of the 2010 Solar and Heliophysics Decadal Survey. Funding for solar sail technology advancement, including continued examination of solar sail mission applications, is highly recommended as an outcome of the 2010 Decadal Survey.

REFERENCES

Burton, R., Coverstone, V., Swenson, G., et al., "Initial Development of the CubeSail/UltraSail Spacecraft," Joint Army Navy NASA Air Force (JANNAF) Spacecraft Propulsion Subcommittee Meeting, Colorado Springs, Colorado, May, 2010.

Edwards, D., et al., "Characterization of Candidate Solar Sail Material Exposed to Space Environmental Effects," *42nd AIAA Aerospace Sciences Meeting and Exhibit*, Jan. 2004, AIAA 2004–1085.

Ellis,J. et al., "A Solar Sail Integrated Simulation Toolkit," *Proc. AAS/AISS Spaceflight Mechanics*, Maui, HI, Feb., 2004, AAS 04-283.

Friedman, L., "Solar Sailing: The Concept Made Realistic," AIAA-78-82, 16th AIAA Aerospace Sciences Meeting, Huntsville, Alabama, January 1978.

Kezerashvili, R. Ya. and Vazquaz-Poritz, J.F., "Solar radiation pressure and deviations from Keplerian orbits," Physics Letters B 675, 18–21 (2009).

Kezerashvili, R. Ya. and Vazquaz-Poritz, J.F., "Bound orbits of solar sails and general relativity," Advances in Space Research 46, 346–361 (2010).

Lichodziejewski, D., et al., "Vacuum Deployment and Testing of a 20-m Solar Sail System," *47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, May 2006, AIAA 2006–1705.

Liewer, P. C. D. Alexander, J. Ayon, A. Kosovichev, R. Mewaldt, D. Socker, A. Vourlidas, "Solar Polar Imager: Observing Solar Activity from a New Perspective" in *NASA Space Science Vision Missions*, (ed. Marc S. Allen), American Institute of Aeronautics and Astronautics, Progress in Astronautics and Aeronautics, Vol. 224 (2009) pp. 1-40.

Liewer, P. C., R. A. Mewaldt, J. A. Ayon, C. Garner, S. Gavit, and R. A. Wallace, "Interstellar Probe using a Solar Sail: Conceptual Design and Technological Challenges," in *The Outer Heliosphere: The Next Frontiers*, (K. Scherer, Horst Fichtner, Hans Jörg Fahr, and Eckart Marsch, eds.) COSPAR Colloquiua Series, 11. Amsterdam: Pergamon Press p. 411 (2001).

MacDonald, M. and C. McInnes, "Solar Sail Mission Applications and Future Advances," in *Proceedings of the Second International Symposium on Solar Sailing*, New York, July 2010.

McInnes, C., *Solar Sailing: Technology, Dynamics, and Mission Applications*, Springer-Praxis series in space science and technology, Chichester, UK, 1999.

McInnes, C. R., M. Macdonald, V. Angelopoulos, and D. Alexander, "*GEOSAIL: Exploring the Geomagnetic Tail Using a Small Solar Sail*," J. Spacecraft and Rockets, 68, 622, 2001.

Mewaldt, R. A. and P. C. Liewer, "Scientific Payload for an Interstellar Probe mission," in *The Outer Heliosphere: The Next Frontiers*, (K. Scherer, Horst Fichtner, Hans Jörg Fahr, and Eckart Marsch, eds.) COSPAR Colloquiua Series, 11. Amsterdam: Pergamon Press , p. 451 (2001).

Murphy, D., "Validation of A Scalable Solar Sailcraft," *53rd JANNAF Propulsion Meeting*, December 2005.

Johnson, L., Young, R., and Montgomery, E., "Recent Advances in Solar Sail Propulsion Systems at NASA," *Acta Astronautica* 61 (2007), pp. 376-382.

The Planetary Society, "What We Do: Projects: LightSail – Solar Sailing," http://www.planetary.org/programs/projects/innovative_technologies/solar_sailing/, accessed Nov 5, 2010.

Thomas, S., Paluszek, M., Wie, B., Murphy, D., "Design and Simulation of Sailcraft Attitude Control Systems Using the Solar Sail Control Toolbox," *AIAA Guidance, Navigation and Control Conference*, Aug. 2005, AIAA 2004–4890.

Thomas, S., M. Paluszek, B. Wie, D. Murphy, "ACS Performance and Stability Validation for a 160-m Solar Sail with Control-Structure Interactions," *41st AIAA Joint Propulsion Conference*, July 2005, AIAA 2005–3926.

Tsuda, Y., et al., "Flight Status of IKAROS Deep Space Solar Sail Demonstrator" IAC-10-A3.6.8, *61st International Astronautical Congress*, Prague, Czech Republic, September 2010.

"**Measuring magnetic free energy in the solar atmosphere**"

Philip G Judge,
High Altitude Observatory,
National Center for Atmospheric Research.
judge@ucar.edu

Abstract: The purpose of this white paper is to draw attention to techniques that have become available which hold promise for the measurement of the free energy in the Sun's corona. These techniques avoid problems associated with extrapolations of photospheric vector magnetic fields by making measurements directly at the coronal base. If this problem were easy, it would have been done decades ago. Nevertheless, the time appears right to pursue new spectropolarimetric capabilities offered by Fabry-Perot instruments operating at near infrared wavelengths, at existing telescopes as well as the upcoming ATST. Success is not guaranteed, but the relatively low risks might be worth the modest investment which would enhance the value of missions such as SDO and also place the problem of the origins of space weather on a firm observational basis.

The problem

The interplanetary and terrestrial environments are governed by the variable solar wind, solar energetic particles and UV/X-ray radiation, coronal mass ejections and flares. All of these phenomena are caused by the release of free energy contained in solar magnetic fields. It is accepted that the interplanetary and terrestrial perturbations originate from magnetic energy stored largely in the corona. This is because it is hard to store significant amounts of energy below the corona, and because the release of energy in the denser solar atmosphere is too slow to account for flares and other dynamic events. Thus in the 1960s it was recognized already that flare energy release required storage and release in the coronal plasma (see the papers in "The Physics of Solar Flares", NASA-SP50, 1964). In 1973, coronagraphs on the Skylab mission allowed scientists to identify for the first time coronal mass ejections as major perturbers of the heliosphere. Later spacecraft such as OSO-8, SMM, Yohkoh, SOHO, RHESSI, STEREO, Hinode and now SDO have all carried instruments to study the dynamic release of magnetic energy by the Sun as it evolves through its magnetic activity cycle.

These experiments have provided the observational basis for the huge community efforts to try to understand the magnetic energy storage and release. It is well-known that information on the free component of the magnetic energy – associated with electrical currents flowing through the corona and cooler atmospheric components - can only be measured by the *vector* magnetic field, not just the line of sight component that has been measured since the pioneering work of Hale. However, it was not until the 1980s that the first credible measurements of the vector field were obtained, even in magnetic elements in the very bright solar photosphere where the fields are strong (> 1 kG). This is because the "transverse" components of the magnetic field make themselves known spectrally only as a second order term in the (usually small) Zeeman parameter (=Zeeman splitting/Doppler width), whereas the line of sight component is a first order term. The transverse fields make their presence known through the linearly polarized components ("Q" and "U") in spectral lines, the line of sight fields are reflected in the circularly polarized ("V") components of the Stokes vector (IQUV, I=intensity).

In principle, measurements of the magnetic field in the photosphere can be extrapolated outwards into

the corona, and these extrapolations should be able to yield the free energy and its evolution and release in the corona. Indeed, one of the goals of the HMI instrument on SDO is to provide global photospheric magnetic field vectors every few seconds for the lifetime of the mission. However, there are fundamental difficulties in using photospheric measurements for studying magnetic free energy. Most notably, the photospheric magnetic field is not close to a "force– free" state, meaning that the magnetic field there is forced by the dynamics of the turbulent convection. Thus, to obtain meaningful extrapolations, a full magneto-hydrodynamic calculation is in principle required. Such calculations are possible, but the physics of the intervening layers between the photosphere and corona, namely the chromosphere and transition region, is poorly understood and can significantly effect the extrapolations. A recent example of a difficulty is evident in the very different states of the magnetic field found by Arber and colleagues (2007) when they simulated magnetic flux emergence through the chromosphere using different assumptions of the microphysics of ion-neutral collisions in the chromosphere. These and other difficulties have led researchers to adopt a simpler approach, in which observations and models are combined judiciously to try to make photospheric data consistent with a simpler (though still challenging) non-linear force-free calculation. A notable example is the well-known nonlinear force free field project (NLFFF, Schrijver et al. 2008; Wiegelman et al. 2008). Unfortunately this approach has essentially failed despite the attempts of several groups to produce credible, reproducible results (de Rosa et al., 2009). A fundamental problem with the approach is that the measured boundary conditions at the photosphere must be altered to be consistent with the calculations themselves. This is, alas, tantamount to changing the data to match the physical conditions assumed in the model.

Measuring magnetic fields at the coronal base

To circumvent difficulties with using photospheric measurements, the obvious solution is to make measurements at the top of the chromosphere- the base of the corona. Here, the plasma is both close to a force-free state and yet dense enough to contribute to spectral features sensitive to the magnetic fields there. The problem is that these are not easy measurements to make or interpret. Spectral features formed in these layers are limited to vacuum UV, radio, and isolated spectral lines in the visible and IR regions. Significant progress has been made at radio wavelengths, but the heights of formation ("τ=1") surfaces are not easily determined, the angular resolution is limited by the long wavelengths, and there are physical limitations imposed by Bremmstrahlung opacity, for example. Vacuum UV measurements are insensitive to the Zeeman effect (small Zeeman parameter). The Hanle effect has been proposed for both UV and visible chromospheric lines, a technique which should be pursued but which is not yet a proven method for chromospheric work. More direct constraints on the free energy in the solar atmosphere are however beginning to be made using chromospheric observations (e.g. Socas Navarro 2005, Solanki et al 2003, Judge et al 2009). Here I suggest that the community consider building instruments devoted to *imaging* s*pectropolarimetry of infrared lines* which already have a demonstrated capability based upon *slit* spectropolarimetry. Imaging capability is critical to provide information on plasma velocity vectors, as discussed below. Such instruments may yield the needed measurements to measure the free energy in the corona, because:

1. In and around active regions, where the effects of magnetic free energy release are most obvious, the magnetic fields are strong enough to control the upper chromosphere's morphology and dynamics. The field there is thus in a low-β state (close to force-free), and the plasma dynamics is essentially a passive tracer of the magnetic field.
2. Such magnetic fields are not only compatible with force-free extrapolations, but also one can

apply Chandrasekhar's virial theorem to derive the free energy in the overlying coronal volume from measurements only at the chromospheric surface, subject to some reasonable assumptions discussed below.

But difficulties naturally arise. The prospect of inferring magnetic fields in the presence of obviously complex thermal structure and dynamics in the chromosphere (e.g. Judge 2010) is indeed daunting, and it is generally believed that vector polarimetry in the chromosphere must be difficult in the extreme. These perceived difficulties have led, in part, to the situation we are in today, in which we are reduced essentially to looking at the *effects* of the free energy release through instruments on the historic fleet of spacecraft launched by NASA and in ground-based observatories. Thus in spite of these efforts, we are not really getting to the root of the problem.

However the time appears right for a concerted effort to measure chromospheric magnetic fields, because of recent advances. Firstly, imaging spectropolarimeters based upon Fabry-Perot interferometers ("FPIs"), e.g. IBIS (Cavallini, 2006), TESOS (Tritschler et al. 2002), CRISP (Scharmer et al. 2008), and GFPI (Bello Gonzalez & Kneer 2008), have the demonstrated capability to perform the chromospheric measurements of close to the required sensitivity. Secondly, tools are now available for the reliable interpretation of a valuable line of neutral helium at 1083 nm.

A way forward

By good fortune, the He I 1083 nm multiplet (actually a triplet) forms in the uppermost chromospheric features (be they spicules, canopy-like fibrils, or the uppermost scale height of the stratified chromosphere). This is because the lines belong to the excited triplet system of helium. Populating this system requires photons or particles of relatively high energy, hence the triplet states are excited in a narrow region between the source of high energy photons and/or particles (the corona) and the denser layers of the chromosphere. While the formation of the *intensity* I of the multiplet is not fully known, for magnetic field measurements using the Stokes IQUV vectors, it need only be recognized that the line picks out certain magnetic field lines in a hydromagnetic structure at the coronal base. Of course this layer is not plane-parallel or simply structured, but observations at the solar limb show that the line is formed over a region of thickness of order 1Mm. Such 1Mm fluctuations in the geometric heights of the formation of this line will translate to errors in the free energy estimated using a planar geometry. However, given that we are interested in the magnetic structure on scales of active regions, i.e. of > 20 Mm, it can be hoped that on such large scales, some important components of the free energy will be reliably measured. General purpose "inversion" programs (from IQUV measurements to vector magnetic fields) have been developed including scattering polarization and the Zeeman and Hanle effects (e.g., Asensio Ramos & Trujillo Bueno 2009). A key assumption being that the maximum optical depth in the multiplet does not significantly exceed unity in the chromospheric formation region. This assumption is certainly valid for essentially all cases of interest, based upon decades of observations. The interpretation of IQUV data of the 1030nm multiplet in terms of vector magnetic field parameters is, practically speaking, now about as simple as any line in the solar spectrum.

A potentially serious issue remains the relative weakness of the polarization (in plages for example, $I >> V >> QU$). Only with stable, mature instruments has the sensitivity been achieved to derive full Stokes vectors of the He I line (Solanki et al. 2003). When QU measurements are noisy, poor information on the full vector field is recovered. In this situation, which will generally be encountered, the data must be augmented in some way. Judge et al. (2009) have proposed to augment polarimetry

with measurements of the fibril morphologies and velocity vectors of plasma inhomogeneities, which have been seen clearly in high resolution (< 1", $\lambda/\Delta\lambda > 200{,}000$) FPI images in other chromospheric lines (e.g. Ca II 854.2 nm). Time series of chromospheric images allow one to track the plane-of-sky components of velocity vectors, Doppler shifts permit measurement of the line-of-sight component. In the low-β, highly conducting environment, the velocities are parallel to the magnetic field. Then, combining the Stokes V and I profiles yields $|\mathbf{B}| \cos\theta$, the derived plasma kinematics yields both θ and φ, which are the field inclination to the line of sight, and azimuth in the plane of the sky, respectively. This analysis can be checked for consistency with the morphology of the fibrils. Another issue is that the He I 1083 nm line forms strongly only in active regions, and there are sometimes active regions which show little He I. In such cases the data can be augmented with Ca~II infrared data (e.g., 854.2 nm), routinely observed using current instruments, even though the diagnosis of magnetic fields from these lines is considerably more complex (Socas-Navarro 2005).

The requirement of high spectral and spatial resolution images, at fast cadence (a 30s cadence suffices, Judge et al. 2009) indicates that, with current technology, FPIs present a promising solution. At present no infrared FPI system has been implemented for solar applications, but there is no practical problem in the development of such FPIs working near 1 micron.

I suggest that the US community consider development of new infrared FPI instruments to enable the science studies outlined here, for implementation at the best solar observatories in the US. While success is not guaranteed, the potential pay-offs might be considered by some to outweigh the risks. Such risks appear to be small. A detailed study by the Arcetri group, which developed the IBIS FPI instrument (Cavallini 2006) operating below 860nm, a work-horse of the National Solar Observatory at the Dunn Solar Telescope, found that such an instrument could be delivered for operation at the Dunn Telescope for $800K with existing technology (unpublished, unfunded proposal to the NSF, 2009). The technological risks are minimal. A bigger risk is that some of the needed magnetic signatures lie close to or below the sensitivity limits achieved to date (some 3 parts in 10,000, Solanki et al 2003)- these are not easy measurements to make. It may also be difficult to apply the Chandrasekhar virial theorem in practice, since it is valid when the magnetic field is force-free at all boundaries enclosing a volume. Thus assumptions concerning the magnetic field outside of the active regions observed will have to be made. However, perhaps the biggest risk is that the community misses the opportunity to build such an instrument when NASA has invested so much on the space missions, past and present, at a time when the ATST will come on line in 2017.

This endeavor is very much at the research forefront and will require the attention from individual researchers to make the measurements and to obtain the first credible measurements of the free magnetic energy (Judge et al., paper presented at the 2010 SHINE meeting). Given the known risks, it is hoped that this approach will mature to the point where it will be possible to routinely estimate the total magnetic free energy of the corona over active regions. This technique will also illuminate the problems of NLFFF extrapolations from photospheric data, which may allow us to make some ad hoc corrections to the calculations based on photospheric measurements, greatly enhancing the value of synoptic datasets such as that from the HMI instrument on SDO.

If successful, such work will not only help to place coronal dynamics, flare physics and space weather origins on a firm observational footing, but also enhance the returns of huge investments by NASA and other agencies in observing the effects of the release of free energy in the Sun, and the subsequent effects on the Earth.

References

Arber, T. D.; Haynes, M.; Leake, J. E., 2007, **ApJ** 666, 541

Asensio Ramos, A.; Trujillo Bueno, J. 2009, In Solar Polarization 5: In Honor of Jan Stenflo, ASP Conference Series, Vol. 405, p.281

Bello González, N.; Kneer, F., 2008, **A&A** 480, 265

Cavallini, F., 2006, **Solar Physics** 236, 415

Judge, P. G., 2010, To appear in the proceedings of the 25th NSO Workshop "Chromospheric Structure and Dynamics. From Old Wisdom to New Insights", Memorie della Societa' Astronomica Italiana, Eds. A. Tritschler et al

Judge, P. G.; Tritschler, A.; Uitenbroek, H.; et al., 2009, **ApJ** 710, 1486

De Rosa, M. L.; Schrijver, C. J.; Barnes, G.; et al., 2009 **ApJ** 696, 1780

Scharmer, G. B.; Narayan, G.; Hillberg, T.; et al., 2008, **ApJ** 689, L69

Schrijver, C. J.; DeRosa, M. L.; Metcalf, T.; et al., 2008, **ApJ** 675, 1637

Solanki, S. K.; Lagg, A.; Woch, J.; Krupp, N.; Collados, M., 2003, Nature 425, 692

Socas-Navarro, H., 2005, **ApJ** 633, L57

Tritschler, A.; Schmidt, W.; Langhans, K.; Kentischer, T., 2002, **Solar Physics** 211, 17

Wiegelmann, T.; Thalmann, J. K.; Schrijver, C. J.; et al., 2008, **Solar Physics** 247, 249

Heliospheric Particle Explorer: Advancing our Understanding of Magnetospheric, Solar Energetic Particle, and Cosmic Ray physics

S. G. Kanekal[1], J. F. Fennell[2], B. Klecker[3], S. Elkington[4],Xinlin Li[4],E. Christian[1], G. de Nolfo[1]
[1] NASA Goddard Space Flight center, [2] The Aerospace Corporation,[3]Max PLanck Institüt für Extraterrestriche Physik,[4]LASP,University of Colorado

1 Introduction

We propose a low earth orbiting satellite with an instrument payload that measure energetic particles over a wide range of energies and species. Charged particles (electrons, protons and heavy ions) in space provide a window into the physical universe encompassing a wide variety of phenomena covering low energy plasmas to highest energy cosmic rays. The proposed mission will address important science science questions in the domains of terrestrial radiation belt dynamics, solar energetic particles, Jovian electrons and cosmic rays. The proposed mission is based on heritage deriving from the very successful NASA mission, SAMPEX (Solar Anomalous Magnetospheric Particle EXplorer) *but extends in a very substantial way the instrument capabilities and mission parameters* thereby enabling significant advances in our scientific understanding of energetic particle dynamics in the heliosphere. The proposed mission spans multiple sub disciplinary areas of research addressing the Sun-Earth system and is a low cost platform worthy of serious consideration for the next decade.

2 Mission Concept

We propose a low Earth orbiting spacecraft carrying instruments capable of measuring charged particles, viz., electrons, protons and ions to advance the scientific knowledge of

i relativistic electron dynamics in the Earth's radiation belts,

ii solar energetic particle sources, acceleration, and propagation, and

iii Jovian electrons and cosmic ray modulation.

The proposed science payload will comprise state of the art instrumentation specifically designed to address the key science topics and will measure particle fluxes, species, and energy spectra with high temporal and spatial resolution. These measurements will significantly further the understanding of relativistic electron dynamics and charged particle acceleration and propagation processes at the Sun and in interplanetary space. Onboard instruments will measure electrons with an energy resolution of $\frac{\Delta E}{E} \approx 20\%$ covering an energy range extending from $\approx$ hundreds of keV to tens of MeV. Additionally high time resolution modes will fully characterize electron microbursts, an important loss mechanism. Measurements of SEP ions will span energies ranging from $\approx$ 1 MeV/nuc to about 100 MeV/nuc and with $1 \leq \mathrm{Z} \leq 28$. The data and science analyses from this mission will complement upcoming in-situ observations of energetic ions in the inner heliosphere.

2.0.1 Mission Overview

The mission we propose will be in a high inclination low earth orbit, e.g., with an altitude of about $\approx$ 600 kms and $\gtrapprox 85^{\circ}$ latitude which will accomplish two important tasks; that of covering all L shells and sampling the Earth's polar regions which have open field lines connecting to the interplanetary space. At these altitudes the orbital period is expected to be about 90 minutes, visiting each L shell about four times in an orbit. The spacecraft orbit will be the sun synchronous dawn-dusk plane. The science payload (described in Section 4) will comprise charged particle sensors with appropriate geometry factors, energy and temporal resolution required to achieve the science goals. The chosen orbit will provide enhanced capabilities for the study of radiation belt dynamics,for example, by covering the magnetospheric regions dominated by EMIC

waves (dusk) and Chorus (dawn) waves. These waves are currently thought to be implicated in loss and energization of relativistic electrons. Since in this orbit, the spacecraft will spend sufficient time over the polar regions thus enabling studies of SEP, Jovian electrons and cosmic rays.

3 Science Gains

The proposed mission will lead to significant advances in our understanding of several key areas of the Sun-Earth system dynamics. The science aspects addressed include, relativistic electron dynamics in the Earth's outer zone, solar energetic particle sources and acceleration processes, and charged particle propagation in the heliosphere. In addition important space weather aspects such as spacecraft anomalies due to deep dielectric discharge, potential hazards to humans in space due to lowered cosmic ray and SEP cutoffs are additional capabilities of the the proposed mission. These topics and the measurements that enable their achievement are summarized briefly below :

- **Radiation belt dynamics**
  - relativistic electron acceleration processes are addressed by the capability of measuring spectral evolution with a high energy and time resolution
  - relativistic electron loss mechanisms due to wave-particle interactions can be characterized by our measurements of microbursts and pitch angle resolution.
  - high time resolution continuous monitoring of the relativistic electron fluxes covering the entire outer zone to asses deep dielectric discharge.
- **Solar Energetic Particles**
  - charge states by utilizing the Earth's magnetic field as a spectrometer.
  - neutral atoms.
  - cutoffs during geomagnetically disturbed times
- **Jovian electrons, and Cosmic Rays**
  - particle transport in the inner heliosphere
  - solar modulation of cosmic rays

3.1 Radiation belt dynamics

The proposed mission in a low Earth polar orbit has a major advantage of being able to sweep across all L shells of the outer zone in a relatively short time. Earlier missions, chiefly SAMPEX have used this advantage and have provided valuable insights. However we note that the instruments onboard SAMPEX were not designed to study radiation belts and had severe limitations such as lack of detailed spectral information, and pitch angle distribution. Furthermore, the PET (Proton Electron Telescope) detector used extensively for radiation belt studies suffered from major contamination due to chance coincidence and count rate saturation. The instrument proposed here to address radiation belt dynamics is designed to provide high quality measurements specifically addressing radiation belt dynamics, for example, the instrument will

- measure electron spectra in multiple differential channels (energy resolution, $\frac{\Delta E}{E} \approx 20\%$)
- be optimized to minimize background, chance coincidence and misidentification
- enable the separation of trapped and precipitating electron fluxes

As described below (see Section 4) the solid state telescopes on board the proposed mission will have a high time resolution mode in order to fully characterize relativistic electron microbursts.

3.1.1 Characterizing electron energization

The proposed mission will measure electron spectra and their temporal evolution and will therefore be able to characterize electron energization. Details such as spatial extent, degree of spectral hardening can

be measured reliably and accurately to help delineate the underlying physical mechanisms. As shown by *Kanekal et al.* [2001]; *Kanekal et al.* [2005] there is a degree of global coherence that implies low altitude measurements capture the dynamics of the relativistic electrons with high fidelity.

Figure 1 shows electron spectra during a geomagnetic storm from days prior to the day of minimum D_{st} (doy 239; -200 nT) to those during the recovery phase. This storm resulted in a strong relativistic electron flux enhancement event. The spectra were measured by the HIST sensor onboard the Polar mission. We note that the measurements by Polar had a low temporal resolution due to its orbit. The proposed mission will have a higher time resolution ($\approx$ 90 minutes) which is needed to discriminate between various electron energization mechanisms. The fast sweep of L shells combined with spectral measurements has the potential to investigate important aspects of energization, such as the temporal and spatial relationship of a seed population to that of the energized fluxes, response of the magnetosphere as a whole to solar wind energy input (*Vassiliadis* [2008]), and characterization of the dependence upon diverse interplanetary drivers. In addition electron losses (see next section) will also be characterized and thus enable a fuller and complete understanding of electron dynamics which is of course a result of the balance between energization and loss.

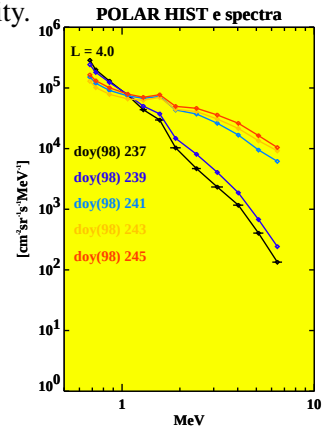


Figure 1: Evolution of electron spectra.

3.1.2 Characterizing electron losses

Relativistic electrons are lost by precipitation, and movement out of the magnetosphere, e.g., magnetopause shadowing. Wave particle interactions are thought to be the leading cause of electron precipitation (reviewed recently by *Millan and Thorne* [2007]). The proposed mission can advance our knowledge of both slow and steady loss of the type resulting in global coherence (*Kanekal et al.* [2005]) as well as rapid loss through electron microbursts.

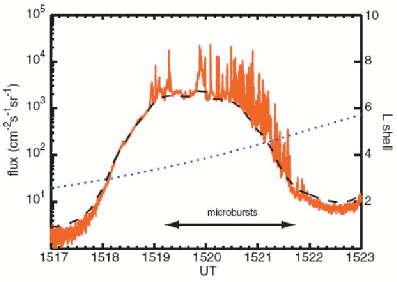


Figure 2: electrons $>$ 1 MeV microbursts on October 19, 1998 (adapted from *Lorentzen et al.* [2001]

Figure 2 shows count rates of relativistic electrons during a radiation belt pass as seen by the HILT sensor onboard SAMPEX. The microbursts are seen as rapid fluctuations in electron count rates in a region from L$\approx$ 3 $-$ 6 (*Lorentzen et al.* [2001]). These microbursts are thought to constitute a significant mode for the loss of relativistic electrons and have been estimated to be capable of draining significant fluxes of electrons (*O'Brien et al.* [2003]). A full and detailed characterization of the microbursts would include spectral,spatial and temporal aspects of these events. All these have important bearings upon the physics of electron loss, for example, spatial characterization of microbursts reveals the spatial extent of the wave-particle interactions causing electron precipitation. The proposed mission with the instrument suite described is ideally suited for characterizing microbursts.

3.2 Solar Energetic Particles

Solar energetic particles provide important information regarding the dynamic nature of their source, i.e, the Sun and the interplanetary medium through which they travel. Their study is important not only from a scientific perspective but also from a space weather point of view. The mission we propose is ideally suited to study SEP acceleration, transport and their impact on humans in space. The solid state particle telescopes onboard the proposed mission (see Section 4) measure electrons, protons and ions over a wide range of energies with excellent energy resolution.

3.2.1 Charge States of Solar Energetic Particles

Measurements obtained over the last decade have shown a large variability of the mean ionic charge of heavy ions in solar energetic particle (SEP) events (e.g., *Klecker et al.* [2007]). This variability is due both to source populations and to variable conditions during acceleration and propagation. For example, the mean ionic charge in impulsive (electron-rich, 3He-rich and Fe-rich) events is strongly energy dependent and increases for Fe from 11-15 at 0.1 MeV/nucleon to 16-20 at 0.5 MeV/nucleon *DiFabio et al.* [2008]. This can only be explained by charge stripping in a sufficiently dense environment during acceleration, placing the acceleration region in the low corona (e.g. *Kartavykh et al.* [2007]). In large (gradual) SEP events where the intensity increase of energetic particles in interplanetary space is related to coronal and interplanetary shocks, the ionic charge at low energies ($<$ 1 MeV) is compatible with solar wind charge states (e.g. Q $\sim$10 for Fe, *Klecker et al.* [2008]), suggesting a solar wind origin. At energies above 10 MeV/nuc, also high charge states (Q $\sim$20 for Fe) are often observed (e.g. *Labrador et al.* [2003]), indicating the contribution of different particle sources to the energetic ions observed in interplanetary space. These results illustrate that the determination of the ionic charge of SEPs is essential for the understanding of composition and energy spectra of SEPs that generally will depend on the particle source, and on acceleration and propagation conditions, that in turn depend on the ionic charge of the particles.

We note that there are no instruments for the in-situ determination of the ionic charge of energetic ions included in the upcoming missions such as the Solar Orbiter and Solar Probe Plus that cover the time period of the decadal Survey 2013-2022. The measurement of the ion parameters energy and mass, together with the determination of the geomagnetic cutoff in a low altitude polar orbit, allows the determination of the ionic charge of solar energetic particles. This technique was successfully used on the SAMPEX spacecraft (*Mason et al.* [1995]; *Leske et al.* [1995]) and allowed for the first time the direct determination of the ionic charge of anomalous cosmic rays (*Klecker et al.* [1995]), and the determination of SEP ionic charge states over the extended energy range of 0.3 to 70 MeV/nuc (*Oetliker et al.* [1997]).

3.2.2 Neutral atoms

Recent observations by *Mewaldt et al.* [2010] have characterized neutral hydrogen related to a flare/CME event and created by proton acceleration at interplanetary shocks. The authors summarized that these neutrals have the potential to advance scientific understanding of solar particle acceleration. From the vantage point of LEO where the geomagnetic field acts as a shield to charged particles, so that the observations of neutral hydrogen are relatively free of background from protons. A dawn-dusk orbit as proposed would have $\gtrapprox$ 50% duty cycle (due to Earth limb shadowing) and over the life cycle of the mission may be expected to gather a fair number of events. In conjunction with other missions imaging the Sun, these *ENA images* may be compared to CME images to obtain insights into the nature of solar neutrals.

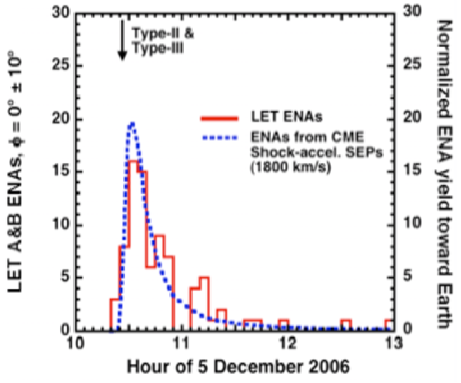


Figure 3: Neutral hydrogen from the Sun (from *Mewaldt et al.* [2010]).

3.2.3 Cutoff variability and trapping

Charged particle cutoffs vary considerably during geomagnetically disturbed times constituting a serious hazard to humans and human assets in space. Figure 4 illustrates the observed proton cutoff variability during a geomagnetic storm (November 1997). Lowered cutoffs expose humans and human assets to radiation

hazards and monitoring them is an important aspect of space weather. Furthermore, the correlation between the cutoff location and the D_{st} index at times shows cutoffs being lowered prior to D_{st} decrease *Kanekal et al.* [2005]; *Leske et al.* [2001].

Thus apart from important space weather aspects there are additional scientific issues that need to be addressed. At times, SEPs penetrate deep into the magnetosphere and get trapped resulting in the formation of new belts which can last from days to months.Details of the physics of the trapping and the role of the interplanetary drivers remain yet to be fully established. The proposed mission in high inclination LEO is ideally suited to carry out these measurements. The onboard charged particle instruments with optimized geometry factors, background suppression and temporal,spatial and energy resolution will significantly further our scientific understanding.

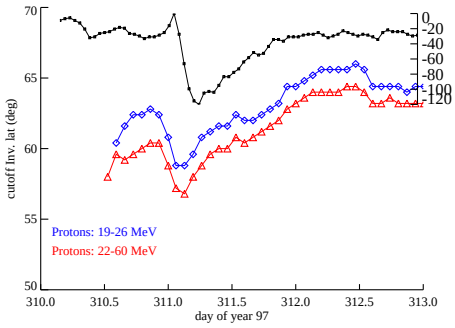


Figure 4: Proton cutoffs and D_{st}

3.2.4 Jovian electrons and cosmic rays

Jovian electrons can be observed over the Earth's poles and can provide information not only regarding the source *Kanekal et al.* [2003] but also about particle transport in the heliosphere, validation of solar wind models, interplanetary field twisting and turbulence *Owens et al.* [2010]. The proposed mission will also monitor cosmic rays during the next solar cycle. Considering the record high GCR intensity during the current unusually extended solar minimum, it will be of great interest to observe the solar modulation during the upcoming maximum. We note that the Voyager mission will measure the interstellar spectrum of ions which can be compared with observations made by the proposed mission to further the understanding of solar modulation. These science topics are well within the scope of the instrument proposed for this mission and increase its multi-disciplinary appeal.

4 Scientific Payload

The science payload comprises of charged particle sensors capable of measuring electron, protons and ions covering the appropriate energy ranges required to address the science goals. The details of the instrument energy ranges and resolution are spelled out in Table 1. The SST1 and SST2 sensors will have the additional capability on select energy channels to measure electron rates with a very high time resolution of the order of tens of milliseconds in order to characterize electron microbursts which result in loss of relativistic electrons. The higher energy sensor, SST1 will be modeled after the REPT (relativistic Electron and Proton Telescope) which will be flown on NASA's RBSP (Radiation Belt Storm Probes) spacecraft. However REPT does not have the high resolution capability and the RBSP orbit is unsuited for the study of microbursts. The payload will comprise two copies of each of the instruments oriented at right angles to each other. This will enable the SSTs to measure the trapped and the precipitating particle populations. One of the ICA units will be sun pointing and will be able to characterize solar neutral atoms (which of course do not follow the spiral interplanetary field). While charged particle telescopes are not a new technology, there is scope for innovation regarding

Instrument Type	Measurement Characteristics	Science Goal
Solid state Particle Telescope (SST1)	$\mathbf{e}^- \gtrsim 2$ to $\gtrsim 10 MeV$ $\Delta E/E \approx 30\%$ $\mathbf{p} \gtrsim 15$ to $\gtrsim 70 MeV$ $\Delta E/E \approx 30\%$	rel. electron energization and loss, inner belt and solar protons
Solid state Particle Telescope (SST2)	$\mathbf{e}^- \gtrsim 0.1$ to $\gtrsim 4 MeV$ $\Delta E/E \approx 30\%$ $\mathbf{p} \gtrsim 5$ to $\gtrsim 20 MeV$ $\Delta E/E \approx 30\%$	rel. electron energization and loss inner belt and solar protons
Ion Composition Analyzer (ICA)	$1 \leq Z \leq 28$ ≈ 1 to ≈ 100 MeV/nuc	charge states energetic neutral atoms

Table 1: Proposed instrument suite.

the associated electronics and instrument configuration that makes them desirable on future missions. The proposed instrument suite is therefore appropriate for future missions considered by NASA and reviewed by this decadal survey.

5 Estimation of Mission cost

Table 2 lists estimated costs for the proposed mission. The instrument costs include design, development, building and calibration. We note that there are two copies of each instrument, one zenith pointing and the other looking toward the sun and perpendicular to the zenith. Launch costs depend upon whether a dedicated launch vehicle is used or piggybacked with another mission. The total mass of the payload and the spacecraft bus is estimated to be less than 200 kilos which can be launched on a pegasus. The instrument costs are based on our experience with RBSP/REPT, SAMPEX/MAST, ACE/SIS, STERO/LET and other instruments. Operational costs are extrapolated from our prior experience with SAMPEX and include software development for mission ops. as well as science data analyses. The estimated total mission costs are $\lessapprox$ \$150M which makes this a small sized mission.

Instrument Launch,Ops.	Total cost \$M	Additional Remarks
SST1	20	two units
SST2	20	two units
ICA	20	two units
s/c bus	50	dedicated
Launch	30	dedicated
	15	piggyback
Operational	5	per year

Table 2: Estimated mission costs.

6 Viability of the Proposed Mission

The proposed mission makes significant contributions to both the SWM and SHP panels addressing as it does, radiation belt response to solar wind and the characterization of SEPs. The study of radiation belt energization and loss, and understanding the acceleration and transport of charged particles in the interplanetary space are important scientific questions for the next decade and are ripe for significant advances. We note that the space weather component of the mission has important societal benefits as human exploration and human assets in space are expected to dramatically increase in the next decade. The proposed mission may be defined as *small* since its total costs are expected to be under \$150M and as such is eminently affordable. Moreover, the proposed instrument suite uses tried and true technologies while being flexible enough to accommodate new advances in solid state detector and associated electronics technologies. The mission can be ready in a reasonably short time span as it is technically straightforward and well within the expertise of major space science institutions in the country. Finally the proposed mission will very strongly complement both radiation belt missions such as the RBSP and interplanetary missions such STEREO.

References

DiFabio, R. Z., E. Guo, E. Mbius, B. Klecker, H. Kucharek, G. M. Mason, and M. A. Popecki, Energy dependent ionic charge states and their connection with ion abundances in impulsive solar energetic particle events, *Astrophys. J.*, *687*, 2008.

Kanekal, S. G., D. N. Baker, and J. B. Blake, Multisatellite measurements of relativistic electron : Global coherence, *J. Geophys. Res.*, *106*(29), 29,721, 2001.

Kanekal, S. G., D. N. Baker, J. B. Blake, M. D. Looper, R. A. Mewaldt, and C. A. Lopate, Modulation of Jovian electrons at 1 AU during solar cycles 22-23, *Geophys. Res. Lett.*, *30*(15), 150,000–1, doi: 10.1029/2003GL017502, 2003.

Kanekal, S. G., R. H. W. Friedel, G. D. Reeves, D. N. Baker, and J. B. Blake, Relativistic electron events in 2002: Studies of pitch angle isotropization, *Journal of Geophysical Research (Space Physics)*, *110*(A9), 12,224–+, doi:10.1029/2004JA010974, 2005.

Kanekal, S. G., M. D. Looper, D. N. Baker, and J. B. Blake, Study of proton cutoffs during geomagnetically disturbed times, *Eos Trans. AGU*, pp. SA2A–0284, 2005.

Kartavykh, Y. Y., W. Dröge, B. Klecker, G. M. Mason, E. Möbius, M. Popecki, and S. Krucker, Evidence of a Two-Temperature Source Region in the ^{3}He-Rich Solar Energetic Particle Event of 2000 May 1, *Astrophys. J.*, *671*, 947–954, doi:10.1086/522687, 2007.

Klecker, B., E. Möbius, and M. A. Popecki, Ionic Charge States of Solar Energetic Particles: A Clue to the Source, *Space Sci. Rev.*, *130*, 273–282, doi:10.1007/s11214-007-9207-1, 2007.

Klecker, B., E. Möbius, M. A. Popecki, and et al., Ionic charge states of heavy ions associated with interplanetary shocks at 1AU, in *International Cosmic Ray Conference*, *International Cosmic Ray Conference*, vol. 1, pp. 83–86, 2008.

Klecker, B., et al., Charge state of anomalous cosmic-ray nitrogen, oxygen, and neon: SAMPEX observations, *Astrophys. J.*, *442*, L69–L72, doi:10.1086/187818, 1995.

Labrador, A. W., R. A. Leske, R. A. Mewaldt, E. C. Stone, and T. T. von Rosenvinge, High Energy Ionic Charge State Composition in Recent Large Solar Energetic Particle Events, in *International Cosmic Ray Conference*, *International Cosmic Ray Conference*, vol. 6, pp. 3269–+, 2003.

Leske, R. A., J. R. Cummings, R. A. Mewaldt, E. C. Stone, and T. T. von Rosenvinge, Measurements of the Ionic Charge States of Solar Energetic Particles Using the Geomagnetic Field, *Astrophys. J.*, *452*, L149+, doi:10.1086/309718, 1995.

Leske, R. A., R. A. Mewaldt, E. C. Stone, and T. T. von Rosenvinge, Observations of geomagnetic cutoff variations during solar energetic particle events and implications for the radiation environment at the Space Station, *J. Geophys. Res.*, *106*, 30,011–30,022, doi:10.1029/2000JA000212, 2001.

Lorentzen, K. R., L. M. D., and B. J. B., Relativistic electron microbursts during the gem storms, *Geophys. Res. Lett.*, *28*, 2573, 2001.

Mason, G. M., J. E. Mazur, M. D. Looper, and R. A. Mewaldt, Charge State Measurements of Solar Energetic Particles Observed with SAMPEX, *Astrophys. J.*, *452*, 901–+, doi:10.1086/176358, 1995.

Mewaldt, R. A., et al., Observations and Interpretations of Energetic Neutral Hydrogen Atoms from the December 5, 2006 Solar Event, *Twelfth International Solar Wind Conference*, *1216*, 592–595, doi:10.1063/1.3395935, 2010.

Millan, R. M., and R. M. Thorne, Review of radiation belt relativistic electron losses, *J. Atmos. Solar Terr. Phys.*, *69*, 362–377, 2007.

O'Brien, T. P., K. R. Lorentzen, I. R. Mann, N. P. Meredith, J. B. Blake, J. F. Fennel, M. D. Looper, D. K. Milling, and R. R. Anderson, Energization of relativistic electrons in the presence of ULF power and MeV microbursts: E vidence for dual ULF and VLF acceleration, *J. Geophys. Res.*, *106*(A8), 15,533, 2003.

Oetliker, M., B. Klecker, D. Hovestadt, G. M. Mason, J. E. Mazur, R. A. Leske, R. A. Mewaldt, J. B. Blake, and M. D. Looper, The Ionic Charge of Solar Energetic Particles with Energies of 0.3–70 MeV per Nucleon, *Astrophys. J.*, *477*, 495–+, doi:10.1086/303679, 1997.

Owens, M. J., T. S. Horbury, and C. N. Arge, Probing the Large-scale Topology of the Heliospheric Magnetic Field using Jovian Electrons, *Astrophys. J.*, *714*, 1617–1623, doi:10.1088/0004-637X/714/2/1617, 2010.

Vassiliadis, D., Response of the radiation belt electron flux to the solar wind velocity: Parameterization by radial distance and energy, *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*, 1810–1828, doi:10.1016/j.jastp.2008.05.019, 2008.

THE SUN AS A STAR

Vinay Kashyap[1], *Jay Bookbinder*[1], *Ofer Cohen*[1], *Jeremy Drake*[1], *Fabio Reale*[2], *Kathy Reeves*[1], *Steve Saar*[1], *Randall Smith*[1], *Paola Testa*[1], *Brian Wood*[3]

We argue for an increased focus on studies that target solar-stellar connections. Theoretical and modeling work developed for the Sun has been explicitly beneficial to stellar studies. The wide range of parameters that are accessible in stars, and the different types of data that are available also help to place constraints on the theoretical foundations of solar physics.

1 Introduction

Advances made in solar astrophysics have been regularly adapted to understand other stars and explain their characteristics. This is a natural consequence of the fact that solar properties can be measured in detail and it thus drives theoretical improvements. A case in point is the "Solar Analogy" that is applied to stellar coronae: it is generally accepted that the atmospheres of stars with convective zones are similar to that of the Sun, in that the atmospheric structure is dominated by magnetic fields that are generated from a self-sustaining dynamo and which plays a significant role in heating the atmospheres to high ($>$ 1MK) temperatures (see reviews by Rosner, Golub, & Vaiana 1984, Güdel et al. 2004; see also Pallavicini et al. 1981, Ayres et al. 1981, Peres et al. 2004). But when examined in detail, the solar analogy is seen to be somewhat simplistic, with stars exhibiting a number of properties in detail that are either at the extremes of the solar behavior, or point to the existence of new physical mechanisms (see below). Such comparative studies of stellar processes increases our capability to better understand the physical processes on the Sun. Stars allow us to explore parameter spaces that cannot be reached on the Sun, thereby providing better constraints to both theoretical and numerical studies. In many instances, the quality of the data available for stars far exceeds that available on the Sun (e.g., energy- and time-tagged photon lists, high-resolution X-ray gratings spectra over a broad energy range), allowing us to apply powerful analysis tools and techniques. The larger sample size available with stars is also useful in establishing the statistical distributions of various phenomena; this is crucial for theory to avoid the trap of dealing with and explaining a single case.

In the following, we describe some of the work that has been done in this context. Our focus here is on outer atmospheres and dynamo studies, and we explore the different ways theoretical, observational, and numerical studies take ideas developed in one field and apply it to the other. We discuss flares and flare modeling in §2, abundance anomalies in §3, the monitoring of activity cycles in §4, and numerical simulations in §5.

There is a need for coordination between solar and stellar researchers to build upon this work. It is necessary to provide funding for research at the interface of solar and stellar astrophysics, and to fund missions that improve the timing and spectral resolutions in solar observations. Current solar missions are not optimized for the type of solar-stellar studies we describe here, but some proposed missions, such as the Solar Spectroscopy Explorer (SSE), which uses a high spectral resolution calorimeter ($\sim$2-5 eV in the soft X-ray) to do photon counting with a precision of $\sim$10 μs will go a long way to bridge this gap.

[1]Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA (vkashyap@cfa.harvard.edu)
[2]Osservatorio Astronomico di Palermo, Palermo, Italy
[3]Naval Research Laboratory, Space Science Division, Washington, DC 20375, USA

2 Flares

Dynamic processes like flare events are useful for testing theories: every event is unique and displays important diagnostic differences from previous events. They are also the most easily studied events due to their higher intensities and relative rarity. Significant progress has been made in understanding the onset and evolution of solar flares (Kopp & Poletto 1984, Tsuneta 1995, Reeves et al. 2007). With the advent of high-cadence coronal imaging (with, e.g., *Hinode*/XRT, *SDO*/AIA; and spectra at high energies with *RHESSI*) with a number of filters that sample a wide range of temperatures, changes to the magnetic topology, timescales of reconnection, the energetics of thermal and non-thermal processes, can all be studied in great detail and within the context of the active region environment in which the flares occur.

While such detailed study is not possible for stellar flares, the larger morphological variety, time-tagged events analysis, better spectral information, and the wider range of intensities displayed by them provide more diagnostics to test against theoretical models. Modeling of stellar flares has thus far relied on assumptions of single loop processes, with a significant amount of analyses limited to scaling law arguments or cooling timescales (see Table 4 of Güdel 2004). More sophisticated hydrodynamical loop models have also been applied to data directly (Reale et al. 1988) or through the derivation of empirical diagnostic tools (Reale et al. 1997, Reale & Micela 1998, Reale 2007), but they remain limited to single loops or arcades of few loops (Reale et al. 2004). Spatial arguments based on flares being eclipsed by a dark companion (e.g., Algol, Schmitt & Favata 1999, Favata & Schmitt 1999) or being modulated by the rotation (Stelzer et al. 1999) have suggested that stellar flares are generally limited to smaller scales and occur relatively low in the corona. In some stars, so-called superflares have been detected, where the the instantaneous X-ray luminosities approach or exceed average bolometric luminosities (Franciosini et al. 2001, Osten et al. 2007, Kashyap et al. 2006, Osten et al. 2010). These flares are characterized by high temperatures, high densities, and have geometries comparable to that expected from solar coronal loops (loop lengths $\sim 0.1 - 0.5 R_*$, aspect ratios ~ 0.1). They also show distinct evidence for non-thermal high-energy emission as in the case of solar X-class flares.

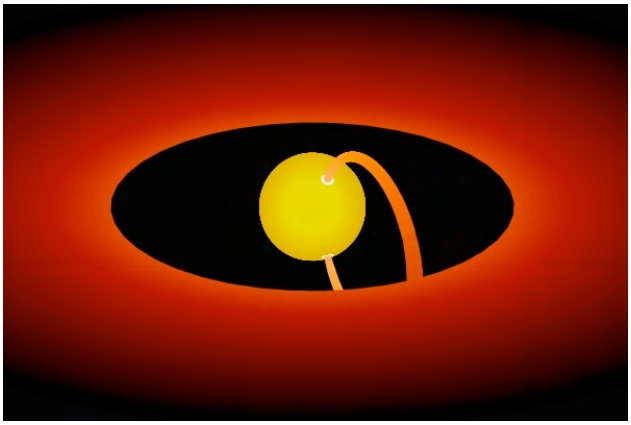


Figure 1: Intense and long-lasting stellar X-ray flares may occur in long loops connecting the star to the circumstellar disk.

A significant correlation between the soft and hard peak X-ray fluxes (measured with *GOES* and *RHESSI* respectively) has been found in a sample of solar flares of all intensities. The resulting scaling law has been compared and found to be compatible with the available stellar flare observations (Isola et al. 2007). An intriguing question here is the existence of very long flaring loops inferred from modeling the intense and long-duration flares observed in young stellar systems (Favata et al. 2005). The question is interesting because such structures might be so long as to connect the stellar surface to the circumstellar disk still accreting on the star (Figure 1). It would then be an indirect evidence of the existence of such long magnetic channels which has been already proposed to describe the disk-star accretion (Calvet & Gullbring 1998). Although these long loops are supported by solid and detailed arguments, they are on a completely different scale from solar coronal structures, raising questions about their lifetime and stability and about the structure and topology of the global star-disk magnetic field. Therefore, they represent a challenge for future studies.

Analysis of stellar flares thus far has been limited to determining the geometric and energetic constraints of the flaring volume. The next stage in the analysis is clearly to apply more sophisticated models that have been developed for the Sun (e.g., Reeves et al. 2007, who model the flare events as a set of cascading loops in an arcade that are sequentially energized). Stellar flare data can be used to set constraints on hitherto unexplored parameters such as footpoint separation, flare height, and magnetic field strength (see Kashyap et al. 2009).

A singular feature of solar flares is the distribution of flare numbers as a function of flare energy. Over many orders of magnitudes, it is seen that the solar flare number distribution follows a power-law (Hudson 1991) with index $\alpha = 1.8$. This is matched by a similar scaling in the magnetic fluxes of observable features (Parnell et al. 2009). In contrast, on many active stars where the flare distribution has been studied, it has invariably been that $\alpha > 2$ (Audard et al. 2000, Kashyap et al. 2002, Güdel 2004). Though there is evidence for an upturn in the solar distribution index for weak flares (see e.g., Parnell & Jupp 2000, Winebarger et al. 2002), this is at energies orders of magnitude lower than the flare energies accessed in stellar observations, and temperature response effects have not been ruled out. The steepness of the distribution is of critical importance, since if $\alpha > 2$, then the entire coronal energy budget could be ascribed to microflares. Studying stellar flares in this context is extremely important, since they can provide exemplar cases where the solar analogy is strongly limited. For instance, there are indications that the flare distribution on the dMe star Ross 154 is better described with a double power-law (Wargelin et al. 2008; Figure 2).

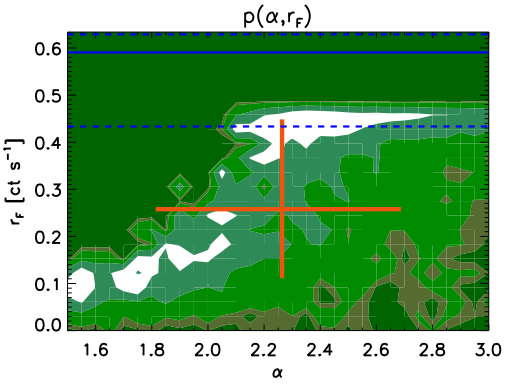


Figure 2: Modeling small flares on dMe star Ross 154 with *Chandra*/HRC-I data. The joint probability distribution of the flare index α and the contribution of flares to the count rate, $r_{\rm F}$, is shown. The cross indicates the average values, and the lighter regions denote higher probability. Two distinct regions, one with $\alpha > 2$ and $r_{\rm F} \sim 0.4$ ct s^{-1}, and another with $\alpha \sim 1.8$ and $r_{\rm F} \sim 0.1$ ct s^{-1}, are visible, suggesting that on this star the flare number distribution is likely not a single power-law (Wargelin et al. 2008).

A complementary investigation considers the contribution of flares to the steady X-ray emission in stellar coronae. Active stellar coronae generally have a different thermal distribution of plasma than on the Sun. The distribution of emission measure with temperature, $EM(T)$ of active stellar coronae usually peaks around $\sim 8 - 20$ MK, and, in some cases, with large amount of emission measure at very high temperatures ($T \sim 20-30$ MK, Güdel et al. 1997; Scelsi et al. 2005). The hot bump in $EM(T)$ in active stars has been observed also in the absence of apparent flaring activity, i.e., during steady emission (e.g., Argiroffi et al. 2003). Conversely, on the Sun and on less active stars, $EM(T)$ peaks at cooler temperatures ($T \sim 1 - 2$ MK, Peres et al. 2000; Raassen et al. 2003), and significant amount of plasma at very high temperatures is observed only during flares (e.g., Reale et al. 2001). From the $EM(T)$ distribution obtained for the solar corona averaged over long time scales and including flares, it has been concluded that the $\sim 8-20MK$ peak in the EM(T) of active stars, could be the result of a large number of unresolved solar-like flares (Argiroffi et al. 2008, Figure 3). If this is the case, quiescent and flare components should follow different scaling laws for increasing stellar activity. In the assumption that the heating of the corona is entirely due to flares, from nano- to macro-flares, then either the flare distribution or the confined

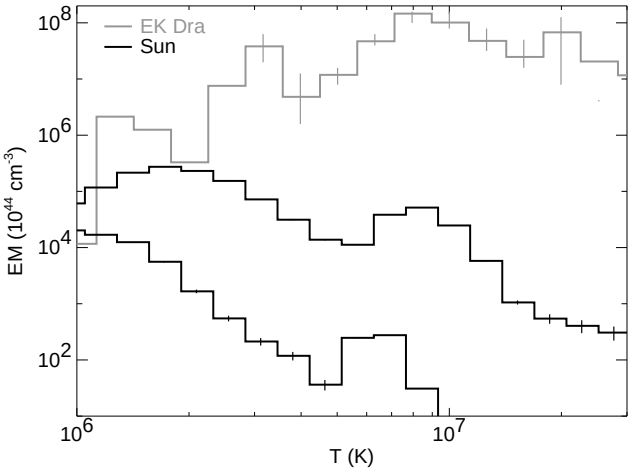


Figure 3: Solar emission measure distributions (flaring + quiescent) near the maximum and minimum phases of the solar cycle (december 1991 and december 1995, both in black) compared with the emission measure distribution of EK Dra (Scelsi et al. 2005, in gray). Hot peaks in the stellar distribution may be due to continuous flaring activity. (From Argiroffi et al. 2008)

plasma response to flares, or both, are bimodal. This is a very important question concerning the basic nature of coronal heating.

There are many questions that can be addressed by a comparative study of solar and stellar flare distributions that could be of fundamental importance to our understanding of coronal heating mechanisms: Why do flare distributions appear to be self-similar over a large dynamic range? Why are stellar flare indices usually > 2? Is there a trend in power-law indices that can be matched to other stellar characteristics? How much of the coronal energy budget on stars can be attributed to microflares, and conversely, why does microflaring not appear to be a dominant contributor to the solar coronal heating mechanism?

3 Abundance Anomalies

Another example of how stellar observations may provide crucial insights into problems in solar physics concerns curious abundance anomalies that have been found in the solar corona. The chemical composition of plasmas in the solar outer atmosphere is not constant nor uniform, and shows evidence for anomalies with respect to the underlying photospheric mixture. In the solar corona and solar wind, abundances are generally found to be dependent on First Ionization Potential (FIP). Relative to the photosphere, elements with low FIP (FIP $<$ 10 eV; Fe, Mg, Si, K, etc.) are generally found to have coronal abundances that are enhanced relative to elements with high FIP ($>$ 10 eV; C, N, O, Ne, etc.) (Meyer 1985, Feldman 1992, Feldman & Laming 2000, Sylwester et al. 2010). This "FIP effect" changes in different types of solar features (e.g., coronal holes, fast solar wind, and newly emerged active typically show abundances close to photospheric), and with time (Widing & Feldman 2001).

Stellar studies allow us to investigate the fractionation processes and reveal trends of abundance patterns in a wide range of stellar parameters and activity levels, not accessible from solar studies alone. For about a decade now, high resolution X-ray spectra from the *Chandra* and *XMM* observatories have provided the high quality X-ray spectra necessary to study the coronal abundances of cool stars, supplementing EUV evidence of FIP effects found for stars with low to moderate activity (Laming et al. 1996; Drake et al. 1997; Laming & Drake 1999). However, for more active stars, the FIP effect is generally either absent or sometimes an inverse FIP effect (IFIP) is observed, where high FIP elements have coronal abundances that are *higher* relative to photospheric values (Audard et al. 2003; Huenemoerder et al. 2001,2003; Sanz-Forcada et al. 2003; Liefke et al. 2008).

These abundance anomalies depend on the still unknown physical mechanisms of chemical fractionation in the process of mass and energy transport to the corona, and therefore have the potential to provide valuable constraints the elusive coronal heating mechanism(s), one of the most fundamental open issues in solar physics and astrophysics in general.

There is evidence for an activity dependence of the FIP effect, which may be an important clue as to the origin of this effect in both stars

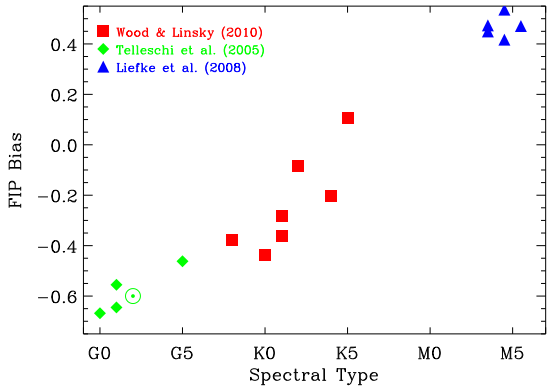


Figure 4: Does the FIP effect depend on spectral type? For stars with moderate activity, the magnitude of the FIP bias appears to be strongly correlated with stellar spectral type (Wood & Linsky 2010).

and the Sun. Studies of solar analogs at different stages of their evolution have suggested that the abundance anomalies of coronal plasma of the "Sun in time" evolve from an IFIP effect in its early active stages (when the corona is hot, luminous, and frequently flaring) to a solar-like FIP effect on short timescales (< 300 Myr; Telleschi et al. 2005). If the extremes of stellar activity are avoided the activity dependence disappears entirely, and one sees a very different empirical picture, one where coronal abundances are correlated with stellar spectral type instead of coronal activity level (Wood & Linsky 2010; see Figure 4; see also Telleschi et al. 2007, Güdel et al. 2007). This dependence however appears to hold only for solar-like stars with low to moderate activity ($L_X < 10^{29}$ erg/s), and might point to different physical mechanisms for chemical fractionation and coronal heating.

The abundance of neon in stellar coronae and in the Sun has been recently subject of a heated debate in the solar and stellar community, also due to the importance of Ne contribution to the opacity for solar interior models (see e.g., Bahcall et al. 2005, Antia & Basu 2005). Stellar coronae with moderate to high activity levels show a Ne/O abundance ratio rather constant and significantly larger than the Sun (Drake & Testa 2005). Low activity coronae are characterized by intermediate Ne/O, closer to the solar values (Liefke et al. 2006, Robrade et al. 2008). The lack of diagnostics of Ne abundance in the photosphere of the Sun and solar-like stars raises the issue of the assumed photospheric abundance. A higher photospheric Ne (consistent with the Ne/O ratio measured in most stellar coronae) might provide enough opacity to help resolve the "solar model problem" (Drake & Testa 2005, Antia & Basu 2005). However, recent studies indicate that it is unlikely that Neon can provide the full solution (Basu & Antia 2008).

Solar full-disk X-ray/EUV spectra are surprisingly scarce (see Laming et al. 1995 and references therein), making it challenging to determine the element abundances for the Sun as a whole. Solar and stellar abundance studies together provide tight constraints on the models. It is plausible that coronal abundances are somehow determined by basic photospheric properties of a star, such as radius, surface gravity, photospheric temperature, convective motions, etc. While most proposed theoretical mechanisms are unable to reproduce the wide range of observed abundance anomalies as a function of stellar parameters (Henoux 1995, Drake 2003), recent efforts, based on the ponderomotive forces associated with the propagation of Alvén waves through the chromosphere (Laming 2004, Laming 2009) provide promising results being able to produce both FIP and IFIP effect. It is thus possible that changes in basic stellar properties could affect Alfvén wave propagation and thereby affect the coronal FIP biases. It is worth noting that such an insight into the nature of the FIP effect would not have been possible without the use of stellar observations to place the Sun into context with other stars. Conversely, the lack of adequate solar data that can be compared with stellar observations hampers progress in modeling coronal abundances.

4 Activity Cycles

The onset of Cycle 24 on the Sun has proven to be surprising from a theoretical standpoint. The long minimum the Sun has undergone has been unanticipated, and none of the predicted sce-

narios based on the previous cycles have played out as expected (see Pesnell 2008). In this context, studying activity cycles on stars can be useful to place constraints on dynamo theories and point to observational methodologies that could be fruitfully applied to the Sun. For example, an effort to track the temperature and component-wise variations in the solar corona over the course of a cycle, in a manner similar to that done for stars in general, was done using *Yohkoh* data by Orlando et al. (2000, 2001; see also Peres et al. 2004, Argiroffi et al. 2008). While this suggested variations of over a magnitude in luminosity at high energies, the behavior at soft X-ray energies is less well understood. Recently, during monitoring of the nearyby solar proxy α Cen A with the X-ray observatory XMM-*Newton*, it was discovered that the star had dropped in intensity by a factor of nearly 50x; but with high-resolution soft X-ray spectra from *Chandra*'s HRC-S/LETG, it was soon found that most of this change could be ascribed to a downward shift in the coronal temperature and not to a change in the luminosity (Ayres 2008). This result points to a need to obtain detailed thermal information of the solar corona as it goes through the cycle at high spectral resolution over a wide range of wavelengths. None of the current solar observatories are capable of acquiring such data, though the recently proposed Solar Spectroscopic Explorer will be able to achieve this with its calorimeter detector.

Analysis of stellar data can also greatly help in constraining how magnetic fields are generated and maintained. Despite decades of work and many advances, solar dynamo theory is still unable to consistently make predictions: witness the wide range of predictions for the shape and strength of the current cycle, and its anomalous minimum, for example (Pesnell 2008). A large part of the problem is that the Sun provides just a single set of physical properties (mass, rotation rate, convection zone depth, etc.) as initial conditions for the dynamo "experiment". Stellar data can provide many more sets of properties with which to test and constrain the uncertain parameters of the problem. Cycle periods (Baliunas et al 1995; Olah et al 2009) and amplitudes (Saar & Brandenburg 2002) and mean flux generation (e.g., via a coronal X-ray proxy; Pevtsov et al. 2003) can be matched using models constrained by measured surface differential rotation rates (Donahue et al. 1996) over a wide range of rotation rates and masses. The evolution of active latitudes over the cycle can be tracked by coupled changes in mean rotation period and mean activity level over time (eg, Donahue & Baliunas 1992), giving constraints on the meridional flow speed (for Babckock-Leighton models), giving further model constraints. Hints of global magnetic topology (mostly toroidal or poloidal, axi- or non-axisymmetric) are emerging (Donati & Landstreet 2009). The stars also offer windows on behavior unknown or only hinted at in the Sun: persistently active longitudes (Korhonen & Elstner 2005); multiple cycle periods (Baliunas et al 1995; Olah et al 2009); multiple active latitude belts (e.g., Donahue & Baliunas 1992); and strong magnetic field production in the absence of significant differential rotataion (Saar 2009). A dynamo theory which could succesfully model both the solar case and *all* these diverse stellar phenomena would truly be well vetted, and the process would help sort out the plethora of competing models.

5 Numerical MHD

Over the past decade numerical models for the solar corona and wind, and the interplanetary environment, have become highly realistic. The information gathered from numerous missions such as Voyager, Ulysses, Pioneer, ACE, WIND, SoHO, etc., has been incorporated into sophisticated MagnetoHydroDynamic (MHD) computer models such as BATS-R-US (see Powell et al, 1999, Cohen et al. 2007,2008) that have been able to successfully match space weather observations, coronal mass ejections, and other aspects of the interplanetary environment. These methods and codes are driven by high resolution magnetograms and provide realistic steady-state solutions for the corona and solar wind to 1 AU and beyond, as well as simulate realistic eruptive events like coronal mass ejections. They are now being applied to stellar cases. In one of the first

such applications, motivated by recent findings of polar active regions on active stars such as AB Dor (Hussain et al. 2007) and Algol (Chung et al. 2004), Cohen et al. (2009) studied the effect of magnetic topology on the mass loss and angular momentum loss due to the stellar wind, and found that the mass- and angular-momentum-losses change in causal mechanism and increases in magnitude as the active regions are moved poleward (see Figure 5).

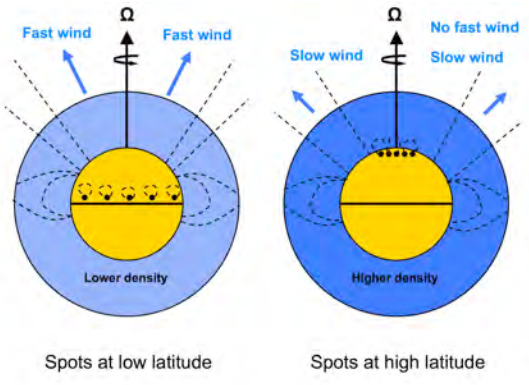


Figure 5: Sketch of the dependence of stellar wind characteristics on surface magnetic topology. With equatorial spots (left) and polar spots (right). As the spots are moved towards the pole, the fast wind is eliminated, leading to increased base density in the corona. This applies more torque on the star due to imperfect co-rotation, leading to a larger angular-momentum loss (Cohen et al. 2009).

One of the major incentives to studying stars is that they allow us access to extremes in parameter values that are not possible to achieve on the Sun. For instance, the G giant FK Com has an equatorial rotational velocity $v\sin i \approx 150$ km s^{-1}. The dynamo exhibits a "flip-flop" behavior where the dominant field appears to switch from one active longitude to another, and is attributed to an axisymmetric dynamo (Berdyugina 2004). The fast rotation has major consequences for the structure of the corona: whereas a potential field extrapolation leads to a complex arcade of magnetic fields that appears as a stronger version of the solar field structure, detailed numerical MHD modeling (Cohen et al. 2010a) shows that the field lines are significantly more twisted and stretched out with a strong toroidal component (see Figure 6, top panels). The twisting and stretching of the fields due to the strong rotation impose a significant strain on the fields, leading to an increase in the magnetic tension (Figure 6, bottom panels), and disconnections in such stretched loops may lead to large-scale flares. Numerical MHD is a new tool whose potential is just now becoming apparent. In combination with the varied parameter space represented by stars, it can transform our physical understanding of all aspects of the solar atmosphere.

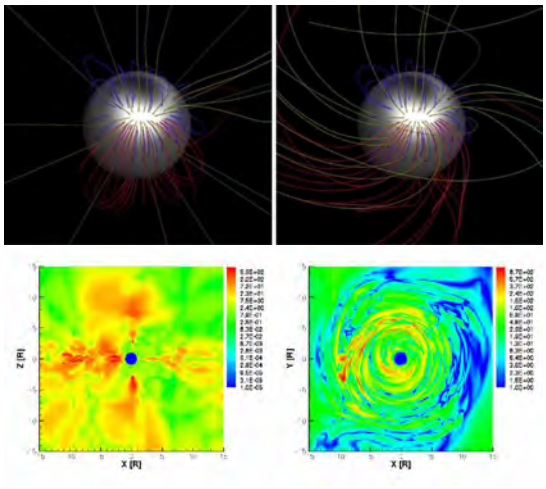


Figure 6: *Top:* Comparison of potential field solution (left) with numerical MHD solution (right) on rapidly rotating star FK Com. The full-fledged, self-consistent solution is dramatically different, with both open (yellow lines) and closed (blue lines) loops stretched and twisted compared to the potential field extrapolation. In many cases, loops that are closed in the potential field solution become open (red lines). *Bottom:* Magnetic tension in the nMHD solution of FK Com, relative to the potential field solution in the $y = 0$ plane (left) and the equatorial $z = 0$ plane (right). The propensity for large scale reconnection events to occur is seen to be significantly higher. (Cohen et al. 2010a)

6 References

• Antia, H.M., & Basu, S., 2005, ApJ, 620, 129
• Argiroffi, C., Maggio, A., & Peres, G., 2003, A&A, 404, 1033
• Argiroffi, C., Peres, G., Orlando, S., & Reale, F., 2008, A&A, 488, 1069
• Audard, M., Güdel, M., Drake, J.J., & Kashyap, V.L., 2000, ApJ, 541, 396
• Audard, M., Güdel, M., Sres, A., Raassen, A.J.J., & Mewe, R., 2003, A&A, 398, 1137
• Ayres, T.R., et al., 1981, ApJ, 250, 293
• Ayres, T.R., Judge, P.G., Saar, S.H., & Schmitt, J.H.M.M., 2008, ApJ, 678, 121
• Bahcall, J.N., Basu, S., Pinsonneault, M., & Serenelli, A.M., 2005, ApJ, 618, 1049
• Baliunas, S., et al., 1995, ApJ, 438, 269
• Basu, S., & Antia, H.M., 2008, Phys.Rep., 457, 217
• Berdyugina, S.V., 2004, Sol.Phys., 224, 123
• Calvet, N., & Gullbring, E., 1998, ApJ, 509, 802
• Chung, S.M., Drake, J.J., Kashyap, V.L., Lin, L.W., Ratzlaff, P.W. 2004, ApJ, 606, 1184
• Cohen, O., Drake, J.J., Kashyap, V.L., Korhonen, H., ELstner, D., & Gombosi, T.I., 2010a, ApJ, 719, 299
• Cohen, O., Drake, J.J., Kashyap, V.L., Sokolov, I.V., & Gombosi, T.I., 2010b, ApJ, 723, 64
• Cohen, O., Drake, J.J., Kashyap, V.L., & Gombosi, T.I., 2009, ApJ, 699, 1501
• Cohen, O., et al., 2007, ApJ, 654, L163
• Cohen, O., Skokolov, I.V., Roussev, I.I., & Gombosi, T.I., 2008, JGR, 113, A03104
• Donahue, R.A., Saar, S.H., & Baliunas, S.L., 1996, ApJ, 466, 384
• Donahue, R.A., & Baliunas, S., 1992, ApJ, 393, 63
• Donati, J.-F., & Landstreet, J.D., 2009, ARA&A, 47, 333
• Drake, J.J., 2003, Adv.Sp.Res., 32, 945
• Drake, J.J., Laming, J.M., & Widing, K.G., 1997, ApJ, 478, 403
• Drake, J.J., & Testa, P., 2005, Nature, 436, 525
• Favata, F., Flaccomio, E., Reale, F., Micela, G., Sciortino, S., Shang, H., Stassun, K.G., & Feigelson, E.D., 2005, ApJS, 160, 469
• Favata, F., & Schmitt, J.H.M.M., 1999, A&A, 350, 900
• Feldman, U., 1992, Phys.Scr., 46, 202
• Feldman, U., & Laming, J.M., 2000, Phys.Scr., 61, 222
• Franciosini, E., Pallavicini, R., & Tagliaferri, G., 2001, A&A, 375, 196
• Güdel, M., 2004, A&ARv, 12, 71
• Güdel, M., Guinan, E.F., & Skinner, S.L., 1997, ApJ, 483, 947
• Güdel, M., Skinner, S.L., Mel'Nikov, S.Y, Audard, M., Telleschi, A., & Briggs, K.R., 2007, A&A, 468, 529
• Henoux, J.-C., 1995, Adv.Sp.Res., 15, 23
• Hudson, H., 1991, Sol.Phys., 133, 357
• Huenemoerder, D.P., Canizares, C.R., Drake, J.J., & Sanz-Forcada, J., 2003, ApJ, 595, 1131
• Huenemoerder, D.P., Canizares, C.R., & Schulz, N.S., 2001, ApJ, 559, 1135
• Hussain, G.A.J., et al., 2007, MNRAS, 377, 1488
• Isola, C., Favata, F., Micela, G., & Hudson, H.S., 2007, A&A, 472, 261 • Kashyap, V., Edmonds, P., Heinke, C., & Gilliland, R., 2006, AAS/HEAD #9, #1.54
• Kashyap, V., Reeves, K., & Wargelin, B.J., 2009, *Chandra's First Decade of Discovery*, 22-25 September, 2009 in Boston, MA, Edited by Scott Wolk, Antonella Fruscione, and Douglas Swartz, #127
• Kashyap, V.L., Drake, J.J., Güdel, M., & Audard, M., 2002, ApJ, 580, 1118
• Kopp & Poletto, 1984, Sol.Phys., 93, 351
• Korhonen, H., & Elstner, D., 2005, A&A, 440, 1161
• Laming, J.M., 2004, ApJ, 614, 1063
• Laming, J.M., 2009, ApJ, 695, 954
• Laming, J.M., Drake, J.J., & Widing, K.G., 1995, ApJ, 443, 416
• Laming, J.M., Drake, J.J., & Widing, K.G., 1996, ApJ, 462, 948
• Laming, J.M., & Drake, J.J., 1999, ApJ, 516, 324

• Liefke, C., Ness, J.-U., Schmitt, J.H.M.M., & Maggio, A., 2008, A&A, 491, 859
• Liefke, C., & Schmitt, J.H.M.M., 2006, A&A, 458, 1
• Meyer, J.-P., 1985, ApJS, 57, 151
• Olah, K., et al., 2009, A&A, 501, 703
• Orlando, S., Peres, G., & Reale, F., 2000, ApJ, 528, 524
• Orlando, S., Peres, G., & Reale, F., 2001, ApJ, 560, 499
• Osten, R.A., Drake, S., Tueller, J., Cummings, J., Perri, M., Moretti, A., & Covino, S., 2007, ApJ, 654, 1052
• Osten, R.A., et al., 2010, ApJ, 721, 785
• Pallavicini, R., Golub, L., Rosner, R., Vaiana, G.S., Ayres, T., & Linsky, J.L., 1981, ApJ, 248, 279
• Parnell, C.E., et al., 2009, ApJ, 698, 75
• Parnell, C.E., & Jupp, P.E., 2000, ApJ, 529, 554
• Peres, G., Orlando, S., Reale, F., Rosner, R., & Hudson, H., 2000, ApJ, 528, 537
• Peres, G., Orlando, S., & Reale, F., 2004, ApJ, 612, 472
• Pesnell, W.D., 2008, Sol.Phys., 252, 209
• Pevtsov, A.A., et al., 2003, ApJ, 598, 1387
• Powell, K.G., et al., 1999, J.Comp.Phys., 154, 284
• Raassen, A.J.J., Ness, J.-U., Mewe, R., van der Meer, R.L.J., Burwitz, V., & Kaastra, J.S., 2003, A&A, 400, 671
• Reale, F., 2007, A&A, 471, 271
• Reale, F., Betta, R., Peres, G., Serio, S., & McTiernan, J., 1997, A&A, 325, 782
• Reale, F., Güdel, M., Peres, G., & Audard, M., 2004, A&A, 416, 733
• Reale, F., Peres, G., Serio, S., Rosner, R., & Schmitt, J.H.M.M., 1988, ApJ, 328, 256
• Reale, F., Peres, G., & Orlando, S., 2001, ApJ, 557, 906
• Reale, F., & Micela, G., 1998, A&A, 334, 1028
• Reeves, K.K., Warren, H.P., & Forbes, T.G., 2007, ApJ, 668, 1210
• Robrade, J., Schmitt, J.H.M.M., & Favata, F., 2008, A&A, 486, 995
• Rosner, R., Golub, L., & Vaiana, G.S., 1984, ARA&A, 23, 413
• Saar, S.H., 2009, *Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology*: GONG 2008/SOHO 21 ASP Conference Series, Vol. 416, p.375
• Saar, S.H., & Brandenburg, A., 2002, AN, 323, 357
• Sanz-Forcada, J., Maggio, A., & Micela, G., 2003, A&A, 408, 1087
• Scelsi, L., Maggio, A., Peres, G., & Pallavicini, R., 2005, A&A, 432, 671
• Schmitt, J.H.M.M., & Favata, F., 1999, Nature, 401, 44
• Stelzer, B., Neuhäuser, R., Casanova, S., & Montmerle, T., 1999, A&A, 344, 154
• Sylwester, B., Sylwester, J., & Phillips, K.J.H., 2010, A&A, 514, 82
• Telleschi, A., Güdel, M., Briggs, K., Audard, M., Ness, J.-U., & Skinner, S.L., 2005, ApJ, 622, 653
• Telleschi, A., Güdel, M., Briggs, K.R., Audard, M., & Scelsi, L., 2007, A&A, 468, 443
• Tsuneta, S., 1995, PASJ, 47, 691
• Wargelin, B.J., Kashyap, V.L., Drake, J.J., Garcia-Alvarez, D., & Ratzlaff, P.W., 2008, ApJ, 676, 610
• Widing, K.G., & Feldman, U., 2001, ApJ, 555, 426
• Winebarger, A.R., Emslie, A.G., Mariska, J.T., & Warren, H.P., 2002, ApJ, 565, 1298
• Wood, B.E., & Linsky, J.L., 2010, ApJ, 717, 1279

A campaign to understand mechanisms responsible for ion heating in magnetic reconnection

A. M. Keesee, P. A. Cassak, and E. E. Scime, *West Virginia University*

Panel themes: Solar Wind-Magnetosphere Interactions (SWM) and Solar and Heliospheric Physics (SH)

Introduction

Magnetic reconnection occurs in several regions of the heliosphere, including in the solar atmosphere, in interplanetary space, at the dayside of planetary magnetospheres, and in planetary magnetotails. In magnetic confinement fusion experiments, the change of magnetic topology can dramatically increase magnetic field components, trigger instabilities, and degrade confinement. Magnetic reconnection is one of the Specific Research Focus Areas pertaining to the Research Objective to Open the Frontier to Space Environment Prediction in the *Science Plan for NASA's Science Mission Directorate 2007-2016*. Because reconnection occurs in the solar atmosphere and at the boundary between the solar wind and Earth's magnetosphere, the proposed campaign contributes to this Panel's themes of Solar and Heliospheric Physics and Solar Wind-Magnetosphere Interactions. Studies of magnetic reconnection are motivated by a quest to understand the essential physics of the Sun and the magnetosphere and to model and predict space weather phenomena that arise from magnetic reconnection.

Laboratory and space studies have found that ion heating is strongly associated with magnetic reconnection. However, the mechanisms responsible for such heating remains an important unanswered question. Because the proposed campaign utilizes previous, current, and planned missions to study ion heating during magnetic reconnection, it is cost-effective and personnel and resources are readily available to participate in the campaign. The proposed campaign complements ongoing research into reconnection in the magnetic fusion community supported by the Department of Energy. We note that although electron heating and acceleration is also an important topic related to magnetic reconnection, the following discussion focuses exclusively on ion heating and acceleration.

Observations of Ion Heating and Acceleration

Satellite observations of ion energization and heating in various locations of the heliosphere present an interesting and puzzling picture. In solar flares, where magnetic reconnection is expected to play the critical role in allowing the conversion of magnetic energy, ions have been observed with energies in the GeV range [*Lin et al.*, 2003; *Emslie et al.*, 2004], which is far greater than the energy associated with the Alfvénic outflow from the reconnection region. In observations by ISEE 1 and 2 at the dayside magnetopause, ion heating was observed in association with accelerated flow events [*Gosling et al.*, 1986]. In the magnetotail, bursts of high energy protons were observed with Imp 7 [*Sarris et al.*, 1976]. High speed ion beams at the plasma sheet boundary layer were interpreted as ion acceleration in the reconnection electric fields [*Hoshino et*

al., 1998]. These observations give indirect evidence that accelerated and heated ions are often found near reconnection sites, suggesting that reconnection plays an active role in energizing the particles. In spatially extended reconnection events in the solar wind, a noticeable but relatively small amount of ion heating was observed [*Phan et al*., 2006]. At present, it is not fully understood how ion energization occurs during magnetic reconnection, nor why it is robust in some settings but not in others.

The heating and acceleration of ions during magnetic reconnection has also been addressed in laboratory experiments in various configurations. *Scime et al.* [1992] measured ion heating associated with reconnection in the Madison Symmetric Torus (MST) reversed field pinch, finding that the heating varied with ion species, suggesting that the heating mechanism has mass and/or charge dependence. *Hsu et al.* [2000] measured ion heating in a 2D reconnection layer in the Magnetic Reconnection Experiment (MRX). They attributed the heating to nonclassical mechanisms, such as wave-particle interactions or collisionless inertial effects because viscous heating by the sub-Alfvénic ion outflow and the energy from Ohmic dissipation were both insufficient to account for the observed heating. *Stark et al.* [2005] performed laser-induced fluorescence measurements of ion velocity distribution functions parallel to the X-line of a laboratory reconnection experiment, showing that the maximum ion heating occurred when the reconnection rate was a maximum. They performed particle simulations that showed ion heating was connected to the in-plane electric field.

Theoretical Work on Ion Energization during Magnetic Reconnection

There have been a number of theoretical and numerical studies of ion energization in reconnection events, including a recent flurry of activity. A long-standing idea is that particles can be accelerated in the reconnection electric field where they are unmagnetized in the dissipation region, going back to the original work on flares by *Giovanelli* [1948]. It is unlikely that this mechanism is a strong contributor to energetic ions in flare events because the dissipation regions are of very small spatial extent and that the charge separation would induce an electric field that would short out the reconnection. However, *Hoshino et al.* [1998] argued that acceleration in the reconnection electric field can lead to observable ion acceleration in magnetospheric events.

The acceleration of ions also occurs in outflow jets of reconnection [*Krauss-Varban and Welsch*, 2007; *Drake et al.*, 2009a]. The idea is that when a cold ion upstream of the dissipation region is on a field line that reconnects, it enters the downstream region and is immersed in the outward directed flow. The ion acts like a pickup particle [*Mobius et al.*, 1985], and gains parallel and perpendicular thermal energy. It has been shown that when there is an out-of-plane (guide) magnetic field present in the reconnection [*Drake et al.*, 2009b], the ion energization is dependent on the charge-to-mass ratio of the ions, which may be important for understanding the observed mass dependent energization in flares and the solar wind and in laboratory experiments .

Ions can also be accelerated in magnetic islands (plasmoids) [*Matthaeus et al.*, 1984; *Kleim*, 1994; *Shibata and Tanuma*, 2001; *Onofri et al.*, 2006; *Drake et al.*, 2006]. The idea is very simple - a newly reconnected field line between two X-lines contracts, and as it does so, it accelerates the particles traversing the field line in a first order Fermi process. This process works for both ions and electrons, is very efficient, and can potentially bring ions and electrons to very high energy if there are sufficient numbers of magnetic islands. A recent study of electron acceleration [*Oka et al.*, 2010] identified self-consistently generated energetic particles in reconnection electric fields within magnetic islands, which may also contribute to ion energization.

Proposed Action

We propose a multidisciplinary campaign to address the mechanisms of ion heating in magnetic reconnection. The campaign will combine analysis from space-based observational measurements (from previous, current, and planned missions), simulations, and laboratory experiments to answer the following questions:

1. In each region, which characteristics of ion heating during reconnection are persistent and which vary? Under what conditions?

   To address this question, statistical surveys must be performed within each region of interest including the Sun, solar wind, magnetopause, and magnetotail. Important questions to address in these surveys include:

   Does ion heating always occur during reconnection? If not, under what conditions does it occur?

   How does ion heating vary by species?

   How much ion heating occurs and what conditions control it?

   Each region of interest has relevant missions (previous, current, or planned) that can be used for these surveys. Ion temperatures can be measured by instruments on Solar Probe Plus near the Sun; WIND and ACE in the solar wind; THEMIS, Cluster, and MMS at the dayside magnetopause; and THEMIS, ARTEMIS, Geotail, and MMS in the magnetotail. The ion temperature measurements can be combined with measurements of relevant conditions such as magnetic field and density to determine what plays a role in driving the ion heating.
2. What are the similarities and differences in the characteristics of ion heating during reconnection at the Sun, solar wind, magnetopause, and magnetotail?

   To address this question, the statistical studies from Question 1 could be compared. Understanding how the heating and the conditions that drive the heating vary among the regions could shed light on whether the heating mechanisms are the same or different in each setting. Similar studies could be performed in laboratory experiments, such as the Magnetic Reconnection Experiment (MRX), the Reconnection Scaling Experiment (RSX), the Madison Symmetric Torus (MST), and the soon to be completed the Florida A&M University Spheromak (FAMUS). A comparison of ion heating characteristics and conditions with the statistical studies described above will help determine which laboratory experiments are relevant to which regions of the heliosphere. For example, the laboratory studies discussed above reported ion heating results that varied greatly in magnitude. These differences may indicate an important set of

conditions that drive more significant levels of heating or even different heating mechanisms. Simulations could also be used to understand the underlying physics for each region.

3. What is the spatial distribution of ion heating?

   Laboratory experiments provide the ability to measure ion temperatures with spatial resolution that is not possible with space-based measurements. Determining where the heating occurs in relation to the reconnection site is important to understanding the heating mechanism. Results from experiments can be used to place multi-point space based measurements (e.g. Cluster and MMS) in context.

4. What is the time scale of the heating?

   Because laboratory experiments can yield simultaneous measurements of reconnection processes and particle heating, they can provide insight into particle heating processes through a detailed understanding of the timescales over which heating occurs. In the reversed field pinch experiments discussed above, there are strong indications that different mass ions are heated at different rates as well as to different final temperatures, as seen in the solar corona. In-situ magnetospheric and heliospheric data can be analyzed in terms of timescales to determine if similar phenomena occur in space plasmas. Simulations can also be used to understand time scales.

References

Drake, J. F., M. Swisdak, H. Che, and M. A. Shay, Electron acceleration from contracting magnetic islands during reconnection, *Nature*, 443, 553, 2006.

Drake, J. F., M. Swisdak, T. D. Phan, P. A. Cassak, M. A. Shay, S. T. Lepri, R. P. Lin, E. Quataert, and T. H. Zurbuchen, Ion heating resulting from pickup in magnetic reconnection exhausts, *J. Geophys. Res., 114*, A05111, 2009a.

Drake, J. F., P. A. Cassak, M. A. Shay, M. Swisdak, and E. Quataert, Magnetic reconnection mechanism for ion acceleration and abundance enhancements in impulsive flares, *Ap. J.*, 700, L16, 2009b.

Emslie, A. G., et al., Energy partition in two solar flare/CME events, *J. Geophys., Res.*, 109, A10104, 2004.

Giovanelli, R. G., Magnetic and Electric Phenomena in the Sun's Atmosphere Associated with Sunspots, *MNRAS*, 107, 338, 1947.

Gosling, J.T., M.F. Thomsen, S.J. Bame, and C.T. Russell, Acceleratied plasma flows at the near-tail magnetopause, *J. Geophys. Res., 91*, 3029, 1986.

Hoshino, M. T. Mukai, and T. Yamamoto, Ion dynamics in magnetic reconnection: Comparison between numerical simulation and geotail observations, *J. Geophys. Res.*, 103, 4509, 1998.

Hsu, S.C., G. Fiksel, T.A. Carter, H. Ji, R.M. Kulsrud, and M. Yamada, Local measurement of nonclassical ion heating during magnetic reconnection, *Phys. Rev. Lett., 84*, 3859, 2000.

Kliem, B., Particle orbits, trapping and acceleration in a filamentary current sheet model, *Ap. J.*, 90, 719, 1994.

Krauss-Varban, D. and B. T. Welsch, Solar flare particle heating via low-beta reconnection, in *Proceedings of the International Astronomical Union, vol. 2, Highlights of Astronomy*, vol. 14, edited by K. A. van der Hucht, Cambridge Univ. Press., Cambridge, U. K., 2007.

Lin et al., RHESSI observations of particle acceleration and energy release in an intense solar gamma-ray line flare, *Ap. J.*, 595, L69, 2003.

Matthaeus, W. H., J. J. Ambrosiano, and M. J. Goldstein, Particle acceleration by turbulent magnetohydrodynamic reconnection, *Phys. Rev. Lett.*, 53, 1449, 1984.

Mobius, E., D. Hovestadt, B. Klecker, M. Scholer, G. Gloeckler, and F. M. Ipavich, Direct observation of He^+ pick-up ions of interstellar origin in the solar wind, *Nature*, 318, 426, 1985.

Oka, M., M. Fujimoto, I. Shinohara, and T. D. Phan, "Island surfing" mechanism of electron acceleration during magnetic reconnection, *J. Geophys. Res.*, 115, A08223, 2010.

Onofri, M., H. Isliker, and L. Vlahos, Stochastic acceleration in turbulent electric fields generated by 3D reconnection, *Phys. Rev. Lett.*, 96, 151102, 2006.

Phan, T.D., J.T. Gosling, M.S. Davis, R.M. Skoug, M. Øieroset, R.P. Lin, R.P Lepping, D.J. McComas, C.W. Smith, H. Reme, and A. Balogh, A magnetic reconnection X-line extending more than 390 Earth radii in the solar wind, *Nature, 439*, 175, 2006.

Priest, E.R., C.R. Foley, J. Heyvaerts, T.D. Arber, J.L. Culhane, and L.W. Acton, Nature of the heating mechanism for the solar corona, *Nature, 393*, 545, 1998.

Sarris, E. T., S. M. Krimigis, and T. P. Armstrong, Observations of magnetospheric bursts of high-energy protons and electrons at ~35 RE with Imp 7, *J. Geophys. Res.*, 81, 2341, 1976.

Scime, E., S. Hokin, N. Mattor, and C. Watts, Ion heating and magnetohydrodynamic dynamo fluctuations in the reversed-field pinch, *Phys. Rev. Lett.*, *68*, 2165, 1992.

Shibata, K. and S. Tanuma, Plasmoid-induced-reconnection and fractal reconnection, *Earth Planets Space*, 53, 473, 2001.

Stark, A., W. Fox, J. Egedal, O. Grulke, and T. Klinger, Laser-induced fluorescence measurement of the ion-energy-distribution function in a collisionless reconnection experiment, *Phys. Rev. Lett.*, *95*, 235005, 2005.

Science and Operation of the Advanced Technology Solar Telescope

Stephen Keil, Thomas Rimmele (National Solar Observatory), Jeff Kuhn (University of Hawaii), Phil Goode (New Jersey Institute of Technology), Robert Rosner ((University of Chicago), Michael Knoelker (High Altitude Observatory), Thomas Berger (Lockheed Martin)

.

1. INTRODUCTION

The Advanced Technology Solar Telescope (ATST) is an identified project in the NSF MREFC program and construction began in 2010. The NSF provided $7 M in FY09 and $13M in FY2010 MREFC funds. In addition Congress provided $146M in American Recovery and Reinvestment Act (ARRA or stimulus) funding. The ATST represents the culmination of the solar community's desire going back over several decades to have a solar telescope capable of resolving fundamental scales and processes on the photosphere and to finally be able to measure the "dark energy" magnetic fields that dominate the solar corona The 2000 Astronomy and Astrophysics Decadal Survey[1] stated that "The first scientific goal for advancing the current understanding of solar magnetism is to measure the structure and dynamics of the magnetic field at the solar surface down to its fundamental length scale – ATST is designed to achieve this angular resolution." The 2003 NSF/NASA Solar and Space Physics Decadal Survey[2] recognized the high priority of the ATST and its synergism with space missions as essential for solving the questions surrounding solar magnetism and solar activity. With its completion in 2017, strong NSF support will be needed to operate ATST and to exploit its scientific capabilities.

2. KEY SCIENTIFIC GOALS

Magnetic fields control the inconstant Sun. The key to understanding solar variability rests with understanding all aspects of these magnetic fields. The majority of magnetic energy, predicted by models and sophisticated simulations to exist on very small spatial scales (< 0"1) and often with weak field strength, cannot be sensed with current telescopes. Up until now it is only from photospheric and chromospheric observations that we have direct measurements of these magnetic fields. As the world's largest solar telescope and coronagraph the photon collecting power of the ATST's 4m aperture is essential to make progress. For the first time we will have the capability to measure solar magnetism from the photosphere to far into the corona, where the solar wind is accelerated and CME's are launched. But even in the lower atmospheric layers, the magnetic fields change so rapidly that current telescopes can not resolve many important aspects of the field before it has evolved and the bulk of the field may exist on scales that have not been resolved. In higher atmospheric layers our current ability to accurately measure the magnetic field is rudimentary at best. ATST observations are not incremental improvements in our understanding of the Sun, but paradigm changing visions

The photosphere is a crucial region where energy is readily transformed from convective motion into thermal and magnetic energy, and electromagnetic radiation. The energy stored in magnetic fields is eventually dissipated at higher layers of the solar atmosphere, sometimes in the form of

violent flares and coronal mass ejections (CMEs) that ultimately affect the Earth and drive space weather. The photosphere, the chromosphere, transition region, and the corona are connected through the magnetic field and therefore have to be treated as one system, rather than as individual layers. The ATST with high performance adaptive optics and state-of-the-art instrumentation is a crucial tool to understand this complex, interconnected physical system. Some of the scientific problems the ATST will address include the origin and generation of magnetic fields, magnetic activity and instability, chromospheric and coronal structure and heating, and sources of solar irradiance variability[3].

2.1. FLUX TUBES, THE BUILDING BLOCKS OF STELLAR MAGNETIC FIELDS

Observations have established that the photospheric magnetic field is organized in small fibrils or flux tubes. This small scale flux may be the dominate form taken by the magnetic field and could be a major contributor to the energy balance of the atmosphere, in particular heating of the outer atmospheric layers and a source of energy for the solar wind. Magnetic structures are mostly unresolved by current telescopes. Flux tubes are the most likely channels for transporting energy into the upper atmosphere, which is the source of UV and X-ray radiation from the Sun, which in turn affects the Earth's atmosphere.

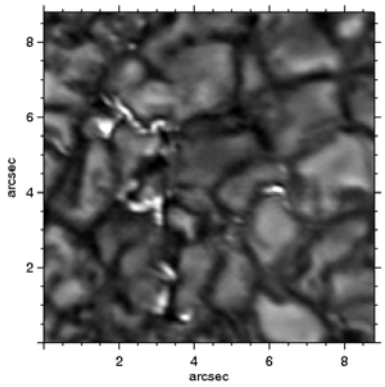


Fig 1. Solar granules and bright magnetic flux regions observed in the g-band of CN with the Dunn Solar Telescope. Measuring the vector magnetic field associated with the rapidly changing bright points needs a larger aperture.

Addressing the issue of small scale magnetic fields requires simultaneous vector magnetometry in multiple layers of the solar atmosphere and drives the spatial (0.03" at 500nm and 0.08" at 1.6 microns) and temporal (periods as short as a second or less) resolution requirements. At this resolution ATST will provide the required spectroscopy and polarimetry at an angular resolution to explore the enigmatic flux tube structures on time scales compatible with their lifetimes. For the first time, ATST will achieve the resolution and photon flux required to investigate the role of flux annihilation and MHD wave modes in generating and transporting energy to the outer solar atmosphere. ATST will provide direct measurements of magnetic flux cancellation rates in the quiet sun from magnetograms that have 16 times better resolution (pixel area) or 30 times better sensitivity (or a compromise between the two) with the 4m ATST than with the current largest high-resolution solar telescopes. The ATST will provide spectroscopic observations of sufficient resolution to study the formation, internal structure, interactions with convection and disappearance of flux tubes and what role they play in the heating of the outer atmosphere

2.2. MAGNETIC FIELD GENERATION AND LOCAL DYNAMOS

To understand solar activity and solar variability, we need to understand how magnetic fields are generated and how they are destroyed. The 11-year sunspot cycle and the corresponding 22-year magnetic cycle are still shrouded in mystery. There are two types of dynamo. The large scale or *mean field dynamo* generates magnetic fields at scales larger than the energy-containing scale of

the turbulence, whereas *the small-scale dynamo* amplifies magnetic fluctuation energy below the energy-containing scale of the turbulence. The small-scale dynamo is usually a much faster process than the mean-field dynamo, and the large-scale field produced by the latter can be treated as approximately constant on the timescale of the small-scale dynamo. However, the mean-field dynamo also gives rise to small-scale magnetic fluctuations because of the turbulent shredding of the larger scale magnetic field, i.e., the large scale and small scale dynamos interact and some solar cycle dependence of the small scale fields might be expected.

The Sun's magnetic machine involves the processes of flux generation, transport, dissipation, instability and ejection. The 11-year sunspot cycle and the 22-year magnetic cycle are still shrouded in mystery. Global dynamo models attempt to explain large-scale solar magnetic fields are based on theories involving large averages. Turbulent convection is believed to stretch and amplify the magnetic field while differential rotation, meridional circulation, and the tachocline shear layer at the base of the convection zone is believed to organize the large-scale solar magnetic field of several 100 kgauss. This field intermittently rises up to the surface and emerges on scales of active regions down to intergranular lanes. However, properties such as diffusion and helicity, that must be assumed in these models are produced by small-scale turbulent processes and have not yet been measured. High spatial, temporal, and spectral resolution is needed to observe the turbulent vorticity and the diffusion of small-scale magnetic fields to measure these properties and how they evolve with the solar cycle.

In addition, local dynamos may produce the majority of small-scale magnetic flux elements recently observed to cover the entire Sun[4,5]. This "magnetic carpet" continually renews itself on a time scale of a few days at most and its flux may be comparable to that in active regions. The ATST will make it possible to directly observe such local dynamo action at the surface of the Sun.

ATST's high time-cadence, high spatial resolution vector magnetograms will provide crucial information about evolution of magnetic field twist during flux emergence. To detect helicity "pumping" by Alfven waves will require a computation of electric currents inside individual "flux tubes" -- a task well outside the scope of existing ground-based telescopes or existing and future space instruments. ATST will provide ultra-high resolution vector magnetograms allowing direct observational test of the hypothesized local dynamo. In addition, the high resolution and high time-cadence white-light images will allow direct comparison of kinetic helicity of local dynamo flows and helicity of magnetic fields generated by these flows

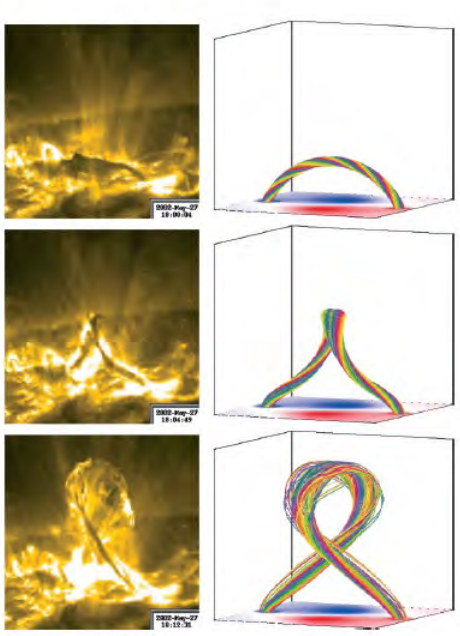


Fig 2. Filament eruption: Trace image sequence (left), model of magnetic field (right.) Actual measurements of the coronal field are missing. What coronal field configurations lead to CMEs and what are the trigger mechanisms?

2.3. INTERACTION OF MAGNETIC FIELDS AND MASS FLOWS

Sunspots are the ideal objects for testing magneto-hydrodynamic theory under astrophysical conditions. In sunspots, the total magnetic field is large enough to completely dominate the hydrodynamic behavior of the local gas, a regime very different from that of the rest of the solar photosphere. Numerical simulations and theoretical models predict dynamical phenomena, such as oscillatory convection in the strong-field regions of sunspot umbrae, flows at the speed of sound along penumbral filaments and oscillations and wave phenomena. To verify the predictions of numerical simulations of sunspots and ultimately answer such fundamental questions as "Why do sunspots exist?" requires an extremely capable instrument. High-resolution (<0.1 arcsec) vector polarimetry combined with high sensitivity (high photon flux) and low-scattering optics are required.

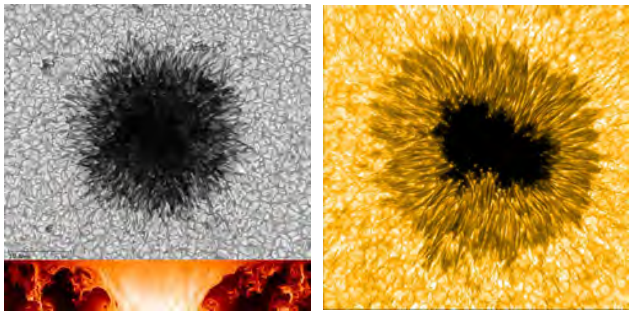


Figure 3. (right) A high resolution sunspot image using adaptive optics and the Dunn Solar Telescope. (left) A state-of-the-art MHD simulation of a sunspot penumbra (Rempel, HAO).

The ATST will be the first telescope able to collect precision polarimetric data with a temporal cadence necessary to capture the evolution of sunspot fine structure and finally understand its physical origin. In sunspot umbrae and penumbrae, optical observations are plagued by contamination of scattered light from the bright surrounding photosphere. The ATST is carefully designed to allow observations of high resolution and low scattered light. In addition, the 4-m aperture will allow high-resolution observations in the infrared (e.g., 0.08 arcsec at 1.6 μm), where scattered light is less of a problem and much more accurate magnetic field fine-structure determination is possible.

2.4. FLARES AND MASS EJECTIONS

An important goal for the ATST will be to study origins and mechanisms of solar activity such as the small-scale and rapid processes in solar flares and configurations leading to flares and CMEs[6]. It is commonly believed that solar flares represent a process of rapid transformation of the magnetic energy of active regions into the kinetic energy of energetic particles and plasma flows and heat. Detection of variations of magnetic field associated with solar flares has been one of the most important problems of solar physics for many years. Such detection would provide direct evidence of magnetic energy release in the flares. An important goal for the ATST will be to study the small-scale processes in solar flares. The ATST will also provide a new set of tools, in particular in the infrared, to measure magnetic fields at higher layers of the atmosphere. There is limited observational evidence that the distribution of electric currents and

current helicity inside an active region varies with flares. Highly uniform sequences of high-resolution vector magnetograms of an active region before and after a flare are required to address this important issue.

Coronal mass ejections (CMEs) originate in large-scale magnetic arcades known as helmet streamers. These structures are known to contain twisted magnetic fields. According to the prevailing view, the arcade becomes dynamically unstable when its fields are twisted beyond some critical point. Field line footpoint motions in the photosphere have long been considered efficient ways to supply (or drain) magnetic shear and energy into (from) the coronal field. MHD simulations have identified critical magnetic shear conditions above which the arcade field will form current sheets, and magnetic reconnection processes will occur to cause active phenomena such as flares, CMEs, and prominence formation and eruption. It is thus important to measure field line footpoint motions and, if possible, the magnetic shear in active regions as well. Although measurements of the footpoint motion (in particular the horizontal plasma flow velocity) have improved considerably, the ATST will offer unprecedented high spatial and temporal resolution in measuring the field line footpoint motion.

More recently this view has been challenged by models in which the arcade emerges as a twisted flux rope or models in which small-scale photospheric reconnection events inject helicity into the corona. These hypothesized reconnection events occur when small (0.1”) photospheric flux elements cancel along the active region's magnetic neutral line. This model seems to contradict those models in which CMEs result from excessive twist in the arcade. Detailed observations of the flux emergence can reveal whether the emerging flux is introducing magnetic twist into the arcade, or changing the arcade's topology through footpoint cancellation. Only with higher spatial resolution vector measurements and good temporal resolution as will be provided by the ATST, however, can it be established, for example, that the rate and orientation of these cancellations is consistent with an observed change in the twist of the overlying arcade. Such observations are critical to distinguishing between competing models.

2.5. INHOMOGENEOUS STELLAR UPPER ATMOSPHERES

Measurements of CO absorption spectra near 4.7 μm show surprisingly cool clouds that appear to occupy much of the low chromosphere[7]. Only a small fraction of the volume apparently is filled with hot gas, contrary to classical static models that exhibit a sharp temperature rise in those layers. The observed spectra can be explained by a new class of dynamic models of the solar atmosphere. However, the numerical simulations indicate that the temperature structures occur on spatial scales that cannot be resolved with current solar infrared telescopes. A test of the recent models requires a large-aperture solar telescope that provides access to the thermal infrared. Spicules, the forest of hot jets that penetrate from the photosphere into the chromosphere, are clearly a MHD phenomenon that is not understood nor adequately modeled. Their role in the mass balance of the atmosphere is uncertain. Combined with UV observations from space, the ATST will allow us to resolve their nature.

2.6. MAGNETIC FIELDS AND STELLAR CORONAE

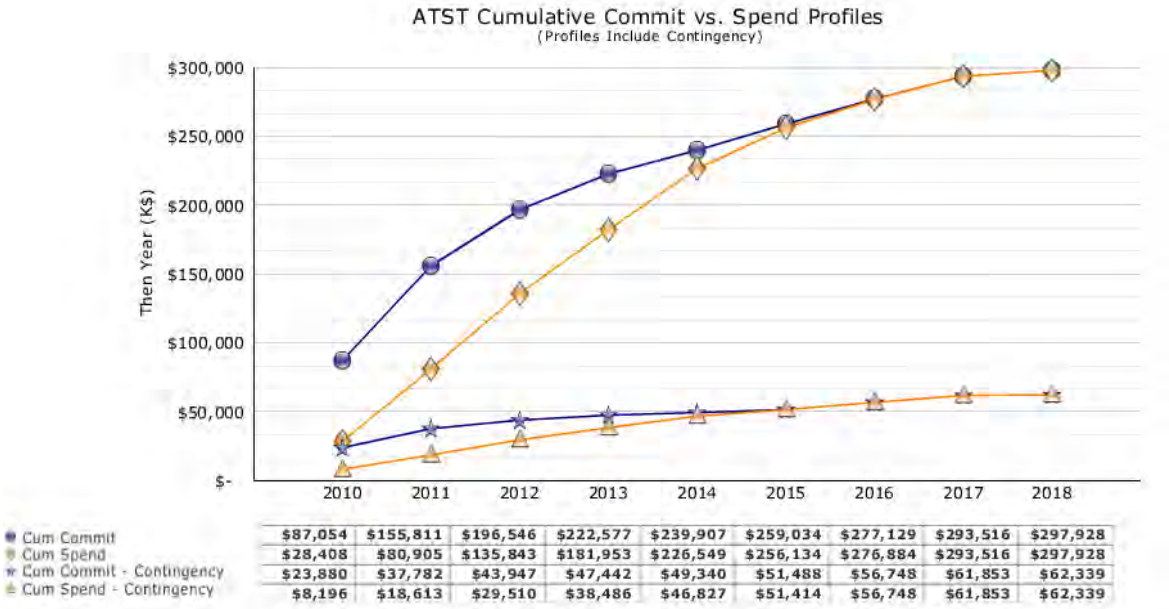
The origin and heating of the solar corona are still mysteries. Most of the proposed scenarios are based on dynamic magnetic fields rooted at the 0.1-arcsec scale in the photosphere. However, none of the processes has been clearly identified by observations or theory. EUV and X-ray

observations have gained in importance, but ground-based observations are still critical, not only to determine the forcing of the coronal fields by photospheric motions, but also for the measurement of the coronal magnetic field strength itself. This is important for developing and testing models of flares and coronal mass ejections, which propel magnetic field and plasma into inter-planetary space and induce geomagnetic disturbances. In particular, precise measurements of the coronal magnetic field strength and topology are needed in order to distinguish between different theoretical models. The ATST with its large aperture, low scattered light characteristics, and the capability to exploit the solar infrared spectrum will provide these critical measurements.

The ATST will also address many other problems in chromospheric and coronal structure and heating. The ATST will have an unprecedented multi-spectral imaging capability with which true three-dimensional mapping will become routine. Simultaneous observations will regularly be conducted at more than one wavelength in the whole range of near-ultraviolet, visible and infrared. Because the optical path in the ATST is fully reflective it is the first big telescope that will be able to observe at all these wavelengths. Its coronal capability will allow it to map coronal fields and plasma properties of coronal loops and structures

3. COST AND STATUS

The costs for the ATST construction phase are based on the Work Breakdown Structure (WBS) down to eight levels of detail, depending upon the complexity and availability of comparable existing systems. The construction phase will cover eight fiscal years, beginning in FY 2010 and ending in FY 2017. The vast majority of the work will be subcontracted competitively to industry and to qualified partners with unique expertise. ATST personnel manage all activities centrally.



Cum Commit	$87,054	$155,811	$196,546	$222,577	$239,907	$259,034	$277,129	$293,516	$297,928
Cum Spend	$28,408	$80,905	$135,843	$181,953	$226,549	$256,134	$276,884	$293,516	$297,928
Cum Commit - Contingency	$23,880	$37,782	$43,947	$47,442	$49,340	$51,488	$56,748	$61,853	$62,339
Cum Spend - Contingency	$8,196	$18,613	$29,510	$38,486	$46,827	$51,414	$56,748	$61,853	$62,339

Current planning, based on an FY 2010 construction start, has calendar year 2017 targeted for obtaining the first scientific data with an ATST instrument. During the first two years of construction, site work, manufacture of the primary mirror blank and completion of the final fabrication designs are primary objectives of the project. Construction of main components such as the enclosure and telescope structure should also be well underway during this period.

Operational costs for ATST have been estimated based from a bottoms-up exercise looking at the task that must be performed for operations, maintenance, and ongoing instrument development. The increased cost of living on Maui was factored into the costs. The payroll to maintain a staff of approximately 40 on Maui is estimated at ~$7.6M. Non-payroll costs are estimated at $4.2M. An addition $2M per year for instrumentation and telescope upgrades would be required, giving a total estimate of ~$13.8M. This is roughly 5% of the capitol cost.

Funding to support the use of ATST by the solar community will be critical. NSF needs to establish a strong grants program that will permit investigators to propose for funding that will permit them to work with ATST data. NSO plans to make all ATST data available to the community similar to data from the HINODE and SDO missions.

4. SUMMARY

High-resolution studies of the Sun's magnetic fields are needed for a better understanding of the fundamental processes responsible for solar variability. The generation of magnetic fields through dynamo processes, the amplification of fields through the interaction with plasma flows, and the destruction of fields are poorly understood. There is incomplete insight into physical mechanisms responsible for chromospheric and coronal structure and heating, causes of variations in the radiative output of the Sun, and mechanisms that trigger flares and coronal mass ejections. Progress in answering these critical questions requires study of the interaction of the magnetic field and convection with a resolution sufficient to observe scale fundamental to these processes. The 4-m aperture ATST will be a unique scientific tool, with excellent angular resolution, a large wavelength range, and low scattered light. Scientific operation and exploitation of the ATST will require strong NSF support starting during the period covered by this decadal survey.

5. REFERENCES

1. McKee, C. F. and Taylor, J. H. (Co-chairs). Astronomy and Astrophysics Survey Committee, Board on Physics and Astronomy, Space Studies Board, National Research Council (2001). Astronomy and Astrophysics in the New Millennium. Washington, DC: National Academy Press.
2. Lanzerotti, L. J. (Chair). Solar and Space Physics Survey Committee, National Research Council (2003). The Sun to the Earth -- and Beyond: A Decadal Research Strategy in Solar and Space Physics. Washington, DC: National Academy Press.
3. Rimmele, T.R, et al. 2008, Advances in Space Research, Volume 42, Issue 1, p. 78-85
4. Rimmele, T.R, Keil, S.L., and Deforest, C., 2009, ASTRO 2010 Science White Paper
5. Schrijver, C. et al., 2009, ASTRO 2010 Science White Paper
6. Giampapa, M.S, et al., 2009, ASTRO2010 Science White Paper
7. Ayers, T., et al. 2009, ASTRO2010 Science White Paper

GENERATION, EVOLUTION AND DESTRUCTION OF SOLAR MAGNETIC FIELDS

(Measuring & Modeling the small-scale magnetic flux of the Sun)

Thomas Rimmele (trimmele@nso.edu) & Stephen Keil (skeil@nso.edu)
National Solar Observatory, Sunspot NM 88349-0062
Craig DeForest
Southwest Research Institution

Introduction

Magnetic fields control the inconstant Sun and play a major role throughout heliophysics. The complex magnetic field that threads the solar photosphere shapes the Sun's entire atmosphere and provides the power for its hearty and frequent explosions (Metcalf et al. 2008). The interaction of interior convection and the magnetic field produces stress that leads to magnetic heating of the layers above. Thus, understanding all aspects of the solar magnetic field is crucial to understanding solar variability and its direct impact on the Earth.

Magnetic fields are important in the heating and energetics of both the quiet and active sun. To understand the magnetically driven energization of the solar atmosphere, it is essential to have an accurate quantitative observational characterization of the magnetic fields and their evolutionary properties, as well as accurate models to obtain physical parameters.

In the following sections, we introduce the Sun's small scale field and how it is measured, describe several current results and questions in the science of the solar magnetic field, and discuss modeling and instrument capabilities required to tackle them

The Solar magnetic field

The Sun exhibits two general patterns of magnetic field: an oscillating global field related to the 11 year sunspot cycle and evidenced by coherent, patterned emergence, breakup, and diffusion of sunspots and their associated active regions; and a universal, apparently random background of small concentrations of magnetic flux on all observable scales to the limit of current observations. The statistical properties of the solar magnetic field are needed to test MHD simulations (Stenflo 2010). Because magnetic structuring continues on scales that are much smaller than the telescope resolution attempts to describe these properties have been ambiguous. Observed magnetic poles vary in size and flux from ~100km diameter features with $\sim 10^{15}$ Mx, through large active regions with a few times 10^{22} Mx. At scales ~10Mm, flux concentrated into downflows between supergranules forms the solar magnetic network of flux concentrations that survive tens of hours; on smaller scales shorter-lived "intranetwork" concentrations exist within supergranules; and below ~1Mm, magnetic flux concentrations can be seen forming in the intergranular lanes, analogous to (but shorter lived than) the supergranular network. This small scale flux appears to be produced independently the solar cycle (e.g. Parnell et al. 2009 & refs therein), and is collectively referred to as the solar "magnetic carpet" (Title & Schrijver 1998). Neither the global nor the background magnetic field is well understood theoretically, though recent advances in modeling and measurement have yielded new insight.

Measurement

Magnetic fields on the Sun are currently measured via the Zeeman splitting, typically in a photospheric spectral line. Spectropolarimetry yields the most sensitive and complete diagnostics of the magnetic field, but yields small fields of view and/or spatiotemporal

resolution. Spectral imaging techniques such as scanned filters or Fourier tachometry provide either line-of-sight magnetic field measurements (via the antisymmetric Zeeman splitting in Stokes V) or full vector measurements (via the symmetric splitting in Stokes Q and U) over an entire image plane simultaneously. New instrument techniques, such as image slicing (Lin 2006) or spectral stereoscopy (DeForest et al. 2004), are under current development to enable higher sensitivity. Magnetic field measurements that use the Hanle scattering-depolarization effect are under active development by several groups (e.g. Stenflo et al. 2002, Trujillo-Bueno et al. 2006).

Understanding the solar dynamo

Understanding the solar dynamo has applications well beyond solar and stellar physics, because of the limitations of current dynamo theory and modeling.

The magnetic and kinematic Reynolds numbers Re_m and Re, and their ratio (the Prandlt number P_m), determine dynamo behavior. In hot rarefied plasmas, such as the warm and hot phases of the ISM, $P_m>>1$ and a small-scale dynamo is theoretically well established. In contrast, $P_m<<1$ in the Sun's convective zone, in planets, and in proto-stellar disks, all of which have disordered fluctuating small scale magnetic fields and are expected to be in a turbulent state (e.g., Brandenburg & Subramanian 2005, Schekochihin et al. 2004,2005,2007). This case is neither as well explored nor as well understood as $P_m>>1$. For example, the presence of the magnetic carpet on the sun is likely related to small-scale dynamo action as numerically simulated by Cattaneo 1999; Cattaneo et al. 2003; Stein and Nordlund, 2006, Voegler & Schuessler 2007. However, these simulations all have $P_m \geq 1$ (Figure 1).

To understand solar activity and solar variability as well as magnetism of other heliospheric and astrophysical systems, we need to understand these processes in detail. Because quite small scale turbulent fields can in principle be resolved on its surface, the Sun provides us with a unique "laboratory" to advance our understanding of dynamo processes.

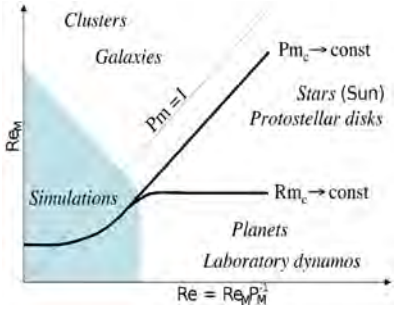


Figure 1: The stability curve of turbulent small-scale dynamo vs. Re and Re_m. Numerical effects limit current simulations to regimes that are more viscous and resistive than astrophysical systems such as the Sun, requiring direct observation to make progress.

There are two generic types of dynamo. A large-scale or *mean field dynamo* generates magnetic fields at scales larger than the energy-containing scale of the turbulence, whereas a *small-scale dynamo* amplifies magnetic fluctuation energy below the energy-containing scale of the turbulence. The small-scale dynamo is usually a much faster process than the mean-field dynamo, and the large-scale field produced by the latter can be treated as approximately constant on the timescale of the small-scale dynamo. It has been hypothesized (e.g. Voegler & Schuessler 2007) that both classes of dynamo exist in the Sun, with a separate small-scale process driven by convective turnover and a global dynamo controlled by bulk flows. These two scales of dynamo are not wholly independent: they are coupled by shredding and aggregation of flux. On the Sun, there is a cross-scale equilibrium between these processes (Schrijver et al. 1997). Currently unresolved physics dominates both the solar network and the smallest currently-resolvable (100km) flux concentrations (Lamb et al. 2008, 2008b).

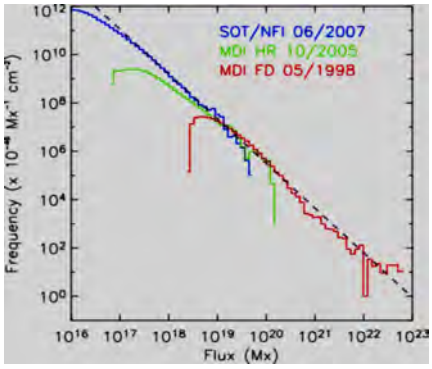


Figure 2. Flux concentration frequency vs. flux, detected by computer vision, follows a -1.8 power law across three instruments and nearly six decades in flux (Parnell et al. 2009). Is the slope break at the small end from instrument effects, or from the Sun?

Small scale activity on the Sun would influence the global dynamo even in the absence of direct coupling between the large and small scale dynamos. Models that attempt to explain large-scale solar magnetic fields are based on theories involving large averages (mean field, Weiss and Thompson, 2008). Turbulent convection is believed to stretch and amplify the magnetic field while differential rotation, meridional circulation, and the tachocline shear layer at the base of the convection zone are believed to organize a deep, large-scale solar magnetic field of several 100 kgauss. Properties, such as diffusion and helicity, that must be assumed in these models are produced by small-scale turbulent processes and have not yet been measured.

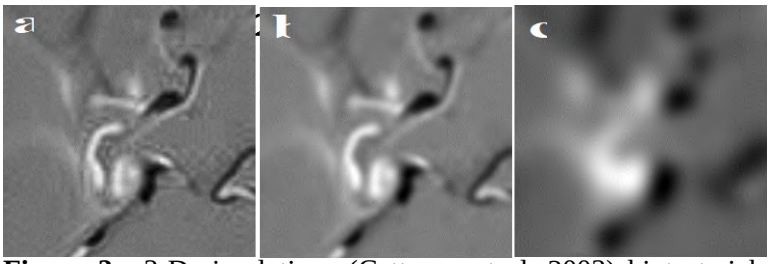


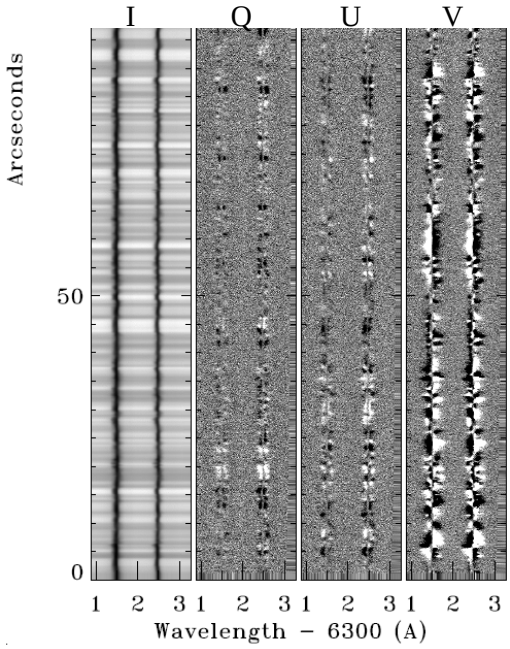
Figure 3. 3-D simulations (Cattaneo et al. 2003) hint at rich, new phenomena just below currently accessible spatial resolution. (**a**) At full (10km grid) resolution, miniature vortices and convective flux tube collapse are visible in intergranular lanes. (**b**) Simulated 4 meter telescope (e.g. ATST) resolves the vortices and horizontal twisted flux tubes. (**c**) Simulated 1 meter telescope fails to resolve the important small-scale dynamics.

High spatial, temporal, and spectral resolution is needed to observe the turbulent vorticity and the diffusion of small-scale magnetic fields to measure these properties and how they evolve with the solar cycle.

Two dynamos, or one?

Although it is commonly hypothesized that at least two separate dynamo mechanisms exist in the Sun (e.g. Leighton 1964, Voegler & Schuessler 2007), the frequency of flux concentrations on the Sun shows a (-1.8) power law versus contained magnetic flux, over nearly six decades in flux, suggesting that the dominant formation mechanism of flux features is scale invariant from global scales to nearly the smallest currently observable scales (Parnell et al. 2009; Figure 2). (If two separate mechanisms were injecting flux on different scales, one would expect a "break" in the power law somewhere between those scales.) Simulations by Stein & Nordlund (2006) also suggest scale invariance on all currently accessible scales.

Figure 4. Stokes polarimetry at 630nm of the quiet Sun shows ubiquitous transient horizontal magnetic fields (Lites et al. 2008) even in quiet Sun. A much larger aperture is needed to resolve the mixed polarity fields and study their evolution on short time scales.



This extremely wide hierarchy of scale poses many questions: How do strong fields and weak fields interact? Does the weak-field component have a large-scale structure? How are they generated? How do they contribute to chromospheric and coronal magnetism and heating? How do they influence determination of elemental abundances?

The curve in Figure 2 shows a break at the small end of the distribution; this may be caused by either proximity of the detection threshold or a change in the physics. There is reason to think that the dynamics of magnetoconvection change below the scale of granulation, which dominates solar convection: granulation imposes an effective diffusion constant on larger flux features, which is absent on smaller scales, allowing convective collapse, vortical mixing, and

related phenomena. 3-D magnetoconvection models (Figure 3) show different behavior on sub-100km scales than on larger ones, hinting that *the energetically dominant spatial scale for the solar dynamo may be immediately below the scales accessible with current (meter-class) solar telescopes.*

To address these questions we must resolve individual magnetic flux concentrations and observe their emergence and dynamics. We must measure the distribution functions of field strength, field direction and "flux tube" sizes and compare these with theoretical models. This also requires observing plasma motions and relating them to the magnetic field.

Horizontal Fields: a "smoking gun" for the small-scale dynamo?

Recent work with GONG+ (Harvey et al. 2007) and Hinode (Lites et al. 2008, Ichikawa et al. 2008) shows a ubiquitous presence of strong, horizontal fields throughout the quiet Sun (Figure 4), even in supergranular interiors (internetwork). The similarity of the density functions and of the occurrence rates in plage and quiet regions of these horizontal fields suggests that indeed a local dynamo process is generating these fields (Ichikawa and Tsuneta 2009). These fields telescope; ~0.2" resolution) observes only part of the distribution, adding further urgency to the need for much finer resolution measurements.

Magnetic and Current Helicity

Helicity plays a fundamental role in evolution and topology and eruption of solar magnetic fields on different spatial and temporal scales (Pevtsov et al., 2008; Pevtsov, 2008, Pevtsov and Longscope, 2009). Helicity is essential for the effective operation of a dynamo, but excessive helicity may suppress the dynamo action. To ensure efficient operation of the dynamo, helicity has to be removed from the dynamo region and transported to the corona (Brandenburg, 2007) where it can result in solar activity. On their way to the surface, magnetic fields can accumulate additional helicity by interacting with convective turbulence. Since coronal fields can store only a limited amount of helicity it needs to be removed from the sun via coronal mass ejections.

Is active region helicity primarily due to interaction with convection or produced by the large-scale dynamo? What is the contribution of surface horizontal flows? The propagation of twist to the corona depends on the emergence of twisted fields, surface flows and Alfven waves and can provide information on the origin of magnetic helicity (Nandy, 2006; Chae, 2007). Detailed modeling combined with high-resolution vector magnetograms of large number of active regions will allow us to separate dynamo and turbulence contribution to the helicity.

We need high spatial resolution vector magnetograms to provide crucial information about evolution of magnetic field twist during flux emergence. To detect helicity "pumping" by Alfven waves will require a computation of electric currents inside individual "flux tubes" -- a task well outside the ability of existing ground and space-based telescopes.

Fine-scale observations of magnetic helicity in the quiet sun are necessary to understand the role, if any, of a small-scale dynamo. The coriolis force affects large scale flows and results in hemisphere-dependent helicity in the global dynamo. In contrast, the small-scale photospheric flows are not strongly affected by the Coriolis force, so helicity from a small-scale dynamo is expected to be hemisphere independent (Pevtsov & Longcope 2009). If the local dynamo simply recycles flux generated by a global dynamo and shredded by convection, the hemispheric helicity rule should be present. Testing the presence of a small-scale dynamo requires observations of vector magnetic fields with spatial resolution significantly better than 0.1 arcsec.

Small-Scale Magnetic Flux Concentrations

The photospheric magnetic field is organized in small fibrils or "flux tubes", which account for a significant fraction of all the magnetic flux threading the photosphere. These structures are

mostly unresolved by current telescopes, yet aggregation of this unresolved flux is the dominant mechanism by which small-scale flux concentrations form (Lamb et al. 2008). Magnetic field lines are the most likely channels for transporting convective energy into the upper atmosphere, which is the source of UV and X-ray radiation from the Sun, which in turn affects the Earth's atmosphere. Detailed observations of these fundamental building blocks of solar magnetic fields are crucial for our understanding of the activity and heating of the outer atmospheres of the sun. Current solar telescopes cannot provide the required spectroscopy and polarimetry at an angular resolution to explore the enigmatic "flux tube" structures (e.g. Figure 3).

Theoretical calculations and numerical simulations suggest that small-scale flux concentrations are formed by convective intensification by the photospheric flow field: flux is continually swept to the intergranular downflow regions where it is concentrated, twisted by local vorticity in the intergranular downflows, and radiatively cooled to form coherent flux bundles. The flux concentrations are highly dynamic, undergoing constant merging, shearing, fragmentation and reformation leading to substructures predicted to be on the order of a few tens of kilometers in scale. Interaction of the flux concentrations with convection buffets, bends, and twists the magnetic field, which heats the upper atmosphere via MHD waves and intertwining of magnetic lines leading to magnetic reconnection and ohmic dissipation of currents.

In spite of the significant progress in adaptive optics (Rimmele, 2004; Langangen et al. 2007), current observations lack the spatial and temporal resolution necessary to verify these theoretical predictions of small-scale flux formation and flow-field interactions (Stein and Nordlund, 2006). Flux-associated bright points are observed to undergo constant splitting and merging on time scales of 10-100 seconds in response to granular flows, similar to the turbulent 3D simulations. However, observations of the formation and decay mechanisms of flux tubes, or of the expected vorticity, are lacking. Such observations require vector magnetic field measurements on spatial scales of a few tens of km combined with a temporal resolution on the order of 10 seconds.

Small-Scale Flux and Solar Atmospheric Heating

Investigating the role of small-scale flux dynamics in powering the Sun's outer atmosphere also requires new levels of spatial and temporal resolution. The TRACE mission (Handy et al. 1999; Schrijver et al. 1999) has shown that coronal fine structure exists on arcsecond scales and very likely below this level as well. However, chromospheric and transition region motions are still not fully resolved in time or space, and the magnetic topology relating the photosphere to the corona remains completely unclear. For example, transition region emissions from hot coronal loop foot points correlate only on the large scales to magnetic field sites in the lower atmosphere (Berger et al. 1999, Nagata et al. 2006). At smaller scales, transition region emissions do not directly overlie Ca II K-line chromospheric bright points or G-band bright points in the photosphere. Thus coronal loops do not appear to be vertically anchored to their associated flux concentrations in the photosphere. Are magnetic field lines entangled by motions as they rise through the atmosphere? Or are the transition region emissions (and perhaps coronal loop heating sources) located at the interface regions of expanding magnetic field lines from the photosphere? Can we correlate flux motions in the photosphere with impulsive events in higher layers of the atmosphere such as spicules and transition region explosive events and ultimately with episodic heating of coronal loops?

MHD waves propagating along magnetic fibrils are a likely candidate for transporting energy to the upper atmosphere (Zagarashvili & Skhirtladze, 2008). In many cases, detailed model predictions, including detailed simulations of observable quantities, such as spectral line profiles or Stokes profiles, are available. The spatial resolution in these models is of the order of a few tens of kilometers and similar resolution data are required to test them.

Small-scale magnetic flux also plays a large role in the total irradiance budget of the Sun. An accurate physical model of flux concentrations, however, is still lacking. Fully resolved

measurements of small-scale flux irradiance as a function of disk position are needed in order to complete a realistic physical model of these important elements.

Addressing these issues requires simultaneous vector magnetometry in multiple layers of the solar atmosphere and drives the temporal resolution requirement even lower, perhaps to one second time scales or less. Only a much larger solar telescope than currently available will achieve the resolution and photon flux requirements and will also allow investigations of the role of small scale flux in energy transport to the outer atmosphere, addressing such issues as chromospheric heating or the source of the fast solar wind in open-field regions.

Magnetoconvection and Flux Transport

The simulations predict that cold plasma forms narrow, turbulent downdraft plumes (Stein and Nordlund, 2006). However, there is no observational evidence for these small-scale (significantly smaller than 0.”1) vortex flows within the intergranular lanes. A large aperture telescope is needed to resolve the small-scale structures and dynamics of convection and verify or disprove the predictions of the models.

Vortical plasma motions, such as those that occur in the downdrafts, will stretch and amplify any seed magnetic fields, producing a small-scale dynamo. The downdrafts may also pump the small-scale magnetic flux produced down towards the base of the convection zone. Convective flows may be important in breaking up the magnetic “flux tubes” into smaller elements and in merging individual “flux tubes” into larger structures. They also may shuffle around the atmospheric magnetic loop foot points and launch MHD waves that propagate into the upper atmosphere.

The ability to resolve and track individual “flux tubes” is needed to understand how magnetic field is organized into larger scale patterns such as meso- and supergranular scales. For example, there is a controversy about the relation of magnetic fields and supergranulation. Does the supergranular scale diverging flow advect magnetic field to the network at the supergranule boundaries? Or, is it the magnetic field that produces the observed supergranulation pattern?

Another example of how a magnetoconvective process that occurs on very small scales (<100 km) causes a global phenomena are the solar p-mode oscillations. Numerical simulations provide detailed predictions on how convective energy is converted into acoustic energy, which may contribute to the heating of the lower chromosphere. While the spatial and temporal resolution of current facilities is sufficient to verify the existence of acoustic events, they are unable to study the underlying physical mechanisms and to verify model predictions. Surface cooling producing low entropy fluid is part of the process, but according to model predictions, turbulence below the surface is the more important driver. One of the remaining challenges is to understand quantitatively the origin of the oscillation mode line asymmetries. This is part of the bigger challenge of understanding the dynamical interaction of oscillations with convection.

Ultra high spatial resolution along with excellent temporal resolution is required to verify the exciting results of numerical simulations. Also needed are measurements of spectral line profiles with sufficient spectral resolution so that the information contained in the details of line profiles can be revealed.

The Advanced Technology Solar Telescope, the tool to make it happen

We have demonstrated that measuring and characterizing evolution of the very fine scale solar magnetic field is crucial to myriad topics, from dynamo theory to solar coronal heating and solar variability and activity. The largest current telescopes are limited by diffraction to ~0".12 (~80km) spatial resolution (at λ= 500 nm), and a new, larger aperture telescope is required to make progress. Because so many science questions surrounding the solar dynamo and related phenomena require the highest possible spatial resolution, and because models predict different

qualitative behavior of flux just below currently observable scales, progress on many crucial issues throughout solar physics is essentially impossible without finer spatial resolution measurements than can be obtained today.

The ATST project will enable breakthrough science that has impact throughout solar and heliospheric physics and the physics of plasmas. Many crucial advances and breakthrough science are anticipated in the transition to the ATST, a 4 meter telescope (0".03, or 20km, ultimate resolution). ATST will provide direct measurements of magnetic flux emergence, twist, oscillations, transport and cancellation rates in the quiet sun from magnetograms that have 16 times better resolution (pixel area) or 30 times better sensitivity (or a compromise between the two) than with the current largest solar telescopes. ATST is appropriately sized to address every single science question outlined in this document, revolutionizing basic understanding of why the Sun has a magnetic field and how that field gives rise to the solar chromosphere and corona and the resultant solar activity. Combining data from ATST with that of high-energy space observations will allow us to develop causal models of solar activity.

References

Berger, T.; et al. 1999, *ASPC* 183. 365

Brandenburg, A. 2007, *Highlights of Astr*. 14, 291

Brandenburg, A.; Subramanian, K. 2005, *Phys. Rept*. 417, 1

Cattaneo, F. 1999, *ApJ* 515, L39

Cattaneo, F.; Emonet, T.; Weiss, N. 2003 *ApJ* 588, 1183

Chae, Jongchul, 2007 *Adv. in Space Res*. 39, 1700

DeForest, C.E. et al.,2004 *ApJ* 616, 600.

Handy, B. N. et al. 1999, *Solar Phys* 187, 229

Harvey, J. W., Branston, D.; Henney, C. J.; Keller, C. U. 2007 *ApJ* 659, L177

Ichikawa et al. 2008, *A&A* 481, L25

Ichikawa, R & Tsuneta, S. 2009, *A&A* (in press), *arXiv*:0812.1631v1

Lamb, D. A. et al. 2008, *ApJ* 674, 520

Lamb, D.A. 2008b, PhD.Th., U. Colorado

Langangen, Ǿ., Carlsson, M. & van der Voort, L.R. 2007, *ApJ* 655, 615

Leighton, R. 1964 *ApJ* 140, 1547

Lites, B. et al. 2008, *ApJ* 672, 1237

Metcalf, T.R. et al. 2008, *Solar Phys*. 247: 269–299

Nagata, S.; Bellot Rubio, L.R.; Katsukawa, Y. 2006, *ApJ* 638, 539

Nandy, D. 2006, *JGR* 111, A12, A12SO1

Parnell, C.P. et al. 2009, *ApJ*, subm.

Pevtsov, A.A., Canfield, R.C.; Sakurai, T. & Hagino, M. 2008, *ApJ* 677, 719

Pevtsov, A.A.; Longcope, D.W. 2009, *ASPC* 369, 99

Pevtsov, A.A. 2008, *J. Astrophys. Astr*. 29, 49

Rimmele, T.R. 2004, *ApJ* 604, 906

Schekochihin, A. A. 2004, *ApJ*, 612, 276

Schekochihin, et al. 2005, *ApJ* 625, L115

Schekochihin et al. 2007, *arXiv*:0704.2002 (accepted New J. Phys.)

Schrijver et al. 1997, *ApJ* 487, 424

Schrijver, C. J. et al. 1999, *Solar Phys* 187, 261

Stein, R. & Nordlund, A. 2006, *ApJ* 642, 1246

Stenflo, J.O. et al. 2002: *A&A* 389, 314

Stenflo, J.O. 2010, A&A 517, 37

Title, A. M., Schrijver, C. J. 1998, *ASPC* 154, 345

Trujillo-Bueno, J., Asensio Ramos, A., Shchukina, N. 2006, *ASPC* 358, 269

Weiss, N.O. & Thompson, M.J. 2008 *Space Sci Rev*, (in press)

Zagarashvili, T.V.; Skhirtladze, N. 2008, *ApJ* 683, L91

A NASA-funded CubeSat Program

Larry Kepko, Doug Rowland and Adrian Daw
NASA Goddard Space Flight Center

Scott Schaire, and Joel Simpson
NASA Wallops Flight Facility

Executive Summary

We propose an augmentation of the Suborbital and Special Orbital Projects (SSOPD) program to include CubeSats as an available science and technology platform, to complement the existing sounding rocket and balloon platforms. A relatively modest increase in the SSOPD budget of ~ $12.5M would allow for a vibrant CubeSat program of up to 4 launches pear year, enabling flight demonstrations of new instrument concepts and opportunities for students to develop the specialized scientific, engineering and management skills required to maintain NASA as a preeminent space agency. The program could be managed at Wallops Flight Facility (WFF), paralleling the highly successful suborbital program, and could leverage the existing WFF contract with NSF to provide technical support for the NSF CubeSat program.

Motivation

For decades the NASA sounding rocket and balloon programs have, and continue to be, indispensable platforms for developing and nurturing the next generation of scientists and engineers, for testing and validating new technologies and instrumentation, and for offering rapid access to space for cutting-edge science experiments. The success and wide community support for the NSF CubeSat Program combined with the increasing number of NASA Explorer proposals that utilize CubeSats demonstrates the maturation of the CubeSat platform. The addition of CubeSats to the NASA portfolio of science mission platforms would enable significant new experiments for all branches of heliospheric and Earth science.

The current manifest of NSF CubeSat missions demonstrates that CubeSats can address a broad array of cutting edge science questions using state-of-the-art instrumentation:

- **Radio Aurora Explorer (RAX)** / Hasan Bahcivan (SRI), James Cutler (University of Michigan) - Understanding the microphysics of plasma instabilities that lead to field-aligned irregularities (FAI) of electron density in the polar lower (80-300 km) ionosphere.
- **Dynamic Ionosphere CubeSat Experiment (DICE)** / Geoff Crowley (ASTRA), Charles Swenson (USU) - Studying storm-time features of ionospheric plasma.
- **Firefly** / Doug Rowland (GSFC), Allan Weatherwax (Siena College) - Determining the source of terrestrial gamma ray flashes.

Colorado Student Space Weather Experiment (CSSWE) / Xinlin Li (University of Colorado, Boulder) - Investigating the relationship between solar energetic particles, flares, coronal mass ejections, and Earth's radiation belts.

Focused Investigations of Relativistic Electron Burst Intensity, Range and Dynamics (FIREBIRD) / Dave Klumpar (Montana State), Harlan Spence (University of Hew Hampshire) - Resolving the spatial scale size and energy dependence of electron microbursts.

CubeSat for Ions, Neutrals, Electrons and MAgnetic fields (CINEMA); TRIO / Bob Lin (UCB), Kyung-Hee University (Korea) - Providing critical space weather measurements of Energetic Neutral Atom (ENA) and in situ suprathermal particles.

Each mission is funded at ~$900k over three years, a relatively small investment that provides for technology and instrument development and training a new generation of scientists and engineers, while simultaneously addressing compelling science questions. Given the continued and growing community support for CubeSats, the maturation of the necessary technology, and the demonstrated benefits to NASA and the heliophysics community, we believe it is time for NASA to include CubeSats as an option in the Suborbital and Special Orbital Projects program. For a small yearly investment NASA would enable:

A) **Instrument & technology test beds**. A major impediment to instrument development is the lack of flight opportunities in which to test and validate new designs. This problem is particularly acute for university researchers who often do not have the access to programs, such as Space Test Program (STP), that the labs do. The lack of inexpensive flight opportunities limits innovation, as new instruments are rarely able to improve the TRL level to the point they could be proposed. For certain types of instrumentation, sounding rockets provide flights of too short duration to test long term instrument or hardware performance, while balloons do not provide the necessary altitudes. CubeSats fill this void, opening a new realm of low-cost investigations, with instrumentation not appropriate for sounding rockets or balloons.

B) **Training the next generation of scientists and engineers**. The sounding rocket and balloon programs have been a primary means by which NASA has trained scientists and engineers for more expensive flight hardware development. Developing a vibrant, well-trained workforce is vital for the long-term health of NASA and the overall space program within the United States. A CubeSat program within the auspices of NASA would expand the explorable regimes of near-Earth space and open up a new realm of instrument development, including solar and terrestrial imaging, thermospheric and ionospheric in situ measurements, and remote sounding. Further, while the sounding rocket and balloon programs allow for often substantial student involvement in designing and building instrumentation, there are comparatively few opportunities to train component or systems engineers within an operational context. Space physics missions often require close collaboration between the scientists, instrument designers, and spacecraft engineers to achieve mission success. Spacecraft charging and magnetic cleanliness are just two examples where choices made in systems designs could adversely affect the ability to produce accurate scientific

measurements. The field would be well served by bringing systems engineers through flights embedded in a space physics program. Because of their small size and low cost, CubeSats offer the ideal platform for student engineers to design, build and fly not just spacecraft subsystems but an entire integrated spacecraft.

C) **Focused science investigations**. As the NSF CubeSat program has demonstrated, the platform is fully capable of a range of science implementations across heliophysics, including ionosphere/thermosphere/mesosphere (ITM), solar and magnetospheric science. The flexibility of the CubeSat platform to be three-axis stabilized or spinning, for example, enables the full-breadth of heliophysics science to be explored. The low cost and complexity allows rapid turn-around from inception to launch, enabling rapid study of focused science questions and also valuable second-tier science that does not justify an Explorer class mission. For example, the NSF Firefly mission is designed to answer very specific questions about energetic electron acceleration over thunderstorms. Despite its small size, its instrumentation is focused to answer new questions that larger satellites, not expressly designed for this mission, could not answer. In a similar fashion to existing suborbital programs, a CubeSat program would allow relatively short concept-to-flight times. Typically, payloads can be flight ready in less than three years from concept and much less than that for existing science.

NASA has typically required closure or significant progress on science questions as a measure of the success of a space flight mission. We believe it would be a mistake to implement a CubeSat program with the same standards of science closure as, for example, the Explorer program. Instead, A NASA sponsored CubeSat program should be implemented to promote flight validation of new advancements, early cutting edges science results, and effective post graduate research, with the acceptance of some risk. A CubeSat experiment that fails early in orbit, for example, has still provided important technical, scientific and engineering development.

Technical Implementation

NASA could implement and manage a CubeSat program in a number of different ways. One option would be to mimic the sounding rocket approach, where much of the technical engineering infrastructure (e.g., communication, attitude control system (ACS), power, mechanical design, and deployment) lies at WFF, leaving the Principal Investigator (PI) responsible for the instrumentation. Another option is to mimic the current NSF implementation, where the PI is largely responsible for all aspects of the CubeSat. Each option carries distinctly different sets of advantages and disadvantages. Relying on WFF for engineering infrastructure frees the PI to focus on their instrument and science, without needing to “reinvent the wheel.” It allows for shared engineering designs and access to flight for small groups that do not have the engineering resources or expertise necessary to design, engineer and build a complete flight system. On the other hand, there is often tremendous value in having the PI and their team design and build a complete system. It allows for the training of systems and subsystem engineers, and can allow for greater flexibility and innovation in spacecraft design.

We describe below the details of implementing a CubeSat program under the auspices of the Suborbital and Special Orbital Projects Directorate (SSOPD) at WFF. This is but one example of an implementation of a NASA-funded CubeSat program, and we encourage the survey panel to examine others, including partnering directly with NSF on the existing program and implementing the NSF-style program within NASA. Regardless of implementation, the program should adhere to the spirit of the CubeSat platform, which is engineering and technical training combined with high science return on a small but powerful platform.

A CubeSat Program office at WFF

The Goddard Space Flight Center Suborbital and Special Orbital Projects Directorate (SSOPD) at WFF could manage the CubeSat Program for a relatively modest investment. SSOPD currently supports the National Science Foundation (NSF) CubeSat Program, manages the Advanced Technologies Program (ATP) for CubeSat and Unmanned Aerial Vehicle (UAV) Applications, the NASA Sounding Rocket (SR) Program and the NASA Balloon Program. SSOPD would manage the overall CubeSat program in a manner similar to the services currently provided in support of the NSF CubeSat Program. This will include coordination with NASA sponsors, interfacing with launch services, offering use of existing test, evaluation, and ground support infrastructure, management of CubeSat bus services, and coordination with selected CubeSat developers. The CubeSat Program would build on SSOPD's long history of successful implementation of NASA's sounding rocket and balloon programs with a culture of responsive schedules, low-cost methods, and miniature payload accommodations. It will also draw from the recent development of CubeSat systems, through the ATP for CubeSat and UAV Applications.

Today, the sounding rocket and balloon programs provide ease of integration for the science instrument developer. The CubeSat Program would be designed in the same fashion to make available to developers CubeSat technologies from the SSOPD ATP, although the proposer would have the option of pursuing their own spacecraft bus solutions. SSOPD has integration and test facilities tailored for small satellites, which will support CubeSat development and testing. These facilities include labs, clean rooms, thermal-vacuum chambers, vibration test stands, Electromagnetic Interference/Electromagnetic Compatibility (EMI/EMC) facilities, antenna pattern test facilities, and spin-balance test facilities. WFF could also provide ground station services.

The current balloon and sounding rocket programs enable the science community to select and prioritize mission funding through a program executive, simplifying the proposal and selection process. Program implementation management is executed through the WFF program offices that provide the integration and launch services. This allows the science PI to focus largely on the delivery of the science instrument should they choose. Typically the PI will have the option to employ extensive use of instrument support "bus" elements that can be provided by the program office. Examples of these are solar pointing for balloons or telemetry and attitude determination for sounding rockets. The CubeSat program could be implemented in very much

the same way. This paradigm allows the proposers the flexibility, if they choose, to focus their limited resources solely on the instrumentation and science. In some cases an integrated spacecraft would be proposed.

As mentioned above, the vision is to provide a small program office at WFF that would be responsive to a program executive at HQ. Flight manifesting would be coordinated through the CubeSat Program Office. Through WFF civil servant and contract engineering, bus and ejector components, integration and test services would be arranged for and provided by the program office as desired from the PI. Current facilities exist, including environmental testing and clean rooms to support these efforts.

Program Cost Estimation

Using the benefits of the existing contractor support organizations and on-site civil servant engineering, the CubeSat program would be able to improve on the current amortized per mission costs of the existing programs. The WFF FY10 costs (not including payloads) are $47M for the sounding rockets, $26.7M for balloons, $3.2M for aircraft, and $19.1M for the range. Each of these suborbital programs is producing approximately 15 missions apiece. Included in these costs are the actual costs of the launch vehicles and the operational costs of the launch effort. By manifesting CubeSat missions on existing host missions, most of these costs are eliminated. CubeSat launch costs are assumed to be $50k/1U, with an additional $100k to handle launch vehicle related documentation, interface, integration, and deployer procurement and testing. Using the numbers above and assuming a program of 12 missions a year once up and running, the per mission program execution cost would be less than $1M.

Modeled after the sounding rocket and balloon programs, there will be a small Program office at WFF to manage the CubeSat Program. The CubeSat Program will work in cooperation with the

Title	Function	Yearly Cost ($K)
Program Manager	Manages overall CubeSat Program	200
Deputy Program Manager/ Technical Manager	Manages Principal Investigator interactions and integration and test support	200
Resource Analyst	Manages Program budget	100
Grants Manager	Supports NASA HQ managed solicitations and awards	100
Safety	Prepares risk analysis and mitigation plans	100
Engineering Support	Supports integration and test and technology development	200
Contract Funding	Electro-mechanical support, integration and test, engineering support, and institutional support	7000
Total		**7900**

existing sounding rocket and balloon programs, using existing contract arrangements and infrastructure. The table above summarizes the yearly program office costs.

The NSF CubeSat program invests ~$900k over three years per experiment, and we assume a similar baseline in this paper. Assuming four new CubeSat awards per year, the program could fund 12 experiments per year (four in their first year, four in their second year, and four in their third) at a yearly cost of $300k/year, plus an additional $250k/launch (four per year). The total yearly science payload cost is then $4.6M. Including the WFF support, the required NASA investment would be $12.5M/year, a relatively modest cost for such a high return-on-investment program.

Summary

With a relatively modest yearly investment of $12.5M NASA could expand the Suborbital and Special Orbital Projects Office to fund 12 simultaneous CubeSat investigations and four launches per year. The program could build upon the highly successful sounding rocket and balloon programs, and leverage existing resources for a high return on investment. We encourage the panel to examine other implementations of a NASA CubeSat program as well, such as partnering directly with the existing NSF program or implementing an NSF-style CubeSat program within NASA. Regardless of how a program is implemented, the benefits for rapid, focused science return, training the next generation of scientists and engineers, and advancing small technologies and instrumentation would be enormous, and are needed to maintain the NASA workforce. The addition of a CubeSat platform alongside the sounding rocket and balloon platforms enables a much larger range of science questions to be addressed, opening up a new area of space exploration to the next generation. Further, the NSF program has proven that the scientific opportunities enabled by CubeSats are in high demand, attracting proposers from the smallest universities to the largest labs. The CubeSat platform is also suited for Earth observing and small astrophysics missions, providing even greater breadth of science return and possible additional funding from other divisions. Finally, we note that each platform – sounding rockets, balloons and CubeSats – provide clearly differentiated science platforms, and including CubeSats in the SSOPD at the expense of funding the existing sounding rocket and balloon programs would be detrimental to the health and vibrancy of the community.

Breakthrough Toward Understanding The Solar Wind Origin

A White Paper submitted to the 2010 Solar and Space Physics Decadal Survey

Yuan-Kuen Ko[1], David H. Brooks[2], Steven R. Cranmer[3], George A. Doschek[1], Antoinette B. Galvin[4], Steve W. Kahler[5], J. Martin Laming[1], Susan T. Lepri[6], Mark Popecki[4], John C. Raymond[3], Yi-Ming Wang[1], Harry P. Warren[1], Peter R. Young[2]

[1]Naval Research Laboratory, [2]George Mason University, [3]Harvard-Smithsonian Center for Astrophysics, [4]University of New Hampshire, [5]Air Force Research Laboratory, [6]University of Michigan

1. Summary

To make a breakthrough in the next decade toward understanding the origin of the solar wind, it is necessary to utilize solar spectroscopic observations and in situ solar wind ion composition measurements, and the synergy between the two is essential. Solar spectroscopic observations have revealed many important physical properties in various structures at the Sun from the chromosphere to the corona, through their powerful diagnostic capabilities in electron/ion temperature, density, abundance and flows. The solar wind ion composition measured in situ is directly related to these physical properties at the Sun where the solar wind is formed and released into the heliosphere. It is therefore clear that combining the unique diagnostic power of the two provides the most stringent constraints on where and how the solar wind is produced and on models of the solar wind formation. To advance our knowledge of the solar wind origin, it is then important to recognize the need in the next decade to continue supporting the collection and analysis of solar spectroscopic and in situ solar wind ion composition data. And we recommend that future mission designs addressing the solar wind origins need to take into account the co-existence and coordination of these two types of instruments.

2. The Diagnostic Power and the Open Questions for Solar Wind Origin

One of the long-unsolved problems in solar and space physics is the origin of the solar wind: where and how it is produced, and what governs its variability. We know this is rooted in the solar magnetic field whose dynamics and interactions with the solar plasma generate the dynamical structures on the Sun, the heating of the corona and the acceleration of the solar wind. These processes manifest themselves in many observables. In this White Paper, we emphasize that the combined diagnostics from solar spectroscopic observations and in situ solar wind ion composition measurements are essential in leading us to the next breakthroughs toward understanding the origin of the solar wind formation.

2.1 Solar Wind Ion Composition Measurements

The solar wind measured in-situ bears signatures of its formation in the solar atmosphere. Unlike particle speed, density and ion/electron kinetic temperatures that may change during transport in the heliosphere, the solar wind ion charge states are “frozen-in” close to the Sun (generally within 5 solar radii, e.g. Owocki et al. 1983, Burgi & Geiss 1986, Ko et al. 1997), and the elemental abundance under normal coronal/solar wind conditions is unlikely to change after it is released into the heliosphere. Therefore the solar wind ion composition

closely reflects the conditions at where the solar wind is being formed at the Sun. It is well known that the solar wind O^{+7}/O^{+6} ratios are anti-correlated with the solar wind speed implying that the coronal electron temperature in the slow wind source region is higher than that for the fast wind (e.g. Zurbuchen et al. 1999, Wang et al. 2009). It is also well known that the First Ionization Potential (FIP) effect exists in the slow wind and less so in the fast wind (e.g. Meyer 1985, Geiss et al. 1995, von Steiger et al. 2000, Zurbuchen et al. 2002). The similar difference in the electron temperature and FIP effect in solar active region (AR)/quiet Sun (QS) and the coronal hole (CH) structures hints at their association with the slow and fast wind, respectively.

However, the variations in the ion composition indicate a far more complicated situation. Here we show two examples. Figure 1 shows the O^{+7}/O^{+6} freezing-in temperature versus proton speed in the solar wind measured by ACE that can be traced back to the boundary of either a north polar CH (NPCH) or a south polar CH (SPCH). It shows a clear distinction (regardless of the solar wind speed!) between the two sources that the solar wind originating from the NPCH has larger O^{7+}/O^{6+} freezing-in temperature than that from the SPCH. Such distinction was also observed by Ulysses (Ko et al. 1999, Zhang et al. 2002) but with opposite trend that the SPCH has larger O^{7+}/O^{6+} freezing-in temperature during those periods of study (years 1992-1997). This implies that there are intrinsic differences in the processes and/or environment for the solar wind formation at the two polar CHs, and it is likely related to the properties in the two CHs such as magnetic field and size (Zhang et al. 2002).

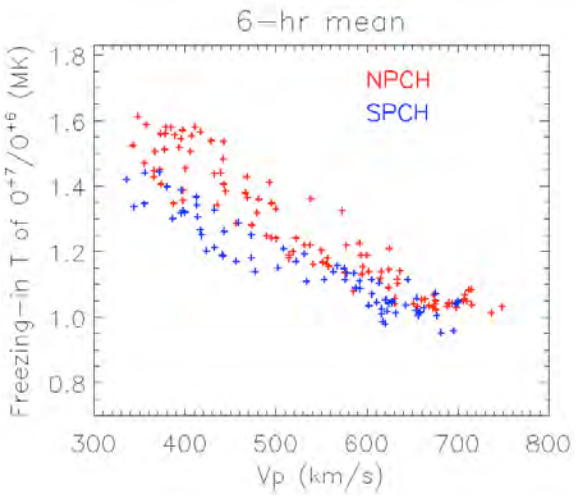


Figure 1 Solar wind freezing-in temperature from the O^{+7}/O^{+6} ratios (6-hr mean) versus proton speed with footpoints at the low-latitude extension of 5 NPCHs and 3 SPCHs. Data are from ACE SWICS and SWEPAM during years 2005-2007 (Ko et al. 2010, in preparation).

Figure 2 shows a time series of solar wind parameters measured by ACE in November 2005. The solar wind from Periods 1&2 are traced back (e.g. by either a PFSS or 3-D MHD model) to the boundary of an equatorial CH with a trailing AR. Period 3 is traced back to the boundary of another equatorial CH with a leading AR. We can see clear distinctions, especially in the solar wind ion composition data, not only between the fast and slow solar wind from the same CH (1&2), but also between the two slow solar wind periods (2&3) from the two CHs. These examples show us that the solar wind ion composition data are powerful diagnostics for the solar wind formation regions, and tell us that:

- We have learned much from past and current space missions about the collective behavior of the solar wind under various solar wind conditions in situ. *To take a leap forward, it is necessary to look into the variations among individual streams, and to search for and understand the underlying cause for such variability. To do this, it is essential to investigate what processes and properties in the associated solar wind source regions contribute to and correlate with these variations.*

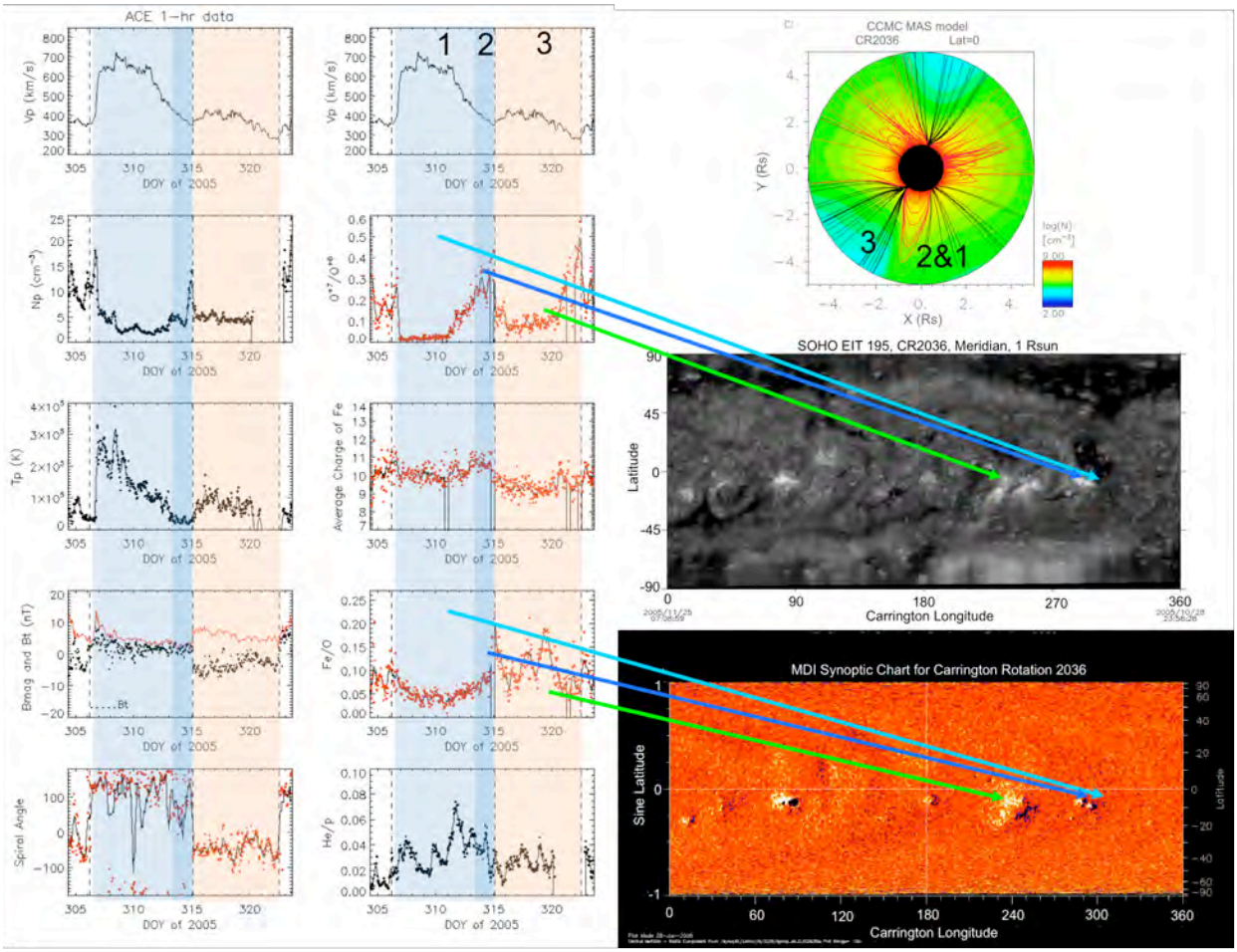


Figure 2 Left panels: ACE measurements of solar wind parameters (1 hour data from SWEPAM, MAG and SWICS. Black curve is from 6-hour smoothing.). Three solar wind periods are marked in different color blocks, and the numbers are written on the top center panel. Period 1 (light-blue) is mostly in the fast solar wind (>500 km/s). Period 2 (dark-blue) and 3 (beige) are in the slow solar wind. The spiral angle indicates that Periods 1&2 and Period 3 are in different interplanetary magnetic field sectors divided by a heliospheric current sheet. Right panels: (Top): Coronal magnetic field configuration on the ecliptic 'cut plane' for this period (Carrington Rotation (CR) 2036) overlaid on the coronal density contours, calculated from the MAS model run at the Community Coordinated Modeling Center (http://ccmc.gsfc.nasa.gov). Black (red) lines represent open (closed) field lines and the associated open field regions corresponding to the 3 periods are marked. Time increases (CR longitude decreases) clockwise starting from right at CR longitude of 360 degrees. (Middle and bottom): SOHO EIT λ195 and MDI magnetogram synoptic maps for CR2036 (EIT map: courtesy of Nathan Rich; MDI map from: http://sun.stanford.edu/synop/) with the approximate footpoint locations for the 3 periods indicated (coordinates are obtained from the PFSS model). We can see that Periods 2 and 3 have similar proton speed and density, but are very different in the proton temperature and ion composition signatures (O^{+7}/O^{+6} ratio, average charge of Fe ions, Fe/O (the FIP bias) and He/p density ratios). This indicates that there are intrinsic differences in the formation of the solar wind between the two CH source regions.

2.2 Solar Spectroscopic Observations

The solar spectroscopic observations, both on disk and off limb, provide rich information about the electron/ion temperature, density, abundance and flows in various structures at the

Sun from the chromosphere to the corona. It is well known that CHs have lower electron density and temperature than the AR/QS structures, and the FIP bias is also lower (Raymond et al. 2001). It is also well known that there are considerable variations within individual types of structures in terms of these properties. In the past decade, there have been increasing efforts to address the problem of solar wind origins from solar spectroscopic observations. Examples are: ion temperature anisotropy and outflows above CHs and in coronal streamers (e.g. Cranmer et al. 1999, Miralles et al. 2001, Strachan et al. 2002, Frazin et al. 2003, Teriaca et al. 2003), outflows at CHs and ARs (e.g. Hassler et al. 1999, Tu et al. 2005, Doschek et al. 2008, Harra et al. 2008, Habbal et al. 2008), abundance variations in solar structures (e.g. Raymond et al. 1997, Ko et al. 2006, Brooks & Warren 2010), relating coronal and solar wind properties from quadrature observations (e.g. Suess et al. 2000, Poletto et al. 2002, Bemporad et al. 2003), and reconnection signatures at CH boundaries (e.g. Madjarska et al. 2004).

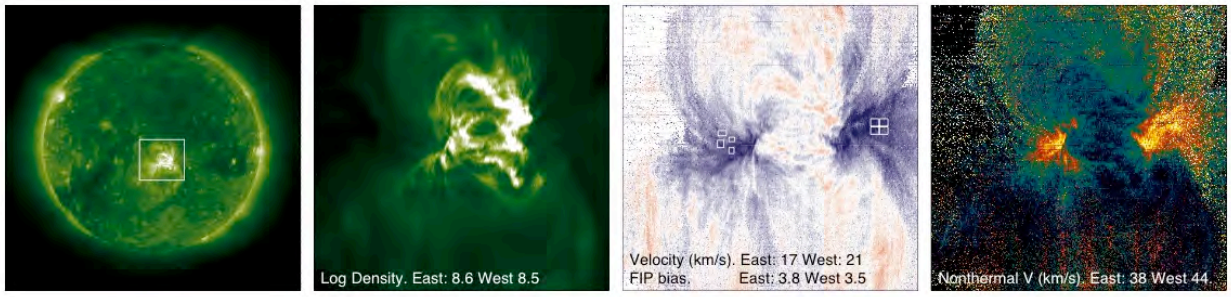


Figure 3 Hinode/EIS observations of AR10978 in the Fe XIII λ202.04 line. From left to right: 1) EIT λ195 image on the day of the EIS observation. The square marks the field-of-view of the EIS raster on AR10978, 2) Intensity image of Fe XIII λ202.04 built from the EIS raster, 3) corresponding Doppler velocity map (blue/red shift indicating up/down flows), and 4) non-thermal velocity map (red/yellow represents higher values than green/blue). Small squares marked on the Doppler velocity map are the upflow regions at both the west and east sides of the AR extracted for detailed study. The average electron density, FIP bias (from the Si/S abundance ratio), upflow velocity, and non-thermal velocity for both sides of the selected upflow regions are written on the panels. Upflows of ~20 km/s and non-thermal broadening of ~40 km/s were seen at both sides of the AR where the 1-2 million degree emission was weak.

Figure 3 shows one recent such attempt from the Hinode/EIS observation of AR10978 on Dec.12, 2007 when the AR was crossing the central meridian (Brooks & Warren 2010). This AR is bracketed by an equatorial CH at its west and the low-latitude extension of the south polar CH at its east. The blue-shift (upflow) regions at the two sides of the AR hint that they can be where the solar wind flows out of the Sun. This is consistent with the idea that these two coronal holes, as indicated from coronal field models (e.g. the PFSS model), are all open field regions that connect to the Earth. The Si/S abundance ratio at the west side of the AR upflow region seems to be also consistent with that in the solar wind measured by ACE three days later. If such one-to-one connection can be established in general by the analysis of many regions, much insight can be gained about how source region properties are reflected in the solar wind properties through the formation process of the solar wind.

- We have learned much about the collective properties of the solar structures and have obtained many clues about how and where the solar wind is formed at the Sun. *Solar spectroscopic observations such as the above, when combined with solar wind ion*

composition measurements, can take us a step further in relating the variations in the solar wind with those in its source regions. It is thus essential to identify and analyze properties in regions at the Sun that are directly associated with the solar wind measured in situ.

2.3 Open Questions for Solar Wind Origin

As we are equipped with much knowledge gained from the wealth of in situ solar wind measurements at one end and solar observations at the other, the time is ripe to connect the dots between the two. And the synergy between the two is essential to help answer some of the most important open questions for solar wind origin that are still facing us.

- *Where is the solar wind formed, and what governs its variability?* What signatures in the coronal hole and/or its boundary with the AR/QS reveal the formation of the, e.g., 700 km/s solar wind versus the formation of the, e.g., 400 km/s solar wind? How do properties, such as temperature and abundance, observed at the Sun lead and relate to properties measured in the solar wind? What in the solar data can reveal/differentiate the solar wind formation processes that result in the solar wind variation in time and space? Can a given theory/model of the solar wind formation make correct predictions for the variations at both the source regions at the Sun and the *associated* solar wind stream?[1]
- *How is the solar wind heated and accelerated* from various processes such as waves and reconnection? What heating and acceleration mechanisms at the Sun can produce the temperature, density, bulk flows, non-thermal velocity profiles, and abundances that are consistent with observations both at the solar source and in the associated solar wind? How does the solar magnetic field manifest itself in the variations of these observables?

3. Next Steps Toward Understanding The Solar Wind Origin

In order to make breakthroughs in the next decade toward understanding the origin of the solar wind formation, we should recognize that solar spectroscopic and in situ solar wind ion composition data are one essential component of *key measurements* in solar and heliophysics missions.

- It is important to continue supporting the collection and analysis of solar spectroscopic and in situ solar wind ion composition data from existing and future space missions.
- Continuing efforts should be encouraged toward the development of high sensitivity and high spectral/time/spatial resolution solar spectroscopic instruments, and high time resolution solar wind heavy ion measurements, that meet the requirements for resolving the fine-scale and dynamic structures relevant for solar wind formation.
- It is important that future solar missions have the ability to measure the composition of both low- and high-FIP elements across a wide range of temperatures and densities. Specific efforts should be taken to identify emission lines of ions that can match the in-situ measurements and provide optimal coordinated diagnostics between the spectroscopic and in situ ion composition data.

---

[1] We also recognize the importance of coronal field and solar wind modeling in order to get the best estimates of the magnetic connection between the source regions and the in situ measurements.

- Co-existence, and coordination in location and timing, between solar spectroscopic observations and in situ solar wind heavy ion measurements will achieve the greatest scientific benefit. Examples are: 1) on-disk observations[2] of structures near the Sun-spacecraft line with the solar wind measured in situ[3], and 2) quadrature observation of the solar corona with the in-situ solar wind off the solar limb[4].

References

Bemporad A. et al. 2003, ApJ, 593, 1146
Brooks, D. H., & Warren, H. P. 2010, submitted to ApJ
Burgi, A., & Geiss, J. 1986, Sol. Phys., 103, 347
Cranmer, S.P. et al. 1999, ApJ, 511, 481
Doschek, G. A. et al. 2008, ApJ, 686, 1362
Frazin, R. A., Cranmer, S. P., & Kohl, J. L. 2003, ApJ, 597, 1145
Geiss, J. et al 1995, Science, 268, 1033
Habbal, S. R., Scholl, I. F., & McIntosh, S. W. 2008, ApJ, 683, L75
Harra, L., K. et al. 2008, ApJ, 676, L147
Hassler, D. M. et al. 1999, Science, 283, 810
Ko, Y.-K., Fisk, L. A., Geiss, J., Gloeckler, G., & Guhathakurta, M. 1997, Sol. Phys., 171, 345
Ko, Y.-K., Gloeckler, G., Cohen, C. M. S., & Galvin, A. B. 1999, JGR, 104, 17005
Ko, Y.-K., Raymond, J. C. , Zurbuchen, T. H., Riley, P., Raines, J. M. & Strachan, L. 2006, ApJ, 646, 1275
Madjarska, M. S., Doyle, J. G., & van Driel-Gesztelyi, L. 2004, ApJ, 603, L57
Meyer, J.-P. 1985, ApJS, 57, 173
Miralles, M. P., Cranmer, S. R., Panasyuk, A. V., Romoli, M., & Kohl, J. L. 2001, ApJ, 549, L257
Owocki, S. P., Holzer, T. E., & Hundhausen, A J., 1983, ApJ, 275, 354
Poletto, G. et al. 2002, JGR, 107, 1300
Raymond, J. C. et al. 1997, Sol. Phys., 175, 645
Raymond, J. C. et al. 2001, in Joint SOHO/ACE workshop "Solar and Galactic Composition". Edited by Robert F. Wimmer-Schweingruber. Publisher: American Institute of Physics Conference proceedings vol. 598 location: Bern, Switzerland, March 6 - 9, 2001, p49
Strachan, L., Suleiman, R., Panasyuk, A. V., Biesecker, D. A., & Kohl, J. L 2002, ApJ, 571, 1008
Suess, S. T., Poletto, G., Romoli, M., Neugebauer, M., Goldstein, B. E., & Simnett, G. 2000, JGR, 105, 25033
Teriaca, L., Poletto, G., Romoli, M., & Biesecker, D. A. 2003, ApJ, 588, 566
Tu, C.-Y. et al. 2005, Science, 308, 519
von Steiger, R. et al. 2000, JGR, 105, 27217
Wang, Y.-M., Ko, Y.-K., & Grappin, R. 2009, ApJ, 691, 760

---

[2] such as Hinode, the proposed Solar-C mission (Doschek et al.), the Solar Spectroscopy Explorer Mission (SSE) (Bookbinder et al.), and 'Fine scale Advanced Coronal and Transition region Spectrograph (FACTS) Mission' (Korendyke et al.). The latter three are white papers submitted to the Decadal Survey.

[3] such as ACE, STEREO, Solar Orbitor, and the proposed Solar Sentinels Mission (Szabo et al., a white paper submitted to the Decadal Survey). Note that there is no future solar wind heavy ion sensors planned at the Sun-Earth line (e.g. L1) post ACE.

[4] such as the proposed 'Mission to the Sun-Earth L5 Lagrangian Point: An Optimal Platform for Heliophysics & Space Weather Research' (Vourlidas et al.), and 'Ultraviolet Coronagraph Spectroscopy: A Key Capability for Understanding the Physics of Solar Wind Acceleration' (Cranmer et al.). Both are white papers submitted to the Decadal Survey.

Zhang, J., Woch, J., Solanki, S. K., & von Steiger, R. 2002, GRL, 29, 1236
Zurbuchen, T. H., Hefti, S., Fisk, L. A., Gloeckler, G., & von Steiger, R. 1999, Space Sci. Rev., 87, 353
Zurbuchen, T. H., Fisk, L. A., Gloeckler, G., & von Steiger, R. 2002, GRL, 29, 1352

Systematic Science for Future Missions

A White Paper submitted to the 2010 Solar and Space Physics Decadal Survey

Yuan-Kuen Ko and George A. Doschek, Naval Research Laboratory

This is to bring to the attention of the steering committee and panels that we should carefully formulate a more systematic approach for future solar and heliophysics missions. Past solar and heliosphysics missions have gone a long way of providing valuable data. And our understanding of the Sun and the heliosphere has been greatly advanced as a result. Future missions are bound to continue this legacy and will aim for further and innovative breakthroughs. As our knowledge continues to accumulate, there are some alarming signs that have/will have negative impacts on the optimal science that can be achieved. Examples are: the lack of a solar wind magnetometer on SOHO, the lack of a solar magnetic field measurement on STEREO, and the non-selection of a solar wind ion composition instrument on Solar Probe Plus. It is understood that resource constraints of budget, power, mass, telemetry, and probably many other factors impacted the selection of the payload. On the other hand, some types of instruments (e.g. spectrometer) would always need a larger share of the resources but constitute an absolutely necessary component for tackling the outstanding and challenging science questions that are facing us. Now comes the opportunity to carefully examine ways and strategy for designing and defining future fleets of solar and heliophysics missions that can achieve systematic, optimal science. This not only includes identifying a set of key measurements that should not be sacrificed, but also the coordination of the locations and timing of data availability from existing and future missions.

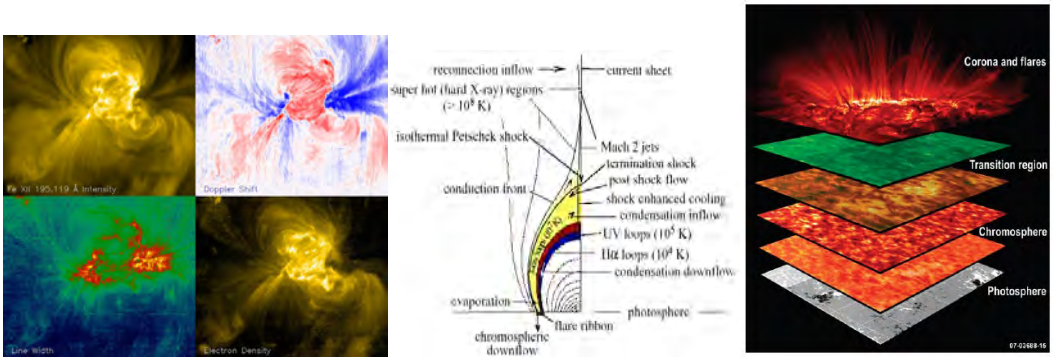
Fine scale Advanced Coronal and Transition Region Spectrometer (FACTS) Mission: An Imaging Spectroscopy Mission to Observe Physical Processes of the Solar Chromosphere, Transition Region, and Corona

By C. Korendyke[1], C. Brown[1], G. Doschek[1], E. Landi[2], P. Young[3], U. Feldman[4], J. Laming[1], J. Mariska[1], J. Morrill[1], J. Seely[1], I. Ugarte[3], A. Vourlidas[1], and H. Warren[1]

[1]Space Science Division, Naval Research Laboratory; [2]University of Michigan; [3]George Mason University; [4]ARTEP

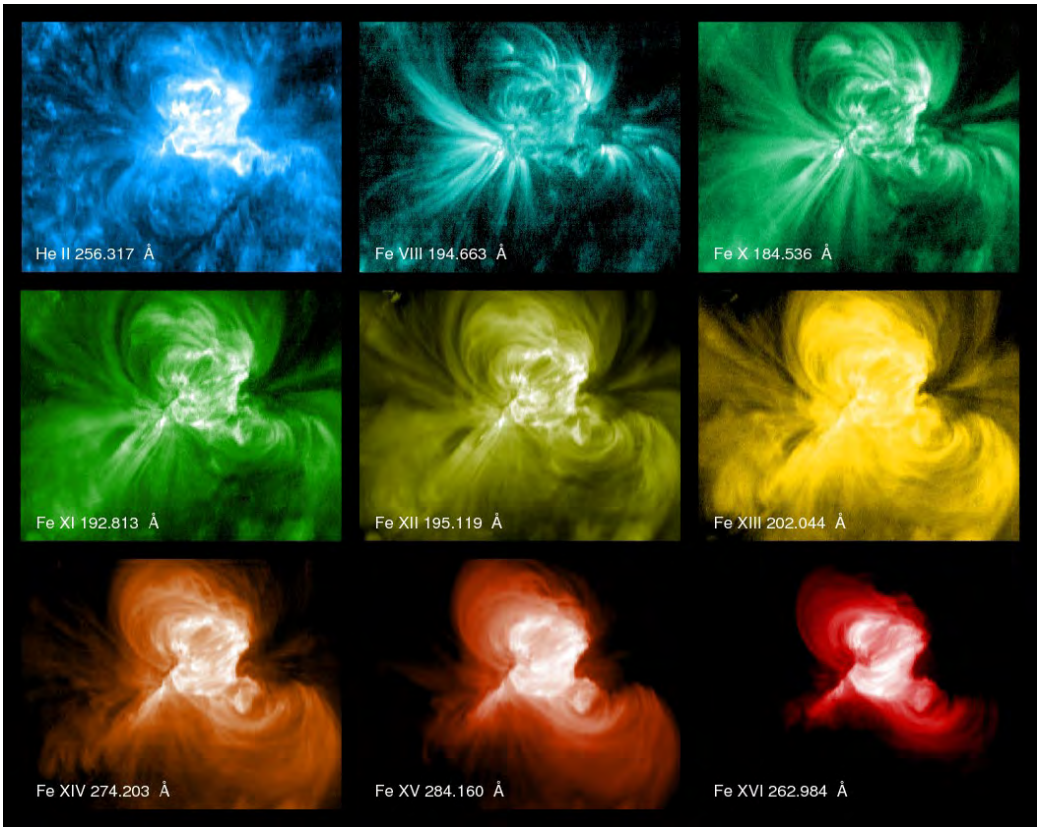
Science Objectives: The science objectives of FACTS are similar to the high resolution Solar-C EUV spectrometer. The FACTS objectives are to determine the dominant physical processes in the solar atmosphere responsible for solar flares and mass ejections, active region loop heating, and general coronal heating. Previous and current observations have shown that these dominant processes, such as magnetic reconnection (e.g., Drake et al. 2006, Nature, 443, 553) are multi-scale in nature. They start as small-scale phenomena and develop into larger-scale phenomena as the energy propagates and is dissipated throughout the solar atmosphere. Filling factors predicts scale sizes from 0.1" in the transition region/chromosphere to 0.3-0.5" in the corona. With both high resolution imaging and spectroscopy, FACTS will directly observe signature of physical processes within the smallest structures of the solar atmosphere. Theoretical and observational evidence directly point to 0.2-0.3" scale structures as being a critical size scale for observing energy release and plasma heating phenomena within the solar atmosphere. Another result of current work is that energy propagates freely over the entire temperature domain of the solar atmosphere. Thus, to observe adequately the flow of energy and dissipation in the atmosphere, an instrument is required with 0.2-0.3" spatial resolution throughout all temperature regions of the upper atmosphere from about 20,000 K to 20 MK. FACTS is designed to have the required properties necessary to determine and characterize the dominant physical processes responsible for the structure, dynamics and evolution of the upper solar atmosphere. These processes are at the root of space weather phenomena.

FACTS observations have the necessary high spatial and spectral resolution, broad but well differentiated temperature coverage, comprehensive spectroscopic diagnostics, and sufficient cadence/context to achieve these objectives. The most relevant RFAs from the NASA Roadmap are: F1: "understand magnetic reconnection … flares, CMEs, …"; F2: "understand processes … that accelerate … particles"; H1: "understand causes…solar activity ... that affects earth"; J2: "Develop prediction capability of … solar activity…".



The left-most figure above shows the power of a spectrometer for measuring the physical conditions in an active region. The upper left image is the intensity of the active region in a spectral line of Fe XII at 195.12 Å formed at 1.4 MK. The lower right image is the electron density map of the active region deduced from spectral line intensity ratios of Fe XII lines. The upper right image is a Doppler map (blue is toward the observer) of the active region determined from the 195 Å Fe XII line, and the lower left image is a turbulence map determined from the Fe XII line width (data from Hinode/EIS). The middle figure is a schematic of the Standard Flare Model which requires diagnostics such as shown in the leftmost figure to test its validity. The right-most figure illustrates the complexity of the solar atmosphere. Energy flow occurs between atmospheric regions of vastly different temperatures and spatial scales. FACTS is designed to observe this entire region with a spatial resolution of 0.2", sufficient for tracking the smallest energy aggregations that appear to be significant for carrying energy in the atmosphere.

The importance of covering all temperatures between about 20,000 K and 20 MK cannot be overemphasized. If an instrument does not cover all the temperature regions, hot plasma can cool, or cool plasma can heat to temperatures outside the range of instrument sensitivity. In this case either the origin or final dissipation of an energy packet cannot be determined. This has been one of the chief stumbling blocks in our ability to determine the physical processes responsible for solar activity. As an example, the figure below shows the different appearance of an active region in different spectral lines formed between 0.1 MK and 3 MK (temperature increases from left to right).



FACTS Spectrograph Instrument Description:

The FACTS instrument design closely parallels the design of the EUV spectrometer for Solar-C. Both of these instruments have a fundamental derivative design based on EIS on Hinode. A portion of the Sun's image is imaged onto a slit or slot by an off-axis parabolic mirror. This mirror must be large to have enough throughput for high-cadence observations. The mirror has multiple coatings in order to select different wavelength regions of the solar spectrum, in order to cover the entire desired temperature domain. The light passing through the slit is diffracted by a multiply-coated grating and focused onto CCDs or ICCDs that cover different wavelength ranges. Below is a possible FACTS instrument summary characteristics.

Table 1: FACTS summary instrument characteristics table

instrument characteristic	**field of view**
spectra plate scale	0.1"/12 micron pixel, λ<2000Å
spectrograph field of view	>200" long slit, >150" raster range
spectrograph wavelength coverage	EUV1 band: 330-370Å (alt. 170-210Å), EUV2 band: 500-1000Å, EUV3 band: 1000-2000Å UV/visible band: 4000-8000Å 1st order 2000-4000Å 2nd order UV/visible band: TBD polarimetry
temperature coverage	photosphere to hot corona (10^7 ° K)
FACTS spectroscopic diagnostics highlights	comprehensive temperature and density line ratio diagnostics

	LOS velocity and turbulence maps abundance measurements UV/visible magnetograms
Spectrograph exposure times	A few seconds to 10s of minutes
UV/VIS broadband filter system	0.2”/12 micron pixel 400” x 400” field of view Ly-α line (121.6nm), Tmin(160nm), Mg II line (280nm), Ca II (393nm), G-band (430.5nm), Hα line (656.3nm), TiO band (705.7nm).

We note that the above table is for a strawman instrument. The wavebands could be chosen differently and still satisfy the science requirements. The final selection of wavebands and mirror and grating coatings would be made during proposal preparation. Continuing with the strawman concept, Table 2 lists details of the FACTS technical design.

Table 2: Subset of scientific/technical characteristics of the FACTS strawman payload.

telescope design parameters	**Baseline or nominal value**
Type	Single paraboloid mirror, 30-40cm dia., mirror has broadband coating with reasonable reflectance in the EUV/UV/Vis spectral range
mirror stabilization mechanism	heritage design with voice coil actuators and integral feedback sensor, <0.03 arcsecond control, 15Hz bandwidth.
slit assembly design parameters	**Baseline or nominal value**
General	>5 slits and slots of variable width (0.1, 0.2, 0.5 and 1 arcsecomds rest – TBD) available for observation, front face with reflective coating, >200” length
UV/VIS broadband filter system	**Baseline or nominal value**
General	reflective design. Projects a high quality image of the slit jaw focal plane onto a CCD camera using selectable bandpasses. A portion of the image is diverted using a beam splitter to feed a dedicated high speed CCD camera used for the focal plane correlation tracker system.
filter bandpass (possible)	Ly-α line (121.6nm), Tmin(160nm), Mg II line (280nm), Ca II (393nm), G-band (430.5nm), Hα line (656.3nm), TiO band (705.7nm).
Spectrograph characteristics	**Baseline or nominal value**
EUV1 bandpass (330-370Å, alt. 170-210Å)	Detector: 4Kx2K CCD, 12micron pixels stray light rejection filter: aluminum on mesh spatial plate scale: 0.1”/pixel spectral plate scale: 0.010Å/pixel, 8.6km/second-pixel@350Å EUV multilayer coating

EUV2 bandpass (500-1000Å)	Detector: 8Kx2K intensified CCD, KBr photocathode material spatial plate scale: 0.14”/pixel spectral plate scale: 0.050Å/pixel, 21.4km/second-pixel @ 700Å
EUV3 bandpass (1000-2000Å)	detector: combination of CCD and ICDD, KBr photocathode material spatial plate scale: ~0.1 and 0.14”/pixel spectral plate scale: ~0.060Å/pixel, 16.4km/second-pixel @1100Å
UV/Visible bandpass (2000-8000Å)	detector: 8Kx1K usable area butted CCDs, spatial plate scale: ~0.2”/pixel spectral plate scale: ~0.030Å/pixel grating assembly includes scan mechanism

In addition to the spectrometer, the FACTS mission implementation includes a high resolution EUV filter imager, to provide context for the spectrometer in situations where the time resolution of the spectrometer is stressed. This imager has been included below in the Mission Implementation Description. The imager will also be capable of 0.1” resolution allowing probing of spatial scales as small as 70km. Spectral channel coverage of the imager will be optimized for the mission science profile.

Mission Implementation Description:

1. *Mission*: single spacecraft mission, 3 axis stabilized, 24 hour solar viewing for most of the year, accommodate 1-5Mbps TM rate.
2. *FACTS spectrograph instrument*: 0.2” goal, two to four channel EUV/FUV/VIS imaging spectrograph.
3. FACTS *EUV telescope*: 0.1” spatial resolution, four imaging channels.
4. *Estimated payload resources required*: 200-250kg, 120W, Taurus size fairing, instrument TRL (6-9 at component level)
5. *Measurement Strategy*: rapid, high spatial resolution, imaging spectra and high resolution imagery.

Lastly, the FACTS instrumentation can also form part of a larger mission to explore the physics of the solar atmosphere at small spatial scales for example the RAM mission or Solar C mission. The expected ROM costs for the FACTS mission implemented with a high resolution EUV imager instrument were estimated and given in Table 3.

Item	ROM cost (FY11 $M)
FACTS spectrograph instrument	$80-120M
High resolution coronal imager	$40-60M
Spacecraft	$80-100M
30% margin	$60-80M
Launch	$180M (varies with NSLS pricing)
MO and DA	$20-50M (sun synchronous, GEO orbit)
Total cost	$460-590M

A Science White Paper submitted in response to the 2010 Decadal Survey Call

Solar Dynamo

A.G. Kosovichev (Stanford), R. Arlt (AIP), A. Bonanno (INAF), A. Brandenburg (NORDITA),
A.S. Brun (CEA), F. Busse (Univ. Bayreuth), M. Dikpati (HAO), H. Hara (NAOJ),
F. Hill (NSO), P.A. Gilman (HAO), I.N. Kitiashvili (Stanford), N. Kleeorin (Ben Gurion Univ.),
R. Komm (NSO), N.N. Mansour (NASA/Ames), M. Miesch (HAO),
A. Nordlund (Copenhagen Univ.), A. A. Pevtsov (NSO), V.V. Pipin (UCLA),
M. Rempel (HAO), I. Rogachevskii (Ben Gurion Univ.), G. Ruediger (AIP),
R.F. Stein (Michigan State Univ.), T.Sekii (NAOJ), R.D. Simitev (Univ. Glasgow),
J.O. Stenflo (ETH), M.J. Thompson (HAO), R.K. Ulrich (UCLA), N. Yokoi (Univ. Tokyo),
T. Yokoyama (Univ. Tokyo), A.A. Wray(NASA/Ames), J. Zhao (Stanford)

Contact information: A.G. Kosovichev, Physics/Astrophysics, Room 128, 452 Lomita Mall, Stanford University, Stanford, CA 94305-4085, tel 650-723-7667, e-mail: sasha@sun.stanford.edu

Solar Dynamo

1 Overview

It is generally accepted that cosmic magnetic fields are produced by dynamo processes operating on various scales. In these processes the magnetic field is maintained against Ohmic dissipation by turbulent motions, and despite the turbulent nature it shows remarkable self-organization properties, forming sunspots and starspots, magnetic loop structures, magnetic spiral arms etc. These phenomena show striking similarity suggesting that the basic physical mechanisms are essentially the same.

The solar magnetism is studied in great details, from global fields of the interior by helioseismology, to the smallest resolved and even unresolved scales by new large ground-based telescopes (SST, GREGOR, BBSO NST) and from space (SOHO, RHESSI, STEREO, Hinode, and SDO). In addition, significant progress has been made in realistic numerical MHD simulations. **The recent progress in observations and modeling provides a good basis for substantially advancing our understanding of the solar dynamo and the formation of active regions. This will have tremendous impact in many fields of astrophysics, heliophysics and developing physics-based forecasts of the solar magnetic activity and sunspot cycles.** However, this new science opportunity requires focused coordinated efforts in observations, modeling and theory. This opportunity can be realized with relatively modest investments, mostly, for supporting the projects that already exist or are under development.

A key element of this opportunity is the understanding of interlinks between small-scale turbulent properties of magnetized plasma and large-scale dynamo processes. It has been long assumed that the turbulent properties, such as turbulent diffusivity and helicity, define the large-scale behavior of the magnetized plasma with some simple back-reaction, but recent plasma experiments and theoretical studies showed that the large-scale flows and structures may significantly alter the turbulence, and that this may cause large-scale organization in plasma[1,2].

In heliophysics, the opportunity of advancing our knowledge of the turbulent dynamo and magnetic self-organization phenomena is based on the following key components:

- **ground-based synoptic observations at the Mount Wilson Observatory (MWO), Wilcox Solar Observatory (WSO) and the National Solar Observatory (NSO/SOLIS) for studying the long-term behavior of large-scale magnetic fields and flows, including the differential rotation, torsional oscillations, and the meridional circulation**;
- **operations and data analysis of the Solar Dynamics Observatory (SDO) mission for the whole solar cycle**;
- **continued long-term helioseismology observations from GONG, and upgrade of the GONG network for high-resolution multi-wavelength observations of solar oscillations and magnetic fields**;
- **out-of-ecliptic space missions for studying the magnetism and dynamics of the polar regions, such as Solar Orbiter, Solar-C (Plan A), and Solar Polar Imager (SPI), which are currently under discussions and initial investigations**;
- **Advanced Technology Solar Telescope (ATST) for studying small-scale turbulent processes with the highest possible resolution and also unresolved properties by using advanced spectro-polarimetry methods**;
- **realistic MHD numerical simulations of the solar turbulence and dynamo at large National Supercomputer Centers**.

2 Recent progress and current status

During the last decade helioseismology observations from SOHO and GONG provided us with a detailed knowledge of the internal differential rotation of the Sun and led to the discovery of two rotational shear zones, at the top and bottom of the convection zone[3]. These zones are believed to be important for shaping the solar dynamo[4]. Helioseismology techniques are rapidly progressing. They have provided initial results on subsurface meridional and zonal flows and formation of sunspots and active regions[5–8]. New data from the Solar Dynamics Observatory allow us to investigate the dynamics of the convection zone and magnetic flux emergence.

The high-resolution observations from the Hinode space mission, Swedish Solar Telescope on La Palma and other large telescopes have provided important details about the interaction between plasma and magnetic fields, and revealed small-scale magnetic structurization[9,10]. New spectro-polarimetry methods based on measurements of scattered polarization provided evidence of "hidden" magnetism of unresolved scales[11]. These studies require large-mirror telescopes, such as the ATST.

Hinode observations of the Sun's polar region made during the periods of high inclination of the solar equator to the ecliptic (7°) have shown that the polar magnetic field consists of small kilogauss-field structures[12]. This structurization is undoubtedly important for understanding the polar field reversals during the solar cycle. In addition, local helioseismology analysis of high-latitude regions revealed large-scale organization of supergranulation ('alignment' patterns)[13], which may be critical for understanding the dynamics of the polar regions and flux transport mechanisms.

Rapidly increasing supercomputer facilities have led to significant advances in numerical simulations of the solar dynamo and magnetic structures, including mean-field dynamo models, MHD turbulence, small-scale dynamo, and the sunspot structure[14–22]. Recent realistic numerical simulations of the global dynamics and dynamo of the Sun are extremely promising, but require further advances in modeling of subgrid MHD turbulence, which controls transport of heat, momentum and magnetic flux. This can be achieved by implementing modern turbulence models and comparing with high-resolution observations.

The ground-based long-term synoptic observations of surface magnetic fields and plasma flows play extremely important role for understanding the cyclic behavior of the solar dynamo. Detailed synoptic data from the Mount Wilson, Wilcox and National Solar Observatories obtained for the last three solar cycles have uncovered a fascinating picture of well-organized processes of magnetic flux emergence, transport and polarity reversal, in the apparently chaotic behavior of magnetic field[20]. These synoptic observations definitely play a major role in our quest for understanding the dynamo mechanisms, and must be thoroughly supported.

Thus, the recent advances in observations and theory provide us a solid basis and a new opportunity to substantially advance our understanding of the dynamo processes and magnetic self-organization. From the observational point of view, it is important to use advances of helioseismology to determine the links between the interior dynamics and the surface and coronal phenomena: emergence and evolution of active regions, magnetic flux transport, polar field reversals and magnetic flux dissipation and escape. From the theoretical point, it is necessary to link dynamo models with observed dynamical phenomena, directly related to the magnetic field generation and evolution, such as variations of the differential rotation rate, zonal and meridional flows.

3 Basic physical questions

How is magnetic field generated and circulated?

In the Sun and astrophysical objects, dynamo action can exist in plasmas with a seed magnetic field and flow fields. However, sufficient conditions for dynamos are not well-determined. For solar and stellar physics it is particularly important that dynamo processes can result in a cyclic behavior. Mean-field MHD theories of solar and stellar dynamos predict the cyclic behavior, which resembles the observed properties such as the butterfly diagram for sunspot formation zone and polar field polarity reversals. However, our understanding of the underlying physical processes is still schematic.

The standard Parker's $\alpha - \Omega$ mechanism[25] has been applied to the Sun and a wide range of other astrophysical objects, stars, galaxies, interstellar medium. The success of this beautiful theory hinges on the turbulent properties, magnetic and kinetic helicities and diffusivity. However, direct numerical simulations, developed during the past decade, revealed significant limitations, such as the catastrophic helicity quenching, which severely restricts the magnetic field growth[16]. The potential solution is in studying the helicity balance including large-scale circulations and helicity loss through coronal mass ejections[26]. This study requires detailed observations of the solar interior, magnetic field and coronal dynamics over the whole solar cycle. Such observations are planned now for SDO, but only for 5 years, a portion of the solar cycle. It is extremely important to extend the operation and data analysis support for this mission. It is particularly important to understand the mechanism of the cyclic polar field reversals revealed in synoptic ground-based observations during the past three sunspot cycles. Because the Sun's axis is almost perpendicular to the ecliptic our knowledge of the polar magnetic field structure and dynamics is very poor. The best opportunity to gain this knowledge can be provided by the out-of-ecliptic solar missions, which are currently under discussion and evaluation[27].

How are heat, momentum and magnetic flux transported in turbulent magnetized plasma?

The main complication in our understanding of many observed phenomena is due to the fact that the solar and astrophysical plasmas are in a highly turbulence state. Turbulence is the major factor determining the heat, momentum and magnetic flux transport. It has been well-established that assessments of the turbulent transport in terms of the traditional (and still widely used) mixing-length theories and order-of-magnitude estimates are insufficient. There is a strong complicated interplay between the turbulence and large-scale plasma flows and magnetic structures. The turbulent transport determines large-scale temporal-spatial characteristics, and, in turn, the turbulent properties strongly depend on the character of large-scale structures and flows. High-resolution solar observations, such as planned for the ATST[23], and realistic MHD modeling[24] can significantly improve our understanding and description of these processes. This requires the development of advanced spectro-polarimetric techniques for observations and of sub-grid MHD turbulence models for modeling. It is nowadays believed that the essential ingredient to understand the physics of the solar dynamo is the meridional circulation, although the strength of the return flow at the base of the convection zone is not determined at the moment. However, by means of long-term helioseismic observations from GONG and SDO it will eventually be possible to put constraints on the return flow.

How does magnetic field act on plasma, and form active regions and sunspots?

One of the puzzling features of solar magnetism is the multi-scale spatial and temporal behavior. High-resolution observations reveal that the magnetic field on the Sun's surface is very structured and consists of small, rapidly evolving magnetic elements, the ultimate scale of which is still unre-

solved. These elements form active regions and sunspots, which seem to emerge randomly. At the same time, the solar magnetic fields show a remarkable degree of organization on the global scale, displaying the 'butterfly' diagram and polarity reversals quite regularly with an 11-year cycle. The recent breakthrough in understanding self-organization processes in plasma experiments offers a clue that the mechanism of self-organization is in the interaction between large-scale flows and turbulence. For instance, shearing flows can stretch turbulent eddies and reduce turbulent dissipation and transport in particular directions resulting in localization of flows and magnetic structures. This discovery encourages further development and search for similar mechanisms in astrophysical plasmas where organized self-maintained structures of various scales are common.

4 Key projects for solving the solar dynamo problem

Synoptic Magnetographic Observations: investigating long-term behavior of the solar dynamo

Most of our current knowledge about the solar dynamo has come from the synoptic observations of magnetic fields and Doppler velocities on the solar surface carried out during the past three cycles. These observations have revealed the general picture of the Sun's magnetic structure and dynamics, spatial-temporal magnetic organization in the form of long-living complexes of activity, migrating zonal shearing flows associated with flux emergence zones. These observations also provide data necessary for predicting future sunspot cycles. This global picture needs to be integrated with the detailed helioseismic data anticipated from SDO and numerical models. Thus, it is extremely important to continue these observations. In the current era of space missions and fast publications, these long-term observational projects are often undervaluated. However, the characteristic period of the Sun's magnetic cycles is 22 years. Thus, for solving the dynamo problem it is very important to maintain the current synoptic observations, at least, at three observatories, MWO, WSO and NSO/SOLIS, keeping their unique features and calibrations. Otherwise, we will loose track of the solar magnetic cycles.

The main value of synoptic observations is in the long-term continuous data sets. The availability of synoptic data enables testing of new models and theories over past solar cycles. For example, it has been recently suggested that solar polar fields may serve as a forecasting tool in predicting the strength of the following solar cycle. This revelation would not be possible without synoptic observations of magnetic fields over several solar cycles from the 150-Foot Solar Tower at Mt. Wilson Observatory, the Kitt Peak Vacuum Telescope, and the Wilcox Solar Observatory.

SDO and GONG: imaging magnetic structures and dynamics inside the Sun

Solar Dynamics Observatory will provide, for the first time, continuous monitoring of the solar interior, surface magnetic fields and coronal structures with a resolution sufficient to investigate the emergence of magnetic flux, formation and decay of sunspots and active regions, their coronal activity, and most importantly, links between these local magnetic structures and global dynamo processes. This mission will provide tremendous amount of data (2TB a day) and has a huge discovery potential. Unfortunately, the budget for this mission is planned only for 3-5 years of operation. For the data analysis only two groups with the annual budget of $700k are organized. This is grossly insufficient. This mission needs more attention and support. In addition, it is very important to modernize the helioseismology network GONG and prepare it for continuing the monitoring of the physical conditions inside the Sun after the end of the SDO mission.

Out-of-Ecliptic Solar Mission: studying the global field polarity reversals

Two out-of-ecliptic solar missions are being actively studied and have a chance of being developed and launched during the next decade. These are the Solar Orbiter and Solar-C (Plan A) missions. They

will provide an important insight into the structure of polar magnetic fields and the mechanism of polarity reversals. The polar fields are largely unipolar, and it was believed that the polarity reversals result from a diffusion or circulation process of magnetic flux transport from the low-latitude zone where the flux emergences. However, recent high-resolution observations from Hinode showed that the polar fields are highly structured, and, while the mean polar field is only a few Gauss, the field in these elements is very strong [12]. This discovery challenges the flux transport models, and opens a new opportunity of studying the relationships between the structurization and global field reversals. The out-of-ecliptic missions will also improve our knowledge of the large-scale convection, differential rotation and meridional circulation, which are the key ingredients of dynamo models.

ATST: studying properties of magnetized turbulence and building blocks of solar magnetism

We expect that most of the progress in observational studies of turbulent properties and formation of magnetic structures in the solar plasma will be made in high-resolution observations with large ground-based telescope, such as ATST. These observations require sophisticated instrumentation for high-spectral resolution and multi-wavelength spectro-polarimetry, and also sufficient photon flux for the scattered polarization analysis, necessary for studying unresolved magnetic elements. This type of measurements will make a breakthrough in our understanding of the nature of the solar magnetism.

Realistic Numerical Simulations: modeling dynamo and magnetic self-organization

Our physical understanding of the complex turbulent MHD phenomena comes mostly from numerical simulations. The modern simulation codes include all essential physics from first principles and are very powerful and can run efficiently on parallel supercomputers. Of course, small turbulent dissipation scales are not resolved. Their modeling requires the development of subgrid turbulence models, particularly, for the MHD case. The subgrid models are usually tested by using simplified simulations of fully resolved turbulence. However, there is an opportunity of developing a synergy between the modeling and the high-resolution ATST observations, and actually test and calibrate the turbulence models for realistic astrophysical conditions (as is done in engineering fluid mechanics through the use of laboratory experiments). During the next decade the numerical simulations will focus on modeling the entire dynamo process without mean-field assumptions [15], and together with the ATST, SDO and synoptic observations and analyses they will substantially advance the understanding of the solar dynamo problem. The simulations are also capable of providing a more detailed knowledge of the formation of self-organized magnetic structures in the solar plasma. The recent initial attempts to model the entire filamentary sunspot structure are quite encouraging [21].

5 Summary

The recent progress in helioseismic and spectro-polarimetric observations and realistic MHD simulations provides an opportunity of substantially advancing our understanding of the physical mechanisms of magnetic field generation and formation of magnetic structures on the Sun. This opportunity is based on the following key components:

1. ground-based synoptic observations of solar magnetic fields and flows from MWO, WSO and SOLIS;

2. long-term helioseismology observations of the solar interior from SDO and GONG;

3. out-of-ecliptic observations of the polar magnetic fields and dynamics from Solar-C/Plan-A and Solar Orbiter;

4. observations of small-scale MHD turbulence from ATST;

5. large-scale realistic MHD simulations.

A significant breakthrough can be made with help of a targeted funding to support coordinated interdisciplinary groups of observers, theorists, and modelers, working together on solar dynamo as a single complex problem.

References

1. Yoshizawa, A., Itoh, S.-I., Itoh, K., Yokoi, N. 2004. TOPICAL REVIEW: Dynamos and MHD theory of turbulence suppression. Plasma Physics and Controlled Fusion 46, 25.
2. Itoh, S.-I., Itoh, K., Shibahashi, H., Diamond, P. H., Yoshizawa, A. 2007. The Possible Magnetic Torus in Stellar Interior. Plasma Physics and Controlled Fusion 49, 809-824.
3. Schou, J., and 23 colleagues 1998. Helioseismic Studies of Differential Rotation in the Solar Envelope by the Solar Oscillations Investigation Using the Michelson Doppler Imager. Astrophysical Journal 505, 390-417.
4. Brandenburg, A. 2005. The Case for a Distributed Solar Dynamo Shaped by Near-Surface Shear. Astrophysical Journal 625, 539-547.
5. Haber, D. A., Hindman, B. W., Toomre, et al. 2002. Evolving Submerged Meridional Circulation Cells within the Upper Convection Zone Revealed by Ring-Diagram Analysis. Astrophysical Journal 570, 855-864.
6. Zhao, J., Kosovichev, A. G. 2004. Torsional Oscillation, Meridional Flows, and Vorticity Inferred in the Upper Convection Zone of the Sun by Time-Distance Helioseismology. Astrophysical Journal 603, 776-784.
7. Kosovichev, A. G., Duvall, T. L. 2006. Active Region Dynamics. Space Science Reviews 124, 1-4.
8. Komm, R., Morita, S., Howe, R., Hill, F. 2008. Emerging Active Regions Studied with Ring-Diagram Analysis. Astrophysical Journal 672, 1254-1265.
9. Scharmer, G. B., Langhans, K., Kiselman, D., Löfdahl, M. G. 2007. Recent High Resolution Observations and Interpretations of Sunspot Fine Structure. New Solar Physics with Solar-B Mission 369, 71.
10. Ichimoto, K., and 10 colleagues 2007. Twisting Motions of Sunspot Penumbral Filaments. Science 318, 1597.
11. Stenflo, J. O. 2004. Solar Physics: Hidden Magnetism. Nature 430, 304-305.
12. Tsuneta, S., and 13 colleagues 2008. The Magnetic Landscape of the Sun's Polar Region. Astrophysical Journal 688, 1374-1381.
13. Nagashima, K., Zhao, J., Kosovichev, A. G., Sekii, T. 2010. Detection of Supergranulation Alignment in Polar Regions of the Sun by Helioseismology. ArXiv e-prints arXiv:1011.1025.
14. Cattaneo, F., Emonet, T., Weiss, N. 2003. On the Interaction between Convection and Magnetic Fields. Astrophysical Journal 588, 1183-1198.
15. Brun, A. S., Miesch, M. S., Toomre, J. 2004. Global-Scale Turbulent Convection and Magnetic Dynamo Action in the Solar Envelope. Astrophysical Journal 614, 1073-1098.
16. Brandenburg, A., Subramanian, K. 2005. Astrophysical Magnetic Fields and Nonlinear Dynamo Theory. Physics Reports, 417, 1-209.
17. Busse, F. H., Simitev, R. 2005. Dynamos Driven by Convection in Rotating Spherical Shells. Astronomische Nachrichten, 326, 231-240.
18. Bonanno, A., Elstner, D., Belvedere, G., Rüdiger, G. 2006. Solar Dynamo Models Driven by a Multi-Cell Meridional Circulation. Memorie della Societa Astronomica Italiana Supplement 9, 71.
19. Rüdiger, G. 2007. Tachocline, Dynamo and the Meridional Flow Connection. Highlights of Astronomy 14, 293-294.
20. Dikpati, M., Gilman, P. A. 2008. Global Solar Dynamo Models: Simulations and Predictions. Journal of Astrophysics and Astronomy 29, 29-39.

21. Rempel, M., Schüssler, M., Knölker, M. 2009. Radiative Magnetohydrodynamic Simulation of Sunspot Structure. Astrophysical Journal 691, 640-649.
22. Kitiashvili, I. N., Kosovichev, A. G., Wray, A. A., Mansour, N. N. 2010. Mechanism of Spontaneous Formation of Stable Magnetic Structures on the Sun. The Astrophysical Journal 719, 307-312.
23. Rimmele, T. R., the ATST Team 2008. The Unique Scientific Capabilities of the Advanced Technology Solar Telescope. Advances in Space Research 42, 78-85.
24. Stein, R. F., Benson, D., Nordlund, A. 2007. Solar Magneto-Convection Simulations. New Solar Physics with Solar-B Mission 369, 87.
25. Parker, E.N. 1955. Hydromagnetic Dynamo Models. Astrophysical Journal, 122, 293.
26. Brandenburg, A. 2007. Why Coronal Mass Ejections are Necessary for the Dynamo. Highlights of Astronomy, 14, 291-292.
27. Alexander, D., and 18 colleagues 2005. Solar Polar Imager: Observing Solar Activity from a New Perspective. Solar Wind 11/SOHO 16, Connecting Sun and Heliosphere 592, 663.

Science Objectives for an X-Ray Microcalorimeter Observing the Sun

J. Martin Laming[1]§, J. Adams[3], D. Alexander[8], M Aschwanden[12], C. Bailey[3], S. Bandler[3], J. Bookbinder[2], S. Bradshaw[8], N. Brickhouse[2], J. Chervenak[3], S. Christe[3], J. Cirtain[3], S. Cranmer[2], S. Deiker[12], E. DeLuca[2], G. Del Zanna[13], B. Dennis[3], G. Doschek[1], M. Eckart[3], A. Fludra[15], F. Finkbeiner[3], P. Grigis[2], R. Harrison[15], L. Ji[2], C. Kankelborg[14], V. Kashyap[2], D. Kelly[3], R. Kelley[3], C. Kilbourne[3], J. Klimchuk[3], Y.-K. Ko[1], E. Landi[6], M. Linton[1], D. Longcope[14], V. Lukin[1], J. Mariska[1], D. Martinez-Galarce[12], H. Mason[13], D. McKenzie[14], R. Osten[17], G. Peres[10], A. Pevtsov[16], K. Phillips[4] F. S. Porter[3], D. Rabin[3], C. Rakowski[1], J. Raymond[2], F. Reale[10], K. Reeves[2], J. Sadleir[3], D. Savin[11], J. Schmelz[9], R. K. Smith[2], S. Smith[3], R. Stern[12], J. Sylwester[5], D. Tripathi[13], I. Ugarte-Urra[7], P. Young[7], H. Warren[1], B. Wood[1]

[1]Naval Research Laboratory, [2]Smithsonian Astrophysical Observatory, [3]NASA Goddard Space Flight Center, [4]Mullard Space Science Laboratory, UCL, [5]Space Research Centre, Polish Academy of Sciences, [6]University of Michigan, [7]George Mason University, [8]Rice University, [9]University of Memphis, [10]University of Palermo, [11]Columbia University, [12]Lockheed Martin, [13]DAMTP, University of Cambridge,[14]Montana State University, [15]STFC Rutherford Appleton Laboratory,[16]National Solar Observatory, [17]Space Telescope Science Institute

ABSTRACT

We present the science case for a broadband X-ray imager with high-resolution spectroscopy, including simulations of X-ray spectral diagnostics of both active regions and solar flares. This is part of a trilogy of white papers discussing science, instrument (Bandler et al. 2010), and missions (Bookbinder et al. 2010) to exploit major advances recently made in transition-edge sensor (TES) detector technology that enable resolution better than 2 eV in an array that can handle high count rates. Combined with a modest X-ray mirror, this instrument would combine arcsecond-scale imaging with high-resolution spectra over a field of view sufficiently large for the study of active regions and flares, enabling a wide range of studies such as the detection of microheating in active regions, ion-resolved velocity flows, and the presence of non-thermal electrons in hot plasmas. It would also enable more direct comparisons between solar and stellar soft X-ray spectra, a waveband in which (unusually) we currently have much better stellar data than we do of the Sun.

1. THE PROMISE OF X-RAY MICROCALORIMETERS IN SOLAR PHYSICS

Coronal heating has been the central problem of the solar outer atmosphere for over half a century (e.g. Klimchuk 2006). The advent of high quality imaging data in the last 15 years (SOHO, TRACE, STEREO, Hinode, and now SDO) has forced the realization that the solar atmosphere is significantly more dynamic and turbulent than had been previously appreciated. Since much of this behavior occurs on few seconds to minutes timescales, comparable to the integration time to record a spectrum or to scan a raster to make an image with grating instruments, what is ideally needed is an instrument capable of imaging spectroscopy, with high throughput, and high spatial and spectral resolution. In this respect, recent advances in microcalorimeter arrays for solar X-ray observations have the potential to revolutionize our understanding of coronal plasmas, specifically with regard to the non-equilibrium plasma dynamics of solar flares and active regions. The microcalorimeter is well-matched to the parameters of interest for detailed studies of the fundamental plasma physics in the Sun's atmosphere, with excellent spectral (<2 eV), spatial (~arcsec) and temporal (millisecond) scales and broad spectral coverage from 0.2 to 10 keV. Such an instrument will be able to:

1. Detect and resolve hundreds of atomic transition lines from different ionization states of many elements in the solar atmosphere

§Contact information: j.laming@nrl.navy.mil; 202-767-4415

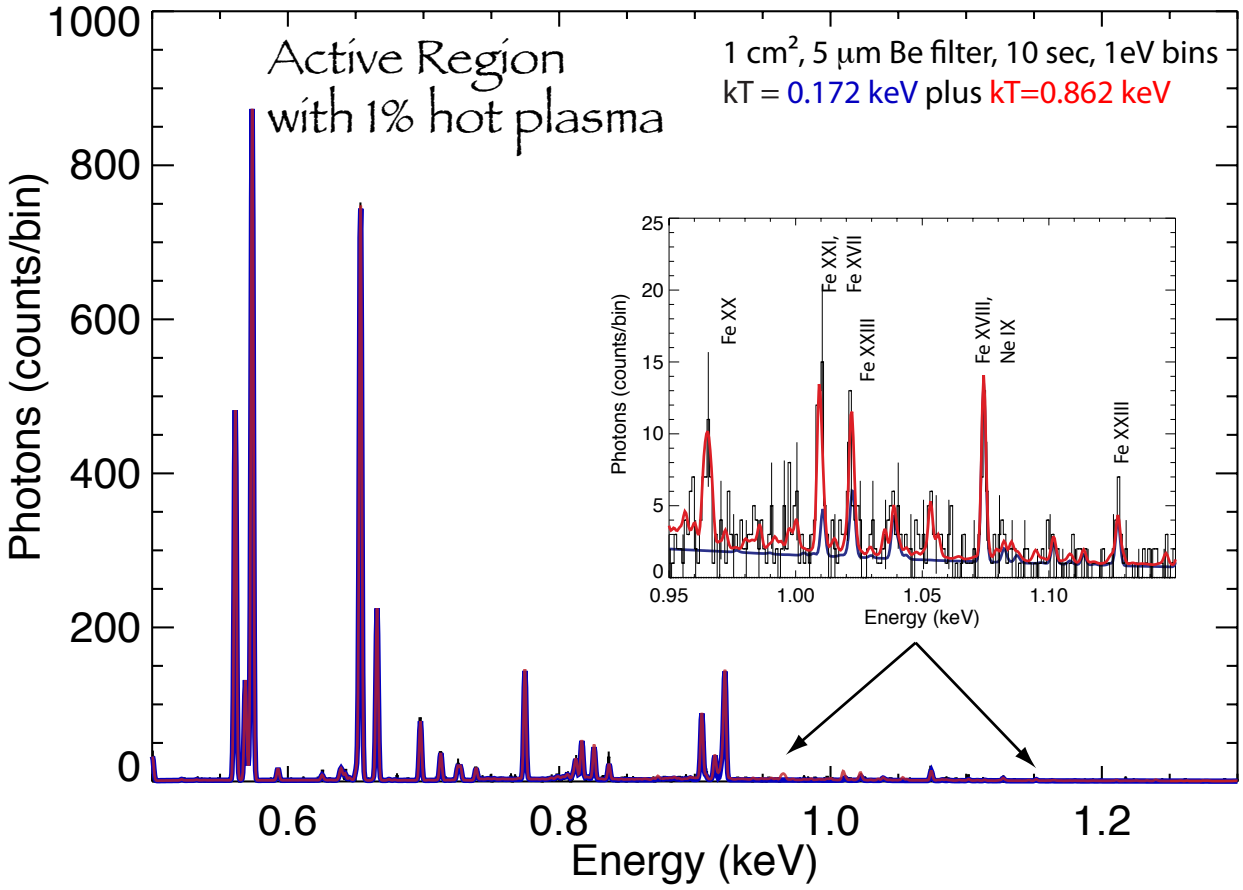


Figure 1 – The spectrum shown here is from a plasma at 2 MK combined with a 10 MK component whose emission measure is 1% of the dominant term (black is the simulated combined data; blue the cool phase only, red the combined model). The zoomed-in region inset shows strong lines of higher charge state iron ions that are characteristic of the higher temperature region. Note that this spectrum is from one square arcsecond, so spatial variations in the hotter plasma could easily be imaged with this technique.

2. Determine thermal and non-thermal electron distributions from lines and continuum
3. Determine plasma densities from ~10^{9}–10^{13} cm^{-3} and temperatures from <0.7 to >10 MK
4. Detect flows with velocities down to < 50 km s^{-1}
5. Determine the abundances of elements from carbon to nickel relative to hydrogen

The simultaneous combination of high resolution spectroscopic imaging observations across this entire energy range is new and tremendously powerful. Solar microcalorimeters will allow all of this information to be obtained simultaneously on time scales that allow us to follow the evolution of the radiative emission during the impulsive phase of flares. This is in contrast to the several minute cadence of previous instruments that sample only a narrow part of the spectrum and require rastering to build up images. The unique capabilities of these detectors will allow us to study how magnetic energy is released in magnetic reconnection events such as nano-flare heating in active regions, solar flares and, with different optics, coronal mass ejections with unprecedented spectral resolution.

With a 1-2 cm^2 mirror and a thin Be filter, a typical active region or C flare region can be observed at full resolution. Gradual spectral resolution degradation will occur at increased incident flux. Nonetheless, brighter flares, up to and including X class flares, can be observed at full spec-

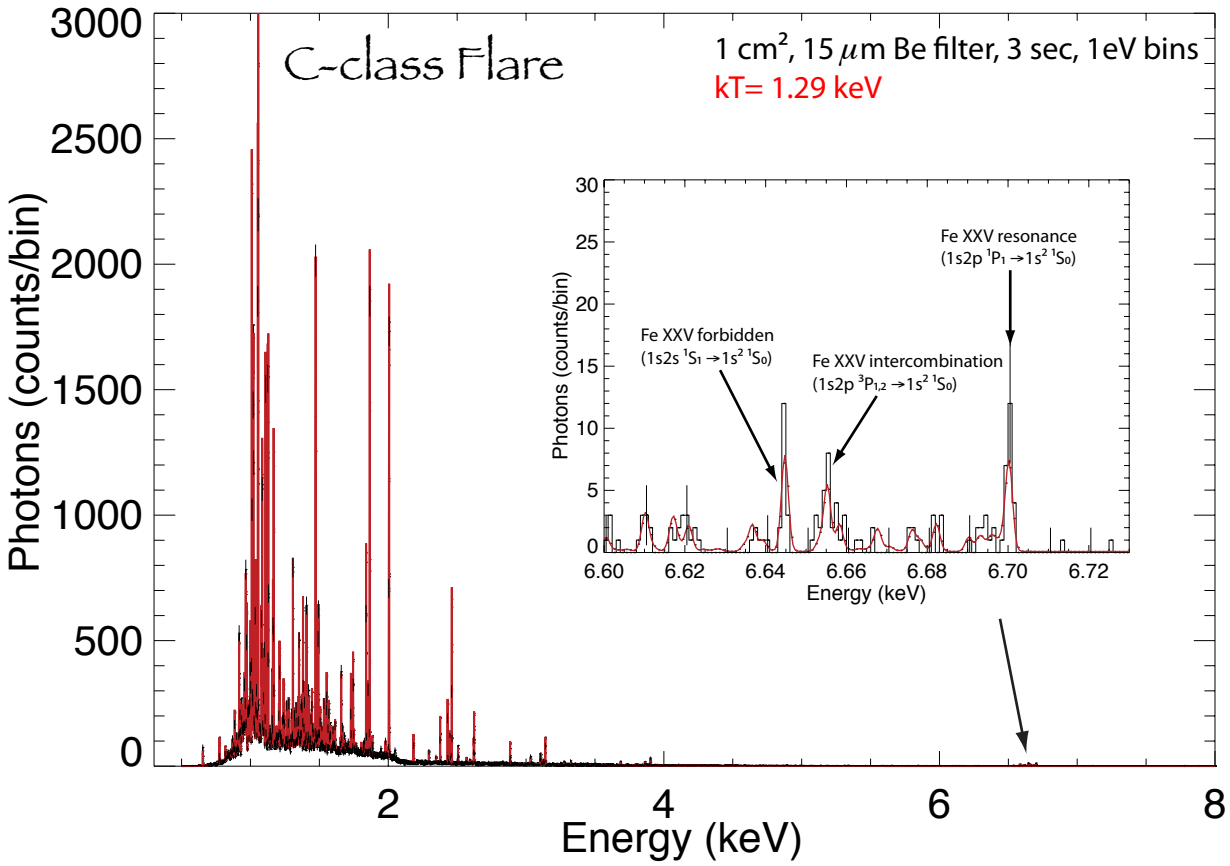


Figure 2 – The full spectrum shows the large number of lines available in a single 3 second observation of one square arcsecond, primarily from the iron L-shell ions combined with the H-and He-like lines of abundant ions. The inset is a blown up segment showing the resolved Fe XXV line complex with the red curve showing the model and the black the simulated data.

tral resolution using a combination of thicker filters (which remove the less-useful soft photons) and possibly a neutral density filter.

2. SOLAR ACTIVE REGIONS

The broadband response of a microcalorimeter is ideally suited to explore modern theories of coronal heating. Following Parker's (1988) initial suggestion, the idea that the corona is heated by short intense releases of energy, commonly known as "nanoflares" or "microflares" has gained currency. A number of indirect observations support this scenario. The scale heights of coronal loops are clearly incompatible with static models. Recent models of the FIP effect[1] where the ponderomotive force from Alfvén waves drives the fractionation (Laming 2004b; 2009) also strongly suggest a coronal source for these waves, again more compatible with impulsive rather than quasi-static coronal heating. Amidst this, the classical "smoking gun", the direct detection of hot plasma (with a low emission measure) signalling the location(s) of the heating event, has remained elusive.

Tantalizing hints are seen when the solar corona is imaged with Bragg Crystal Spectrom-

[1]*The coronal enhancement in abundance of elements with First Ionization Potential (FIP) less than about 10 eV, i.e. Mg, Si, Fe, over their photospheric values, while elements with higher FIP remain essentially unchanged.*

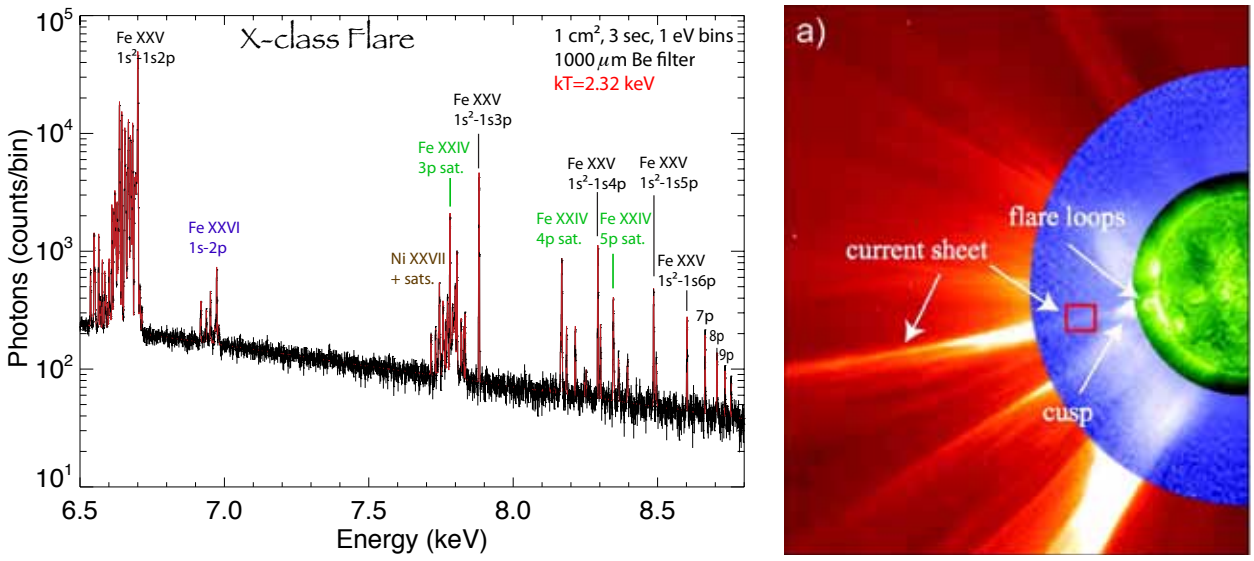


Figure 3 – Left: Theoretical solar flare spectrum, showing the Fe (6.7 keV) and Fe/Ni (8 keV) spectrum at a few eV resolution. RHESSI observes these Fe and Ni lines as single broad features at 1 keV resolution. The continuum is free-free and free-bound. The assumed temperature is 20MK (typical of medium-to-large flare peak temperatures) Right: Composite LASCO/C2, Mauna Loa/Mark IV, and EIT/Fe XII image showing the current sheet related to the CME that occurred on 2003 November 18 (taken from Bemporad 2008).

eters in the light of Mg XII 8.42 Å (Zhitnik et al. 2006, Urnov et al. 2007), in emission measure analyses of spectra taken with Hinode/EIS (Patsourakos & Klimchuk 2009)[2], with Hinode/XRT using different filter ratios (Schmelz et al. 2009, Reale et al. 2009a) and/or in quiet sun observations with RHESSI (McTiernan 2009, Reale et al. 2009b). On the other hand Warren et al. (2010) study the Hinode/EIS spectrum from the core of a solar active region and find no evidence for very hot plasma which presents a challenge to the theory of nanoflares. The observational evidence to date all suffers from the use of either filter ratios with a broad band telescope, where low temperature emission has to be corrected for, or the observation of an intrinsically weak signal at low spatial resolution. The imaging capabilities of a microcalorimeter array, combined with its broad bandpass and dynamic range, will allow detection of small admixtures of recently-heated plasma in a region otherwise dominated by cooler material.

Fig. 1 shows two views of a simulated single pixel calorimeter spectrum of a solar active region. The imaged emission measure is 4×10^{38} cm^{-3} at 2×10^{6}K, with a small contribution 4×10^{36} cm^{-3} at 10^{7} K, to model the anticipated nanoflare component. The left panel shows the broad band 0.5-1.3 keV region. The right panel shows a blow-up of the 0.95-1.15 keV region, and the number of hot lines (Fe XVII, XVIII, XX, XXI, XXIII) that immediately become visible. Outside this region the Mg XII 8.42Å Lyman α transition already used (Zhitnik et al. 2006, Urnov et al. 2007) becomes available, along with other He- and H-like systems.

The microcalorimeter resolution < 2 eV means that Lyman α from H-like ions Mg XII and up will be resolved, and the intensity ratio between the two components will be measurable. Several of these ratios varying from 2:1 have been previously detected by instruments on solar flare

[2]*Lines from hot plasma also appear in the EUV, but are weak and difficult to isolate from other emission from lower temperatures.*

missions (P78-1, Solar Maximum Mission, Hinotori, all without any imaging or other context instrumentation), and generally interpreted as due to effects of radiation transfer in an asymmetrical plasma, such that lines with higher opacity escape preferentially in one direction; we note this is easily tested with a calorimeter via observations of higher Lyman series lines. Laming (1990) suggested that an alternative explanation could lie with directed electrons, presumably accelerated along the magnetic field direction, which impart some polarization to the levels they excite. Upon radiative decay, the polarized light is naturally emitted in an anisotropic fashion, giving rise to different Lyman α intensity ratio for different angles of observation with respect to the "beam" axis. The relaxation time of ~keV electrons in coronal plasma is sufficiently short to make this unlikely in the case of a flare produced electron beam; the suprathermal electrons would have to be continuously replenished, however a nanoflare model of coronal heating might be expected to yield just such a signature, especially if the heating events are associated with small scale magnetic reconnection.

Accompanying the energy release in the corona by nanoflares should be episodes of evaporation from the chromosphere, following heat conduction downwards. This upflow is predicted to be 50-100 km s^{-1}, i.e. detectable with the microcalorimeter. To date, such upflows have generally only been convincingly detected during the impulsive phase of solar flares (e.g. Milligan et al. 2006ab) as a "shoulder" on the blue wing of an otherwise stationary line profile. Czaykowska et al. (1999) see an upflow of 50 km s^{-1} during the late gradual flare phase, through the shift of the entire line profile.

3. SOLAR FLARES

Solar flares have traditionally been observed with Bragg Crystal Spectrometers flown on missions at solar maximum around 1980 (P78-1, Solar Maximum Mission, Hinotori) and again around 1991 (Yohkoh). The microcalorimeter brings all of the attributes of these missions together with broad bandpass and X-ray imaging. The higher throughput compared to earlier missions will allow observations of the early and preflare stage of the event, before the hard X-ray burst. Battaglia et al. (2009) argue that this early emission stems from conduction-driven chromospheric evaporation following heating in the corona. With a spectral resolution of < 2 eV, the thermal Doppler widths of Fe lines in evaporating plasma would be easily resolved; most likely there is always also a "non-thermal" or turbulent component to the line width but it appears to be < 2 eV, from SMM BCS spectra of the main phase of flares (it is much larger at the flare impulsive stage, e.g. Alexander et al. 1998).

During this phase it would also be key to look for non-ionization equilibrium effects through line ratios in particularly the 6.7 keV Fe line complex which is made up of He-like Fe (Fe XXV) lines (resonance, intercombination, forbidden) and Fe XXIV and lower ionization stage dielectronic satellites. Intensity ratios of satellites formed by dielectronic recombination and the resonance line give electron temperature, those formed by inner-shell excitation give the Li-like/He-like Fe ion abundance ratio. If there is ionization equilibrium, the temperatures from the two ratios should be equal but non-equality would indicate either an ionizing or recombining plasma. Fig. 2 shows a simulated spectrum from one calorimeter pixel of a C class flare, obtained in 3 sec integration.

The 6.7 keV "Fe line" complex consists of many closely spaced lines. However, the higher-excitation 8 keV "Fe/Ni line" feature (see Phillips 2004), consisting mostly of Fe XXV $1s^2 - 1snp$ lines (n = 3,4,5...) and satellites which crowd together to form single spectral features, is

easier to resolve, and would therefore be sensitive to lower upflow velocities. This feature has been observed with solid-state detectors (e.g. on the NEAR spacecraft, resolution ~600 eV) and with RHESSI (1 keV resolution) but never with crystal spectrometers, in space or in the laboratory. The left panel of Fig. 3 shows a simulation of this spectral region. Ratios of satellites/parent (Fe XXV) lines give electron temperatures and possibly departures from thermal distributions. Evidence exists for high (approaching 10^{12}–10^{13} cm^{-3}) electron densities in solar flares (Phillips et al. 1996), using diagnostic line ratios in Fe XXI and Fe XXII observed with crystal spectrometers on SMM. The microcalorimeter would provide similar spectral resolution, but with significantly higher bandpass, throughput and spatial resolution. Hence the spatial variation of several different density diagnostic line ratios could be evaluated, with potentially important consequences for our understanding of the physics of solar flares.

4. OFF-LIMB OBSERVATIONS OF CMES AND THE SOLAR WIND

Off-limb solar observations will be perhaps where the combination of microcalorimeter attributes finds its optimum application. Observations of disk flares have revealed large blue shifts at onset in 10^7 K plasma (Innes et al. 2001, Innes et al. 2003), reaching as high as 650 - 1000 km s^{-1}. These velocities are similar to the Alfvén speed expected, and so could be the Alfvénic outflow from the reconnection region, or due to turbulence. Off limb current sheets have been imaged with Hinode/XRT (Savage et al. 2010), and modeled (Reeves et al. 2010). Abundances have been previously measured in CME current sheets (Ciaravella et al. 2002, Ko et al. 2003), with the result that the FIP effect is present and strong. Microcalorimeter observations should make substantial progress on outstanding unsolved issues surrounding magnetic reconnection; the partitioning of energy between thermal energies of electrons and ions, bulk kinetic energy and turbulence.

Evidence is also mounting that CMEs must be heated as they erupt, and that this heating represents a significant fraction of the CME energy budget. Analyses of ion charge states detected in situ (Rakowski et al. 2007), and of remotely sensed UV emission (Akmal et al. 2001; Ciaravella et al. 2001; Lee et al. 2009; Landi et al. 2010) reveal that previously cold (i.e. ~10^6 K coronal material or cooler) is heated to temperatures approaching 10^7 K as the filament erupts, and microcalorimeter observations ought to be able to detect such hot emission.

The shock wave driven ahead of the CME would also make an interesting target. Spectroscopic diagnostics exist within the calorimeter bandpass for the measurement of postshock electron temperature (the "G" ratio of He-like C V, N VI, O VII, etc). Ion temperatures can be measured from the line broadening, especially for shocks approaching 1000 km s^{-1}. The derived electron and ion internal energies should add up to the total shock energy (derived from kinematics), unless significant energy is being lost to the acceleration of solar energetic particles.

There is also evidence that the fast solar wind is heated in the region between about 1.5 and 5 $R_\odot$ heliocentric distance where it is accelerated, from analyses of ion charge states collected *in situ* (Laming 2004a, Laming & Lepri 2007) and the observations of optical and near IR lines from different Fe charge states during eclipses (Habbal et al. 2007; 2010). A microcalorimeter should be able to detect C V, VI, N VI, VII, and O VII, VIII in this region as confirmation of these inferences, and to compare with *in situ* observations. Although the electron temperature in this region may be as high as a few $\times 10^6$ K, Fe charges remain those of M-shell ions, because ionization equilibrium is rapidly lost in the low-density solar wind plasma.

5. CONCLUSIONS

While great progress has been made from X-ray imaging and UV and EUV imaging/spectroscopy, in the next decade we should return to the natural spectral band — the X-ray band — for studying high energy coronal phenomena in detail. This offers potentially decisive contributions to the emerging nanoflare model of coronal heating, through the detection of hot plasma associated with the heating. The microcalorimeter is also ideally suited to observations off-limb, and appears poised to advance this new field as well. The combination of spectral, spatial and temporal resolution over a broad spectral band afforded *simultaneously* by the microcalorimeter offers a major advance in diagnostic capability. Current instruments generally give either spatial or temporal resolution, and then over a relatively narrow spectral bandpass if high spectral resolution is required. The broad bandpass will also allow the acquisition of solar spectra that are much more comparable to stellar X-ray spectra returned by instruments on satellites like Chandra and XMM-Newton. Currently, we are in the unusual situation of having better stellar spectra in this spectral region than we do of the Sun (Osten 2010); a microcalorimeter will redress this mismatch, enabling direct comparisons with stellar observations in addition to addressing the issues in solar physics discussed above.

6. REFERENCES

[1] Akmal, A., et al. 2001, ApJ, 553, 922
[2] Alexander, D., Harra-Murnion, L. K., Khan, J. I., & Matthews, S. A. 1998, ApJ, 494, L235
[3] Battaglia, M., Fletcher, L., & Benz, A. O. 2009, A&A, 498, 891
[4] Bemporad, A. 2008, ApJ, 689, 572
[5] Ciaravella, A., et al. 2001, ApJ, 557, 351
[6] Ciaravella, A., et al. 2002, ApJ, 575, 1116
[7] Czaykowska, A., De Pontieu, B., Alexander, D., & Rank, G. 1999, ApJ, 521, L75
[8] Habbal, S. R., et al. 2007, ApJ, 663, 598
[9] Habbal, S. R., et al. 2010, ApJ, 708, 1650
[10] Innes, D. E., et al. 2001, ApJ, 549, L249
[11] Innes, D. E., McKenzie, D. E., & Wang, T. 2003, Solar Physics, 217, 267
[12] Klimchuk, J. A. 2006, Solar Physics, 234, 41
[13] Klimchuk, J. A., Patsourakos, S., & Cargill, P. J. 2008, ApJ, 682, 1351
[14] Laming, J. M. 1990, ApJ, 357, 275
[15] Laming, J. M. 2004a, ApJ, 604, 874
[16] Laming, J. M. 2004b, ApJ, 614, 1063
[17] Laming, J. M., & Lepri, S. T. 2007, ApJ, 660, 1642
[18] Laming, J. M. 2009, ApJ, 695, 954
[19] Landi, E., Raymond, J. C., Miralles, M. P., & Hara, H. 2010, ApJ, 711, 75
[20] Lee, J.-Y., Raymond, J. C., Ko, Y.-K., & Kim, K.-S. 2009, ApJ, 692, 1271
[21] Milligan, R. O., et al. 2006a, ApJ, 638, L117
[22] Milligan, R. O., et al. 2006b, ApJ, 642, L169
[23] Osten, R. 2010, Heliophysics Decadal White Paper
[24] Parker, E. N. 1988, ApJ, 330, 474
[25] Phillips, K. 2004, ApJ, 605, 921
[26] Phillips, K. et al. 1996, ApJ, 466, 549
[27] McTiernan, J. M. 2009, ApJ, 697, 94
[28] Patsourakos, S., & Klimchuk, J. A. 2009, ApJ, 696, 760
[29] Rakowski, C. E., Laming, J. M., & Lepri, S. T. 2007, ApJ, 667, 602
[30] Reale, F., Testa, P., Klimchuk, J. A., & Parenti, S. 2009a, ApJ, 698, 756
[31] Reale, F., McTiernan, J. M., & Testa, P. 2009b, ApJ, 704, L58
[32] Reeves, K., et al. 2010, ApJ, 721, 1547
[33] Savage, S. L. et al. 2010, ApJ, 722, 329
[34] Schmelz, J. T., et al. 2009, ApJ, 693, L131
[35] Warren, H. P., Brooks, D. H., & Winebarger, A. R. 2010, arXiv1009.5976

Understanding the Coronal Abundance Anomalies of the Sun

J. Martin Laming[1], Yuan Kuen Ko[1], Enrico Landi[2], Vyacheslav Lukin[1], Rachel Osten[6], Kenneth Phillips[3], Cara Rakowski[1], John Raymond[4], Daniel Savin[5], Brian Wood[1]

[1]Naval Research Laboratory, [2]University of Michigan, [3]Mullard Space Science Laboratory, UCL, [4]Smithsonian Astrophysical Observatory, [5]Columbia University, [6]STScI

Abstract We would like to emphasize the importance of abundance anomalies in regions of the solar corona, wind, and accelerated particles, for the insights they provide into wave-particle interactions in the solar atmosphere and how these might inform models of coronal heating.

Concentrating on the FIP effect, we argue that this will be comprehensively explained by a model in which the ponderomotive force due to Alfvén waves in the chromosphere preferentially lifts chromospheric ions, but not neutrals, into the corona. This model appears to require a coronal origin for the Alfvén waves, suggesting coronal heating by nanoflares or Alfvén resonance. The FIP effect is also relevant to discovering the sources and origins of the fast and slow speed solar winds, and possibly the variable helium abundance therein. It impacts modern research on the solar elemental composition, and the variability of FIP fractionation also impedes attempts to forecast the solar EUV irradiance. The variation of coronal abundance anomalies observed in a carefully selected sample of solar-like stars a reveals new insights about the Sun's place in the stellar context, and properties of waves observed in helio- and asteroseismology.

The study of element abundances requires spectroscopic and *in situ* observations. Spectroscopic observations need to be at high spatial resolution (to resolve individual loops) and with appropriate diagnostics. This work will also require significant investment in theory/modeling/data analysis.

1. Introduction

In 1859, Bunsen and Kirchhoff, working in Heidelberg, identified the element sodium in the visible spectrum of the solar atmosphere obtained earlier by Fraunhofer in 1814. This provided the first step along the route to understanding the elemental composition of the 6000 K part of the solar atmosphere now known as the photosphere. A little over a century later, Pottasch (1963) inferred element abundances in the 1 MK hot outer atmosphere of the Sun, known as the corona, and noted that abundances of elements such as Mg, Si, and Fe were significantly higher than their photospheric values. This coronal abundance anomaly began to be taken seriously in the 1980's (e.g. Meyer 1985a,b) and is now commonly known as the "FIP Effect", owing to its dependence on First Ionization Potential (FIP). Briefly, elements with FIP less than about 10 eV (low FIP), which comprises those elements which can be photoionized by hydrogen Lyman α, are observed to be enhanced in abundance in the solar corona and slow speed solar wind by a factor of about 3-4, while high FIP elements retain their photospheric composition. Helium, the element with the highest FIP (24.587 eV) is actually depleted in the solar wind (both fast and slow), and, as far as

we can tell, in the corona and coronal holes too (Laming & Feldman 2001, 2003). In solar coronal holes and the fast wind, the FIP effect is much reduced.

In this white paper, we argue for a renewed effort in understanding this phenomenon. Recent observations have revealed that the chromosphere, the part of the solar atmosphere most likely as the site of the fractionation where low FIP elements are ionized and high FIPs are neutral, is exceedingly dynamic in nature. This largely rules out models of the fractionation based on various forms of diffusion, because it is an inherently slow process. We will argue below that the most likely agent by which fractionation proceeds is the ponderomotive force, which arises as Alfvén waves encounter the steep density gradients of the chromosphere. This places the FIP effect at the center of a number of issues surrounding the problems of coronal heating and the origin of the solar wind.

2. The Ponderomotive Force Model

Turbulence, as revealed through non-thermal line broadening, is ubiquitous in the solar upper atmosphere (Feldman & Doschek 1977), and chromosphere (De Pontieu et al. 2007). Due to the strong damping of compressional modes, this is almost certainly in the form of Alfvén waves. Due to their lack of compression, Alfvén waves are hard to observe and decisively identify as such; progress along these lines has been made only recently (Tomczyk et al. 2007, Jess et al. 2009). The FIP effect, if due to the ponderomotive force, possibly offers the most robust, albeit indirect, observable of Alfvén waves. The ponderomotive force arises from a second order term in $\delta \mathrm{J} \times \delta \mathrm{B}$ in the momentum equation, and can be manipulated (e.g Litwin & Rosner 1998; Laming 2009) for waves of frequency $\omega << \Omega$, the ion gyrofrequency, into the form

$$F = \frac{q^2}{4m\Omega^2}\frac{\partial E_\perp^2}{\partial z} = \frac{q^2}{4m\Omega^2}\frac{V_A^2}{c^2}\frac{8\pi U_{waves}}{h} = \frac{m}{2\rho}\frac{U_{waves}}{h} \tag{1}$$

where $E_\perp$ is the perpendicular wave electric field and q is the ion charge. In equation 1, h is the chromospheric scale height and U_{waves} is the wave energy density, assumed constant under these conditions. The dependence on the Alfvén speed, V_A, means that the ponderomotive force is usually strongest at the top of the chromosphere. The ponderomotive acceleration, F/m, is independent of the ion mass, leading to the mass independent fractionation that is observed. Waves incident from below on the footpoints of a closed coronal loop are generally reflected back down again unless special conditions (i.e. resonance between the waves and the loop) are met. Such reflected waves generally give rise to an inverse FIP effect, seen in later-type stars (see below) but not in the Sun. Coronal waves incident on the loop footpoint from above and reflecting back into the corona give rise to upwards pointing ponderomotive force and the FIP effect as observed in the Sun. Such coronal waves must be generated in the corona, strongly suggesting that models of coronal heating by nanoflares of Alfvén resonance should be favored. The Laming (2004, 2009) model comes about as a natural extension of existing work on Alfvén wave propagation in the solar atmosphere with essentially no extra physics required. It is also the model most worked out

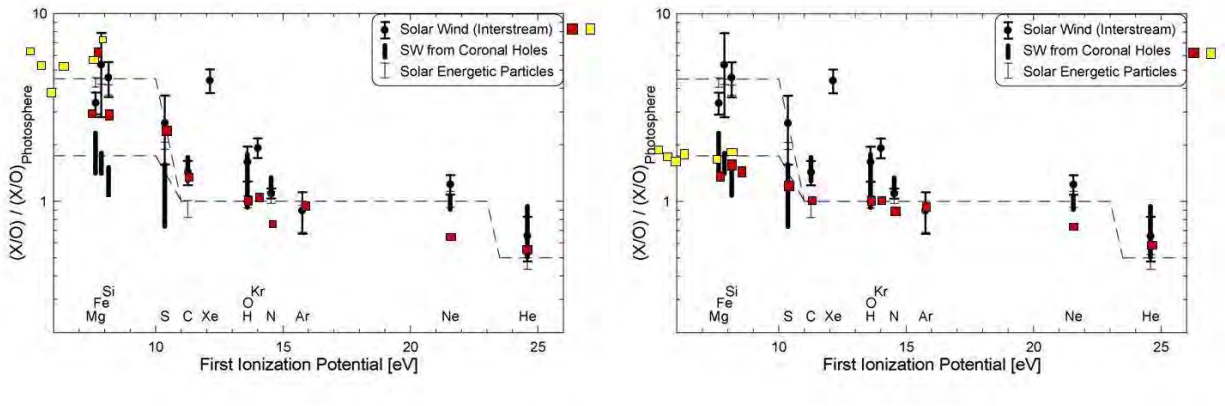


Fig. 1.— Comparison of model results (red and yellow symbols) with observations of von Steiger et al. (2000), shown in black, for the slow (left panel) and fast (right panel) solar wind. Red symbols are for direct comparison with von Steiger et al. for all of their elements listed except Xe. Yellow symbols indicate modeled fractionations for (from left) K (FIP 4.341 eV), Na (5.139 eV), Al (5.986 eV), Ca (6.113 eV), Ni (7.640 eV), and W (7.864 eV). The dashed lines are drawn by von Steiger et al. (2000) to guide the eye through the different fractionation patterns.

in detail to interpret observations, giving it unique potential for diagnosing wave processes in the corona and chromosphere.

Figure 1 (left panel) shows the FIP fractionation calculated for a closed magnetic loop with an Alfvén wave of approximately 50 km s^{-1} rms amplitude. This is slightly higher than generally observed from non-thermal line broadening (typically 20 - 30 km s^{-1}), but is supported by some recent observations which reveal a less intense broad component to strong coronal emission lines (e.g. Peter 2010, though whether this a truly a wave remains unclear). This suggests that a small part of the imaged coronal volume is oscillating with large amplitude, also consistent with coronal heating by e.g. nanoflares or Alfvén resonance. Figure 1 (right) shows FIP fractionations calculated for an open field region, assuming the turbulence spectrum calculated by Cranmer et al. (2007), with an additional contribution from Isobe et al. (2008) to represent chromospheric reconnection, possibly forming polar plumes. The reduction in FIP effect observed in the fast solar wind, included in Figure 1 from von Steiger et al. (2000), arises naturally in the ponderomotive force model of the FIP effect.

3. The Relevance of the FIP Effect

3.1 Origins of the Slow and High Speed Solar Wind: Observations are beginning to identify open field regions as possible sources of the slow speed solar wind by detecting the FIP effect in such outflows. Brooks & Warren (2010) and Ko et al. (2006) find fractionated outflows on the boundaries of equatorial coronal holes. Similarly, within the unfractionated plasma of polar coronal holes and the fast solar wind, polar plumes are observed also to be fractionated to varying degrees. It appears above that turbulence in an open field region induced by kink mode oscillations

of the photospheric region of a flux tube produces no fractionation, so these observations imply the existence of another process. Low altitude reconnection has frequently been suggested as a means of driving outflows, or generating an electron heat flux to excite ion cyclotron waves. We argue that the "extra" fractionation observed in these regions derives from waves also generated by the reconnection. One school of thought holds that reconnection alone provides the fractionation, as ions in the reconnection region are heated and neutrals are not, and that such a process can naturally lead to a FIP fractionation as observed (Arge & Mullan 1998). However applying the collision rates in Laming (2004) to their scaling relations suggests that a mass dependence should be present in the fractionation, with heavier ions enhanced in abundance proportional to their mass. This is not observed in the FIP effect, but may be relevant to the enhancement of trans-Fe elements observed in impulsive SEP events (e.g. Reames & Ng 2004)[1].

3.2 The Helium Abundance in the Solar Corona and Wind: Figure 1 demonstrates that as well as providing a compelling description of the abundance enhancement of low FIP ions (the FIP effect), the observed depletion of He is also reproduced. This arises because He, having the highest FIP of all elements, remains neutral the longest in the chromosphere, and consequently is the last element to be picked up by the ponderomotive force and accelerated into the corona. We argue that this is a much more likely origin for the depletion of He in the solar wind than the usually discussed "inefficient Coulomb drag". This would require a solar wind in which protons flow faster than α particles, in contradiction to observations at 1 AU, and should not be restricted solely to He, again in conflict with observations.

The He abundance in the slow speed solar wind is observed to be variable through the solar cycle (Aellig et al. 2001; Kasper et al. 2007). The slowest wind at solar minimum has the lowest He abundance, down to about 2%, rising to 4-5% at solar maximum. The fastest slow speed wind is much less variable. Also, CME plasma observed *in situ* is well known to have higher He abundances than other regions of the solar wind. We anticipate that these variations will arise as the chromospheric ionization balance responds to heating in the corona and subsequent downwards heat conduction.

3.3 Solar Elemental Composition and Helioseismology: Some elements like the noble gases have no photospheric absorption lines, and so may only be detected spectroscopically in the corona. Using such observations to infer solar abundances for these elements relies on the assumption of no fractionation between photosphere and corona. A case in point is Ne.

A reevaluation of the solar composition (Asplund et al. 2005) revised the solar metallicity down from 0.0170 (Grevesse & Sauval 1998) to 0.0122, largely driven by improved analysis of lines of C, N, and O. The lower metallicity decreases the sound speed at the base of the solar convection zone,

[1]Such chromospheric reconnection also appears incapable of producing an inverse FIP effect, as seen in active stars (section 3.6), though this is possible with the ponderomotive force. Reconnection would not automatically give lower fractionation in the fast wind compared to the slow wind, as observed, and as also happens naturally with the ponderomotive force.

yielding now unacceptable agreement between observed and modeled sound speeds (see e.g. Fig. 1 in Guzik & Mussack 2010). The reduced abundances also decrease the depth of the convection zone in solar models, again worsening agreement between models and helioseismic inversions (e.g. Basu & Antia 2004).

One early solution proposed was to increase the abundance of Ne to compensate (Antia & Basu 2005; Bahcall et al. 2005). Asplund et al. (2009) argue for an Ne/O abundance ratio of 0.175 ± 0.031 following the measurements of Young (2005) in the solar corona. A survey in nearby active stars had previously yielded the abundance ratio Ne/O in the range 0.3 - 0.4 (Drake & Testa 2005), significantly higher than the solar coronal value and closer to that suggested by helioseismology. Drake & Testa (2005) argue that the relative constancy of this ratio in a sample of over 20 stars suggests that the "true" Ne/O abundance ratio should be around 0.4, and that the variation in the solar corona must be due to some unknown fractionation.

Our models for the solar FIP effect described above, appear necessarily to produce simultaneously *coronal depletions of Ne and He*. The depletion is absent when the Inverse FIP effect prevails, strongly supporting, at least in part, the conjecture of Drake & Testa (2005). The depletion of He in the solar wind has been a longstanding problem, now solved by our FIP effect model. We argue that Ne should behave in a very similar way, as seen in Fig. 1 above.

3.4 Coronal Heating: The observed FIP effect in the solar corona and wind only arises in realistic situations if the Alfvén waves have an origin *in the corona*. Alfvén waves originating for example from photospheric p-modes[2], would only give rise to positive FIP effect if they happened to be in resonance with the coronal loop, i.e. the wave frequency $\omega = V_A/2L$ where V_A is the coronal Alfvén speed, and L is the loop length. Otherwise, the more common Alfvén waves off resonance will be reflected back down again, and the associated ponderomotive force would give an "Inverse FIP Effect", i.e. a depletion of low FIP ions in the corona. This has never been observed on the Sun (though see below for a discussion of active stars). It seems likely that upward propagating p-modes *do not* convert to Alfvén waves, presumably because they are mostly reflected lower down at an acoustic cut off.

A coronal origin for Alfvén waves constrains coronal heating models. Nanoflares or Alfvén resonance scenarios are strongly preferred. In each case, the Alfvén waves are not the cause of the heating, but are released as a "by-product" of the heating mechanism. For example Sturrock (1999) reviews the mechanisms by which various wave modes may be excited by reconnection. The reconnected field line is generally distorted, and this can either propagate away from the reconnection site as an Alfvén wave, or emit magnetoacoustic waves traveling perpendicularly to the magnetic field direction. Kigure et al. (2010) find that a significant fraction of the magnetic energy released (several tens of %, depending on geometry and plasma β) can be carried off by

[2]Which generate Alfvén waves at the chromospheric layer where the sound speed and Alfvén speed are equal, by mode conversion (e.g. Cally 1995) or parametric conversion (e.g. Zaqarashvili & Roberts 2006).

Alfvén or magnetoacoustic (fast or slow mode) waves, with Alfvén waves dominating for $\beta < 1$. Higher frequency waves may be associated with "plasmoids" or magnetic islands in the current sheet (e.g. Sturrock 1999). The potential importance of this plasmoid instability (in increasing the reconnection rate at suitably high Lundquist number) has been recognized for some time (Lee & Fu 1986; Biskamp 1986; Shibata & Tanuma 2001), but it is only recently that the linear and non-linear properties of this instability are being quantified (e.g. Loureiro et al. 2007; Bhattacharjeee et al. 2009; Lapenta 2008; Daughton et al. 2009). The associated turbulence remains to be investigated.

3.5 Solar EUV Irradiance: The solar EUV irradiance depends linearly on the abundances of elements that when ionized in the solar corona become the strongest emitters of EUV radiation. For the Sun these are elements such as Mg, Si, and most importantly, Fe. Since these are all low FIP elements, the degree of FIP fractionation present in, for example, an active region directly affects its radiative output. Any attempt to forecast the EUV irradiance must incorporate some model of the FIP fractionation. The monitoring of space debris orbits requires knowledge of the earth's atmospheric density at 400 km altitude to ±5% (see e.g. Storz et al. 2005). This in turn translates to knowledge of the solar EUV/X-ray irradiance to similar accuracy, i.e. a precision far greater than our current ability to predict the abundances of the important emitting elements in the solar corona.

3.6 The Sun's Place in the Stellar Context: In a survey of stars restricted to X-ray luminosities less than 10^{29} ergs s^{-1}, Wood & Linsky (2010) discover a correlation between spectral type and degree of FIP fractionation, shown in Fig. 2. From a logarithmic FIP bias, $(\log(X/H)_{phot} - \log(X/H)_{Cor})$ of -0.6 at spectral type G2 (corresponding to the Sun), the FIP effect steadily diminishes, disappearing at spectral type K5, and becoming an Inverse FIP Effect for the M dwarfs in the sample. We interpret this as being due to an increasing energy density of Alfvén waves being reflected from loop footpoints as one moves to later spectral type. This arises as stronger magnetic fields in these stars place the layer where mode or parametric conversion of p-modes to Alfvén waves occurs *below* the acoustic cut off. Thus we can possibly infer from the behavior of coronal abundances details of wave phenomena of relevance to helio- and asteroseismology. The FIP and Inverse FIP effects connected in this way are just one example of how a solar-stellar research connection can benefit both communities. A more extended discussion on this theme is given in another white paper submitted to the decadal survey (Osten 2010).

4. Conclusions

Our present knowledge of the FIP effect derives from spectroscopic observations and significant efforts over the past 30 years in data analysis, theory and modeling. The situation echoes that of the discovery of the solar photospheric composition, beginning with the "laboratory astrophysics" of Kirchoff and Bunsen. The main message of this white paper is that potentially significant conclusions can be reached about the physics of the solar atmosphere from the serendipitous discovery that the elemental composition is variable, combined with the interpretation that waves and tur-

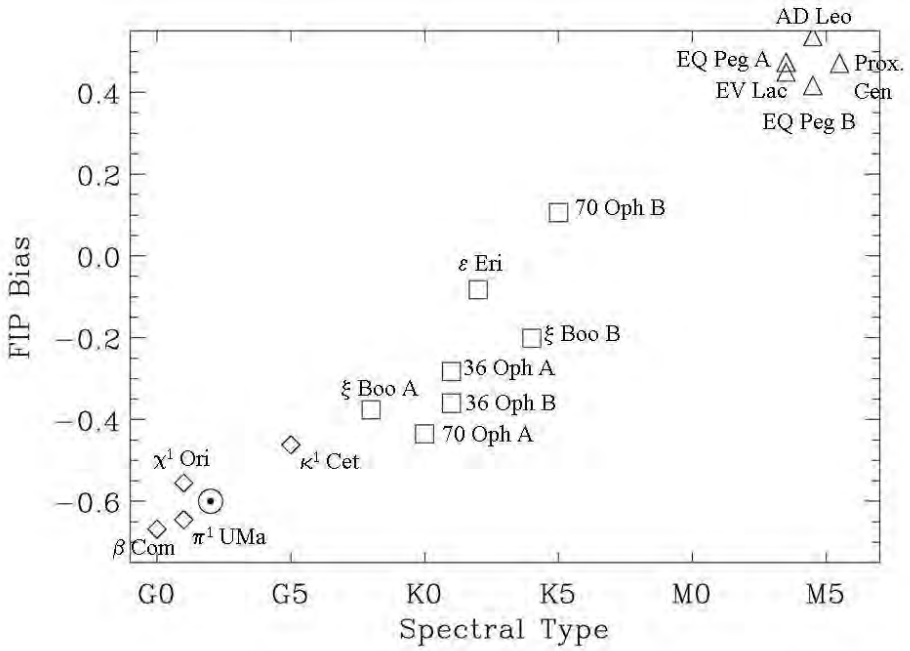


Fig. 2.— Survey of FIP fractionation ($\log(X/H)_{phot} - \log(X/H)_{Cor}$) observed in a sample of dwarf stars, from Wood & Linsky (2010), where values below zero indicate a solar-like FIP effect and values above zero indicate an inverse FIP effect. Diamonds indicate measurements from Telleschi et al. (2005), triangle from Liefke et al. (2008), with a solar value from Feldman & Laming (2000). For all GK stars the FIP bias calculations include corrections for stellar photospheric abundances from Allende Prieto et al. (2004), but for the M stars there are no stellar photospheric measurements available so we have to simply assume solar photospheric abundance apply. Among these, EV Lac is the strongest case, due to chromospheric evaporation of photospheric abundance material observed during flares (Laming & Hwang 2009, and refs therein). For the purposes of this figure, we avoid extremes of stellar activity, confining our attention to stars with $\log L_X \leq 29$.

bulence in the solar atmosphere are responsible[3]. However such ideas and discoveries can only be exploited if funding is available for such Research and Analysis. This topic is elaborated on in much more generality in another white paper (Klimchuk et al. 2010), also to be submitted to the Heliophysics Decadal Survey. Here, we have tried to give an example of the type of science that could be investigated, and its range of implications for solar physics.

We also argue for more emphasis on high resolution (both spatial and spectral) spectroscopic missions with suitable diagnostics to further exploit such insights. A possibility is described in further white papers (Laming et al. 2010; Bandler et al. 2010; Bookbinder 2010) describing science, instrumentation and possible missions motivated by developments in microcalorimeter technology. When concerned with element abundances, spectroscopy is absolutely essentially, since it is the only means by which emissions from different elements may be identified.

[3]Also reached an in quasi-serendipitous manner; it certainly wasn't the product of a carefully planned and executed scientific investigation!

5. References

Aellig, M. R., Lazarus, A. J., & Steinberg, J. T. 2001, Geophys. Res. Lett., 28, 2767
Antia, H. M., & Basu, S. 2005, ApJ, 620, L129
Allende Prieto, C., Barklem, P. S., Lambert, D. L., & Cunha, K. 2004, A&A, 420, 183
Arge, C. N., & Mullan, D. J. 1998, Sol. Phys., 182, 293
Asplund, M., Grevesse, N., & Sauval, A. J. 2005, ASP Conf. Ser. 336, 25
Bahcall, J. N., Basu, S., & Serenelli, A. M. 2005, ApJ, 631, 1281
Bandler, S. R., et al. 2010, White Paper submitted to Heliophysics Decadal Survey
Basu, S., & Antia, H. M. 2004, ApJ, 606, L85
Bhattacharjee, A., et al. 2009, Phys. Plasmas, 16, 112102
Biskamp, D. 1986, Phys. Fluids, 29, 1520
Bookbinder, J. A., et al. 2010, White Paper submitted to Heliophysics Decadal Survey
Brooks, D. H., & Warren, H. P. 2010, arXiv:1009.4291
Cally, P. S. 1995, ApJ, 451, 372
Cranmer, S. R., van Ballegooijen, A. A., & Edgar, R. J. 2007, ApJS, 171, 520
Daughton, W., et al. 2009, Phys. Rev. Lett., 103, 065004
De Pontieu, B., et al. 2007, Science, 318, 1574
Drake, J .J., & Testa, P. 2005, Nature, 436, 525
Feldman, U., & Doschek, G. A. 1977, ApJ, 216, L119
Feldman, U., & Laming, J. M. 2000, Phys. Scripta., 61, 222
Grevesse, N., & Sauval, A. J. 1998, Space Science Reviews, 85, 161
Guzik, J. A., & Mussack, K. 2010, ApJ, 713, 1108
Isobe, H., Proctor, M. R. E., & Weiss, N. O. 2008, ApJ, 679, L57
Jess, D. B., et al. 2009, Science, 323, 1582
Kasper, J. C., et al. 2007, ApJ, 660, 901
Kigure, H., et al. 2010, PASJ, 62, 993
Klimchuk, J. A., et al. 2010, White Paper submitted to Heliophysics Decadal Survey
Ko, Y.-K., et al. 2006, ApJ, 646, 1275
Laming, J. M., & Feldman, U. 2001, ApJ, 546, 552
Laming, J. M., & Feldman, U. 2003, ApJ, 591, 1257
Laming, J. M. 2004, ApJ, 614, 1063
Laming, J. M. 2009, ApJ, 695, 954
Laming, J. M., & Hwang, U. 2009, ApJ, 707, L60
Laming, J. M., et al. 2010, White Paper submitted to Heliophysics Decadal Survey
Lapenta, G. 2008, Phys. Rev. Lett., 100, 235001
Liefke, C., Ness, J.-U., Schmitt, J. H. M. M., & Maggio, A. 2008, A&A, 491, 859
Litwin, C., & Rosner, R. 1998, ApJ, 506, L143
Lee, L. C., & Fu, Z. F. 1986, JGR, 91, 6807
Loureiro, N. F., Schekochihin, A. A., & Cowley, S. C. 2007, Phys. Plasmas, 14, 100703
Meyer, J.-P. 1985a, ApJS, 57, 151

Meyer, J.-P. 1985b, ApJS, 57, 173
Osten, R. A., et al. 2010, White Paper submitted to Heliophysics Decadal Survey
Peter, H. 2010, A&A, 512, A51
Pottasch, S. R. 1963, ApJ, 137, 945
Reames, D. V., & Ng, C. K. 2004, ApJ, 610, 510
Shibata, K., & Tanuma, S. 2001, Earth Planets Space, 53, 473
Storz, M. F., et al. 2005, Advances in Space Research, 36, 2497
Sturrock, P. A. 1999, ApJ, 521, 451
Telleschi, A., et al. 2005, ApJ, 622, 653
Tomczyk, S., et al. 2007, Science, 317, 1192
von Steiger, R., et al. 2000, JGR, 105, 27217
Widing, K. G., & Feldman, U. 1992, ApJ, 392, 715
Wood, B. E., & Linsky, J. L. 2010, ApJ, 717, 1279
Young, P. R., & Mason, H. E. 1997, Solar Physics, 175, 523
Young, P. R. 2005, A&A, 444, L45
Zaqarashvili, T. V., & Roberts, B. 2006, A&A, 452, 1053

Using Solar Neutrons to Understand Solar Acceleration Processes

David J. Lawrence[1], William C. Feldman[2], Dennis Haggerty[1], George Ho[1], Ralph McNutt[1], James Miller[3], Richard Miller[3], Patrick Peplowski[1], Edmond Roelof[1]

[1]Johns Hopkins University Applied Physics Laboratory
[2]Planetary Science Institute
[3]University of Alabama-Huntsville

1. Introduction: Importance of Solar Neutrons for Understanding Solar Acceleration Processes

The development of an understanding of energetic-particle acceleration mechanisms has long been a goal of heliospheric physics. Although much progress has been made since the space-flight era began, many gaps in our knowledge remain. The largest gaps involve energetic-ion acceleration in solar flares. These ions are difficult to detect remotely if they do not have access to magnetic field lines that connect the corona to interplanetary space. Yet, 90% of all ions accelerated in large impulsive flares remain tied to the Sun on closed magnetic field lines (Hua and Lingenfelter, 1987) and those that do escape remain confined to spatially-limited magnetic flux bundles in interplanetary space limiting the ability to directly observe these ions. Whereas energetic electrons can be imaged indirectly through bremsstrahlung, gyrosynchroton radiation, and hard x-rays, energetic ions are nearly invisible except through their coupling to secondary neutral radiation in the form of gamma rays and/or neutrons.

Several recent studies (see, e.g., Miller et al. (1997); Ryan (2000); Hua et al. (2002); Murphy et al. (2007), and references therein) have shown that combined analyses of gamma rays and neutrons from solar flares are essential to adequately constrain the many parameters needed to account for the total number, energy spectrum, time dependence, and angular distribution of all ions accelerated. This combined analysis is also essential for constraining the magnetic and plasma structure of the flare neighborhood in the corona that control trapping and transport, such as the elemental composition, density and temperature scale height, magnetic loop structure, acceleration altitude, and the MHD and plasma turbulence near the flare site.

Because of sensitivity issues stemming from the necessity of measuring neutrons at 1 AU[1], to date, many more flares have been observed in gamma rays than in neutrons. This circumstance has put a filter on the types of impulsive flares that have been amenable to a joint neutron and gamma ray analysis. For the most part, these flares have been necessarily large. Another element of the observational filter stems from the fact that all previous neutron detections have been at energies larger than 25 MeV. However, all neutrons produced through interactions between accelerated ions and coronal plasma ions result from two dominant types of nuclear reactions: knock-on/charge-exchange, and spallation/ evaporation. In the first case, the emerging neutrons have a range of energies that extend up to that of the incident ion energy per nucleon, and in the second case, the neutron energy distribution is a Maxwellian having a temperature of several MeV. Lately, there has been a realization that even relatively low energy ions (E_i < 30 MeV/nucleon) can produce significant fluxes of low energy neutrons (E_n < 10 MeV). The most

[1] Neutrons in free space have a lifetime of ~15 minutes (Nieto et al., 2008). Therefore, most neutrons of moderate energy (less than few 10s of MeV) decay before ever reaching the Earth and are not detected.

significant reactions in this third class of reactions are the exothermic reactions induced by alpha particles involving isotopes ^{13}C, ^{25}Mg, and ^{26}Mg, (and their inverse reactions) and to a much smaller extent, of low threshold alpha reactions involving isotopes ^{14}N, ^{18}O, ^{22}Ne, ^{29}Si, ^{54}Fe, and ^{56}Fe (and their inverse reactions) (Share et al., 2010). Detection of the first reaction type will come from solar sites located preferentially near the limb of the disk to facilitate detection of the high-energy portion of the neutron spectrum. Because of the large pitch angles of trapped protons after initial acceleration, these ions preferentially convert to neutrons having directions perpendicular to the radial, yielding a higher detection probability at Earth if the flare is near the limb. On the other hand,

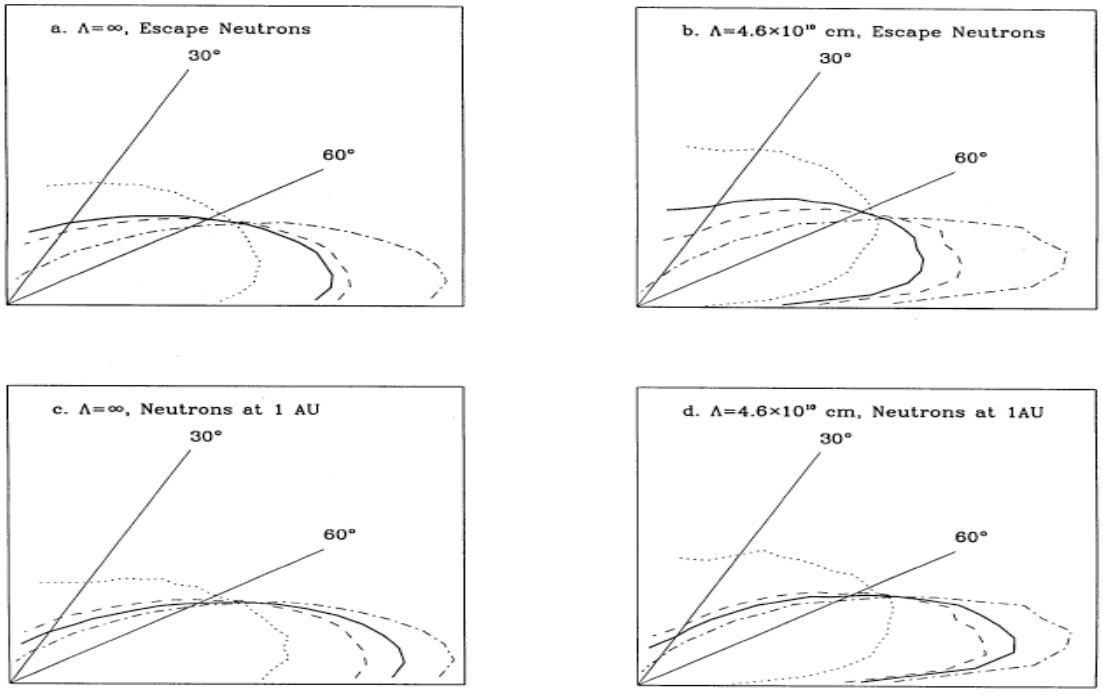


Figure 1. Zenith angular dependence of escaping neutrons (top two panels); zenith angular dependence of neutrons at 1 AU (bottom panels). Left panels have no ion pitch angle scattering; right panels are for saturated pitch angle scattering. Figures from Hua et al., (2002).

detection of low energy neutrons in the second and third class of reactions, which is only possible at distances much closer to the Sun than 1 AU is facilitated from flare sites nearer to disk center because boil-off neutrons from highly excited spallation nuclei and exothermic reactions are nearly isotropic. Here, the coronal escape probability is maximum in the radial direction where the neutrons will encounter a reduced column density of overlying coronal protons. This difference in the neutron angular distribution of escaping neutrons is illustrated in Figure 1, from simulations carried out by Hua et al. (2002).

In summary, the objectives of studying solar neutrons is to fill the gaps in our understanding of ion acceleration, trapping, and transport in the corona during solar flares. This can be accomplished through analysis of combined neutron, gamma ray, extreme ultraviolet (EUV), radio wave and energetic particle data sets. The significance of these studies is to: 1) better understand magnetic reconnection as revealed in solar

flares and coronal mass ejections; 2) to understand the plasma processes that accelerate and transport particles; and 3) to explore the full range of flares that produce neutrons. Up to now, only the very largest impulsive flares have been amenable for study at 1 AU. All of these objectives are specifically called out in the 2005 Heliophysics Roadmap for Science and Technology.

2. Current and Potential New Measurements of Solar Neutrons

At present, the Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) spacecraft (Lin et al., 2002) provides the most comprehensive source of high-energy X-ray and gamma-ray data that are suitable for studies of ion acceleration in solar flares. Solar-flare neutrons have for some time been, and are currently, monitored by ground-based neutron detectors (e.g., (Ryan, 2000) and references therein]) The only source of solar neutron data from space at present is the Neutron Spectrometer (NS) sensor of the Gamma-Ray and Neutron Spectrometer (GRNS) instrument (Goldsten et al., 2007) on the MErcury Surface, Space ENvirnoment, GEochemistry, and Ranging (MESSENGER) spacecraft (Solomon et al., 2007). This spacecraft is presently orbiting the Sun between 0.30 AU and 0.55 AU en route to orbit insertion about Mercury in March 2011. Between now and the end of the mission, the MESSENGER orbit will decrease from its present aphelion of 0.55 AU to that of the aphelion of Mercury, 0.47 AU. Its orbit about the Sun is presently scheduled to last one Earth year beyond orbit insertion now, scheduled for 18 March, 2011. However, the termination date of the MESSENGER mission is presently under consideration to be extended for at least one additional year to include the period up to solar maximum.

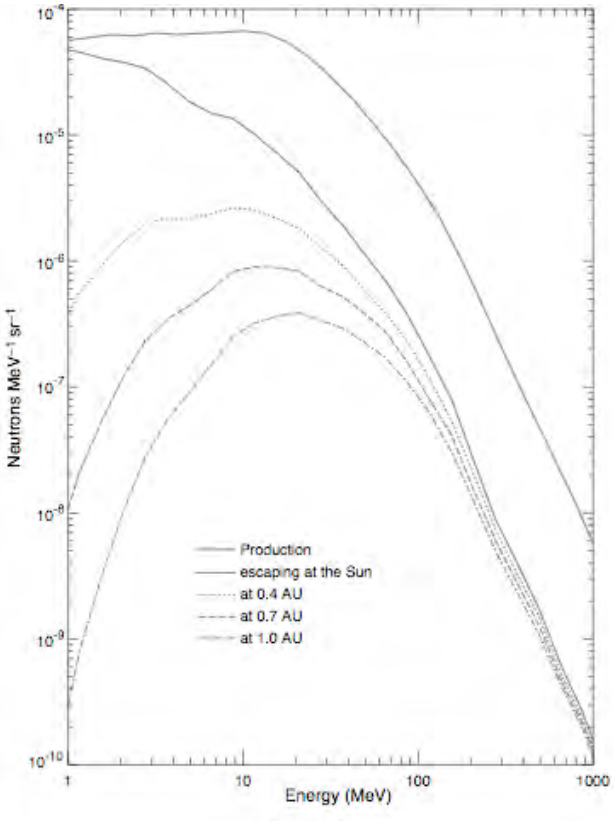


Figure 2. Energy spectrum of neutrons at various stages of production and propagation to 1 AU (from Murphy et al., 2007).

The Neutron Spectrometer aboard MESSENGER is sufficiently sensitive to solar neutrons as to complement the measurements of solar flare gamma rays that are now provided by RHESSI. The energy range of the MESSENGER NS covers the maximum flux portion of the neutrons generated by ions accelerated during solar flares, between 0.5 and 7.5 MeV. This range nicely brackets the maximum flux portion of flare neutrons that was simulated by Murphy et al. (2007), as shown in Figure 2. This new NS data set will then effectively provide a separate scientific mission of opportunity for MESSENGER that could function as an active addition to the Heliophysics Great Observatory.

A neutron and gamma-ray spectrometer were included as part of the baseline mission design for the Solar Probe Plus (SPP) mission (STDT,

2008). However, the final instrument selection that took place in mid-2010 did not include either instrument. For the neutron measurements in particular, the SPP mission easily provides the best opportunity for making robust measurements of solar neutrons because the close distance of SPP to the sun (~0.05 AU) results in an increased sensitivity for 5 MeV neutrons by six orders of magnitude from 1 AU[2]; the SPP mission provides for a factor of ~800 increased sensitivity from 0.3 AU, which is the minimum distance achieved by the MESSENGER NS.

Finally, neutron and gamma-ray instruments are listed as potential measurements for the Solar Sentinels mission, which would approach the sun with a distance of 0.25 AU (Solar Sentinels STDT, 2006).

3. Observations of Neutrons at 0.48 AU with the MESSENGER NS

A recent report by Feldman et al., (2010) illustrates the type of interplanetary-space neutron measurements that can be made with the MESSENGER NS. On 31 December 2007, neutrons were detected by the MESSENGER NS during an M2 flare (see Figures 3 and 4). This flare contained multiple acceleration episodes as seen in Type III radio bursts. After these bursts ended, both the energetic-particle and neutron fluxes decayed smoothly to background with an e-folding decay time of 2.84 hours, spanning a nine-hour time period. This time is considerably longer than the mean lifetime of a neutron, which indicates that either the observed neutrons were generated in the spacecraft by solar-energetic-particle protons, or that they originated on the Sun. As part of this study, Feldman et al. (2010) investigated the possibility that the detected neutrons might have been generated on the spacecraft due to energetic ion interactions with spacecraft material. Unfortunately, the MESSENGER spacecraft does not have any instruments to measure energetic ions having energies greater than about 30 MeV per nucleon, which is the general threshold energy for producing neutrons. The MESSENGER NS is sensitive to energetic charged particles, but at present the response of the NS is not sufficiently modeled to discriminate between energetic electrons or ions. Even with these limitations, a full spacecraft particle transport model that was validated with neutrons measured at Mercury (Lawrence et al., 2010), was used to determine that if all the NS detected charged particles were ions (instead of the more likely mix of electrons and ions), then at most only 20% of the detected neutrons could have been locally generated. As a consequence, Feldman et al. (2010) concluded that a likely explanation of their long duration is that energetic ions were accelerated over an extended time period onto closed magnetic arcades above the corona and then slowly pitch angle–scattered by coronal turbulence into their chromospheric loss cones. Because of their relatively low energy loss in the Neutron Spectrometer (0.5–7.5 MeV), most of these neutrons beta decay to energetic protons and electrons close to the Sun, thereby forming an extended seed population available for further acceleration by subsequent shocks driven by coronal mass ejections in interplanetary space.

---

[2] This rough scaling assumes that the detected neutron flux scales as $F \sim e^{-(t/\tau)}/D^3$, where t is the neutron time-of-flight from the sun to the sensor, τ = 886 seconds is the neutron lifetime (Nieto et al., 2008), and D is the distance from the sun to the measurement. The factor of $1/D^3$ is included to account for the time spreading of the neutrons ($1/D$) and an areal flux factor ($1/D^2$). A full sensitivity analysis requires a detailed understanding of a given instrument sensitivity and mission profile (e.g., time spent at a given sun distance) and is beyond the scope of this white paper.

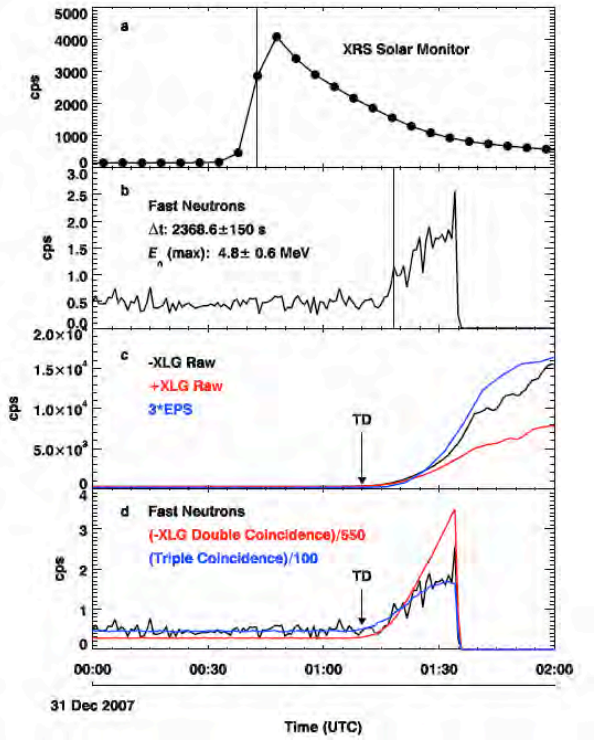


Figure 3. (a) Response of the MESSENGER X-ray Spectrometer (XRS) solar monitor to X-rays during the 31 December 2007 solar flare. (b) NS detector response to fast neutrons produced by the solar flare. (c) Rates of energetic particle counts at MESSENGER measured using the two Li-glass (LG) scintillators and the MESSENGER Energetic Particle Spectrometer (EPS) EPS detector. (d) An overlay of fast-neutron, lithium glass double-coincidence, and lithium glass/borated plastic triple-coincidence counting rates before a data dropout due to an instrument safing from high rates. The energy range of the EPS is 65 keV to 1.0 MeV; the energy range of the single raw LG counting rates for electrons is from 300 keV to about 2 MeV and protons up to about 30 MeV; the energy range of the double-coincidence rates for electrons is from about 2 to 20 MeVand protons from about 30 to 120 MeV; and the energy range for triple-coincidences of electrons is greater than about 20 MeV and protons greater than about 120 MeV (figure taken from Feldman et al., 2010).

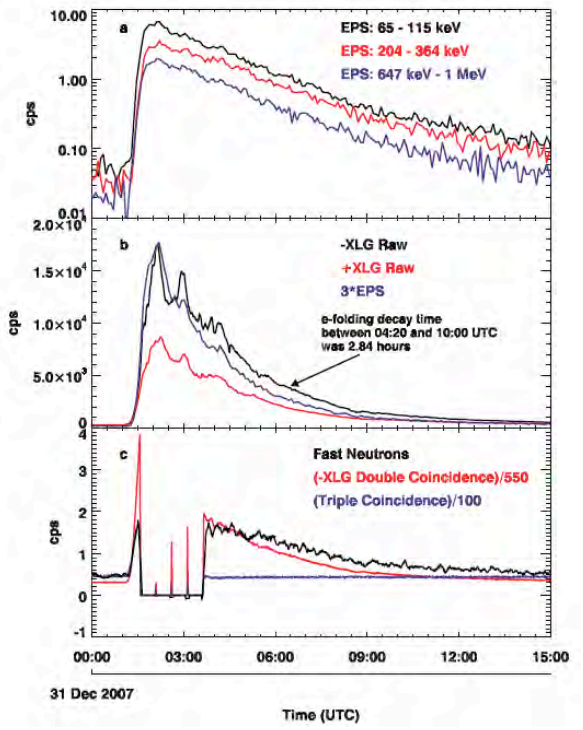


Figure 3. (a) Time-dependent counting rates of energetic electrons measured with the MESSENGER EPS detector. The top (black) curve is for 65 to 115 keV, the middle (red) curve is for 204 to 364 keV, and the bottom (blue) curve is for 647 keV to 1 MeV. (b) Rates of energetic particle counts at MESSENGER measured using the two Li-glass scintillators and the EPS detector. The EPS counting rates have been scaled to match that registered by the LG scintillator. (c) An overlay of fast-neutron (black curve), double-coincidence (red curve), and triple-coincidence (blue curve) counting rates throughout the entire fast-neutron event. The three vertical lines during the safe mode record three failed attempts to exit the safe mode before the BP counting rate dropped below the preprogrammed safe level. (figure taken from Feldman et al., 2010).

In response to the Feldman et al. (2010) study, Share et al. (2010) have suggested that the most likely explanation for the detected neutrons is that they were generated locally on the spacecraft from flare-related ions and alpha particles. The particular concerns of

Share et al. (2010) include the following: 1. The inferred number of accelerated protons at the Sun for this M2-class flare was 10X larger than any flare observed to date; 2. The onset and duration of the 'solar' neutron count rate was similar to that of the solar energetic particles (SEPs) (albeit with some exceptions); and 3. They argue that Feldman et al. (2010) underestimate the local neutron production from energetic ions. There is clearly more work to be done to sort out and understand these conflicting interpretations of the MESSENGER measurements.

Further progress can be made in understanding and resolving these discrepancies through multiple tasks: 1) Further modeling of the MESSENGER NS and spacecraft can be carried out to understand the response of the NS to charged particles (to help discriminate between energetic electrons and ions), as well as better understand the local generation of neutrons. 2) The ability to distinguish between solar and local neutron production will be helped considerably after Mercury orbital insertion with the aid of the 12-hour orbital period, which should effect a significant modulation of SEP protons incident on the spacecraft during flare events. Specifically, one can monitor the modulation of SEP protons by the Mercury magnetosphere using its simultaneously measured plasma and fields structure by onboard plasma, magnetic field, and energetic particle experiments and determine if any detected neutrons have or do not have a similar modulation. 3) Now that the sun is becoming more active, many more flare events will be detected with the MESSENGER NS. It is likely that future events will have coincident energetic particle measurements from other spacecraft assets such as STEREO and/or ACE. Starting in March 2011, the MESSENGER spacecraft will be in orbit about Mercury for at least one Earth year. If the mission proceeds nominally, it will have enough fuel to enable an extended mission of up to a few years. Therefore, there will be ample opportunity to measure multiple solar flares with possible neutron detections.

4. Future of Solar Neutron Studies

Measurements of solar neutrons provide a unique opportunity to enhance our understanding of solar acceleration processes in all types of flares having an expanded range of strengths. The only mission currently operating that can provide any information about solar neutrons in the neutron energy range of 0.5 – 10 MeV is the MESSENGER mission. However, since it is a planetary exploration mission, there is no explicit funding to carry out the analysis and modeling tasks described in Section 3. In particular, as more neutron events are detected, more effort will be needed to analyze and interpret these events along with comparative event studies. The required funding for such studies should be modest and on the order of standard Science, Research and Technology (SR&T) grants.

Second, flight opportunities for neutron detectors to fly inside of 0.3 AU should be strongly encouraged. SPP clearly provided the best opportunity likely to be seen in a generation, as it will be traveling to within 9.5 R_S (or 0.05 AU). If other such flight programs are proposed (e.g., Solar Sentinels, which may fly to within 0.25 AU), a neutron instrument should be strongly considered. In summary, to fully understand solar acceleration mechanisms, it is required to have robust neutron measurements in concert with coordinated observations of gamma-rays, energetic ions and electrons, EUV, and radio waves.

References

Feldman, William C., et al. (2010) Evidence for Extended Acceleration of Solar-Flare Ions from 1–8-MeV Solar Neutrons Detected with the MESSENGER Neutron Spectrometer, Journal of Geophysical Research, 115, A01102, 10.1029/2009JA014535.

Goldsten, J. O., et al. (2007), The MESSENGER Gamma-Ray and Neutron Spectrometer, Space Sci. Rev., 131, 339-391.

Hua, X. M. and R. E. Lingenfelter (1987), Solar flare neutron production and angular dependence of the capture gamma-ray emission, Solar Physics, 107, 351 – 383, doi:10.1007?BF00152031.

Hua, X.-M., B. Kozlovsky, R. E. Lingenfelter, R. Ramaty, and A. Stupp (2002), Angular and energy-dependent neutron emission for solar flare magnetic loops, Astrophys. J. Suppl., 140, 563-579.

Lawrence, David J., et al. (2010) Identification and Measurement of Neutron-absorbing Elements on Mercury's Surface, Icarus, 10.1016/j.icarus.2010.04.005, 209, 195 – 209, 2010.

Lin, R. P., et al. (2002), The Reuven Ramaty High-Energy Solar Spectroscopic Imager (RHESSI), Solar Phys., 210, 3-32.

Miller, J. A., P. J. Cargill, E. A. Gordon, G. D. Holman, B. R. Dennis, T. N. LaRosa, R. M. Winglee, S. G. Benka, and S. Tsuneta (1997), Critical issues for understanding particle acceleration in impulsive solar flares, J. Geophys. Res., 102, 14631-14660.

Murphy, R. J., B. Kozlovsky, G. H. Share, X.-M. Hua, and R. E. Lingenfelter (2007), Using gamma-ray and neutron emission to determine solar flare accelerated particle spectra and composition and the conditions within the flare magnetic loop, Astrophys. J. Suppl., 168, 167-194.

Nieto, Michael Martin et al. (2008), Testing the Unitarity of the CKM Matrix with a Space-Based Neutron Decay Experiment, Modern Physics Letters A, Vol. 23, #21, 1735 – 1743.

Ryan, J.M., Long-duration solar gamma-ray flares, Space Sci. Rev., 93, 581-610, 2000.

Share, Gerald H. et al. (2010) Physics of Solar Neutron Production: Questionable Detection of Neutrons from the 2007 December 31 Flare, http://arxiv.org/abs/1007.2349, submitted for publication to J. Geophys. Research.

Solar Probe STDT Report (2008) Solar Probe Plus: Report of the Science and Technology Definition Team, NASA Report NASA/TM—2008–214161.

Solar Sentinels STDT Report (2006) Solar Sentinels: Report of the Science and Technology Definition Team, NASA Report NASA/TM—2006–214137.

Solomon, S. C., R. L. McNutt, Jr., R. E. Gold, and D. L. Domingue (2007), MESSENGER mission overview, Space Sci. Rev., 131, 3-39.

Magnetospheric Emission from Extrasolar Planets

J. Lazio,[1] Jet Propulsion Laboratory
T. Bastian, NRAO
G. Bryden, Jet Propulsion Laboratory
W. M. Farrell, NASA/GSFC
J.-M. Griessmeier, ASTRON
G. Hallinan, National U. of Ireland
J. Kasper, Harvard-Smithsonian Center for Astrophysics
T. Kuiper, Jet Propulsion Laboratory
A. Lecacheux, Observatoire de Paris
W. Majid, Jet Propulsion Laboratory
R. Osten, STScI
E. Shkolnik, Carnegie Institute of Washington
I. Stevens, University of Birmingham
D. Winterhalter, Jet Propulsion Laboratory
P. Zarka, Observatoire de Paris

[1] 818-354-4198; *Joseph.Lazio@jpl.nasa.gov*



Executive Summary

The magnetospheric emissions from extrasolar planets represent a science frontier for the next decade. All of the solar system giant planets and the Earth produce radio emissions as a result of interactions between their magnetic fields and the solar wind. In the case of the Earth, its magnetic field may contribute to its habitability by protecting its atmosphere from solar wind erosion and by preventing energetic particles from reaching its surface. Indirect evidence for at least some extrasolar giant planets also having magnetic fields includes the modulation of emission lines of their host stars phased with the planetary orbits, likely due to interactions between the stellar and planetary magnetic fields. If magnetic fields are a generic property of giant planets, then extrasolar giant planets should emit at radio wavelengths allowing for their *direct* detection. Existing observations place limits comparable to the flux densities expected from the strongest emissions. Additional sensitivity at low radio frequencies coupled with algorithmic improvements likely will enable a new means of detection and characterization of extrasolar planets within the next decade.

1 Frontier Question: What is the Luminosity Function of Extrasolar Planetary Magnetospheric Emissions?

Looking to the next decade, the luminosity function of magnetospheric emissions represents a scientific frontier that is ripe for discovery and would present a new means for characterization for extrasolar planets. A combination of algorithmic improvements and developments for existing telescopes as well as the new telescopes under construction should yield at least an order of magnitude of additional sensitivity.

2 Science Opportunity: Extrasolar Planetary Magnetospheric Emission

Search for and exploit extrasolar planetary magnetospheric emissions as a means of directly detecting and characterizing those planets.

2.1 Planetary Magnetospheric Emission

The Earth and gas giants of our solar system are "magnetic planets" because they contain internal dynamo currents that generate planetary-scale magnetic fields. These magnetic fields are immersed in the solar wind, a supersonic magnetized plasma. The solar wind deforms the planetary magnetic field, compressing the field on the front side and elongating it on the back, forming a "teardropped"–shaped magnetosphere aligned with the solar wind flow (Figure 1). The magnetopause forms the boundary between the magnetosphere, in which the planet's magnetic field is dominant, and the solar wind. The stellar wind incident on the magnetopause is an energy source to the planetary magnetosphere.

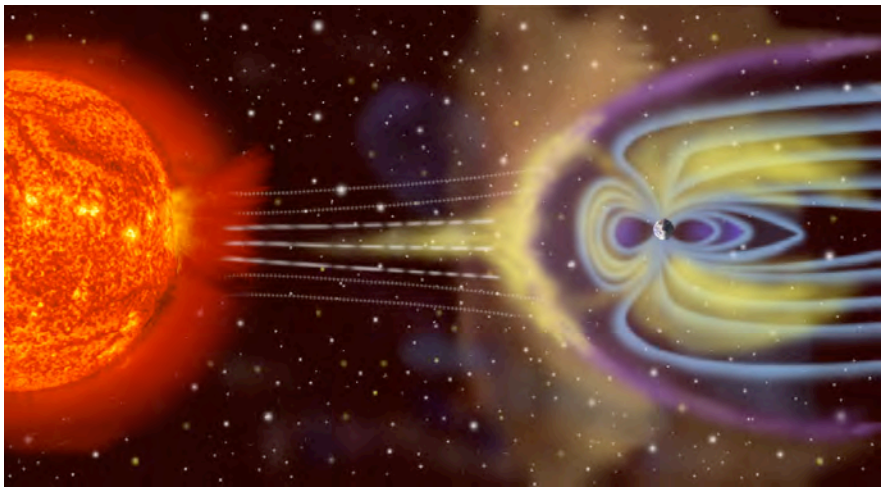


Figure 1. *Illustration of the interaction between the Sun and Earth's magnetosphere (not to scale). In this example, an eruption from the Sun has reached the Earth, compressed the magnetosphere, and injected energetic particles.*

Credit: NASA

The magnetospheres of the solar system's magnetic planets host radio-wavelength masers, generated by electron cyclotron radiation ("electron cyclotron masers"). The magnetosphere-solar wind interaction produces energetic (keV) electrons that then propagate along magnetic field lines into auroral regions, where an electron cyclotron maser is produced. Specific details of the electron cyclotron maser emission vary from planet to planet, depending upon secondary effects as the planet's magnetic field topology. Nonetheless, applicable to all of the magnetic planets is a macroscopic relation relating the incident solar wind power P_{sw}, the planet's magnetic field strength, and the median radio luminosity L_{rad} (Figure 2). Various investigators (e.g., Millon & Goertz 1988) have found

$$L_{rad} = \varepsilon P_{sw}^{x}$$

with ε the efficiency at which the solar wind power is converted to radio luminosity, and $x \approx 1$. The value for ε depends on whether one considers the magnetic energy or kinetic energy carried by the stellar wind. The strong solar wind dependence is manifest in the fact that the Earth's luminosity is larger than that of either Uranus or Neptune, even though their magnetic fields are 10–50 times stronger than that of the Earth.

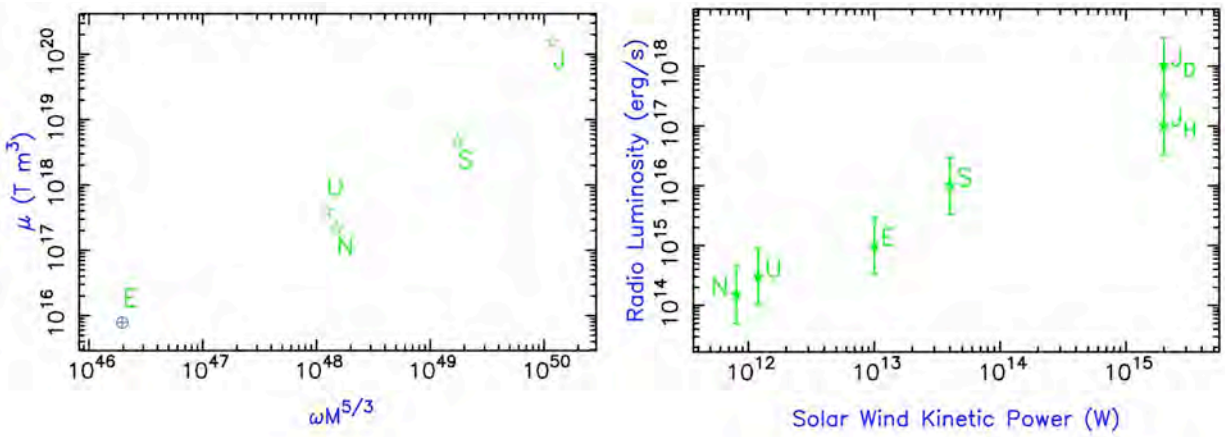


Figure 2 *(Left) Magnetic moments of solar system planets as a function of the rotation rate ω and mass M. (Right) Radio luminosities of solar system planets as a function of the incident solar wind kinetic power. (Adapted from Zarka et al. 2001.) Even though the magnetic field strength of Earth is less than that of Uranus or Neptune, it is more luminous because it is closer to the Sun and a larger solar wind power is incident. These relations can be both descriptive and predictive, as the luminosity of Uranus and Neptune were both predicted successfully before the arrival of the Voyager 2 spacecraft.*

The electron cyclotron maser occurs below a characteristic emission frequency determined by the cyclotron frequency in the magnetic polar region, which in turn depends upon the planet's magnetic moment or magnetic field strength. Using similar scaling laws based on the solar system planets, one can predict this characteristic frequency. For reference, the characteristic frequency of Jupiter is approximately 40 MHz (≈ λ7.5 m).

These scaling relations are not only descriptive but also *predictive*. Before *Voyager 2* reached both Uranus and Neptune, their luminosities were predicted (Desch & Kaiser 1984; Desch 1988; Millon & Goertz 1988). For both planets, the predictions were in excellent agreement with the measurements.

2.2 Extrasolar Planetary Magnetic Fields

Observations over the past 5 years show strong evidence suggesting the presence of a magnetic interaction between a star and its hot Jupiter (Shkolnik et al. 2005, 2008). For instance observations of HD 179949 and υ Andromeda show chromospheric hot spots that rotate with the orbital periods of their inner planets. Further, two other stars, HD 189733 and τ Bootis show similar evidence of chromospheric variability, which most probably are tied to flaring activity modulated with orbital period of their planets (Walker et al. 2008).

The solar system scaling laws enable quantitative predictions for an extrasolar planet's luminosity (Zarka et al. 1997; Farrell et al. 1999; Zarka et al. 2001; Lazio et al. 2004; Stevens 2005; Zarka 2007; Griessmeier et al. 2007). In the case of a known planet, these estimates depend upon measured parameters—the planet's mass and orbital semi-major axis—or upon parameters that can be estimated reasonably—for example, the rotation period can be assumed to be of order 10 hr, or for hot Jupiters taken to be the orbital period assuming that the planet is tidally locked to its host star. In the case of stars not yet known to have planetary companions, radio limits can be inverted to obtain constraints on the presence of planets, especially useful for active stars for which the radial velocity method is limited.

Planets around young stars may be more luminous than the solar system planets. Changes in the stellar wind can produce significant changes in a planet's luminosity (Griessmeier et al. 2005a, 2007). Wood et al. (2002, 2005) find the stellar mass loss as a function of age, $dM/dt \propto t^x$, with $x \approx -2$; the stellar wind around a 1 Gyr old star may be 25 times as intense as the solar wind. Young stars are often not observed in radial velocity surveys because the high stellar activity levels make it problematic

to isolate a planetary signal. Thus, a blind survey for magnetospheric emissions is a search methodology that could mitigate a selection bias in the current census. As an illustration, early predictions for the flux density of the planet orbiting τ Boo were of order 1–3 mJy (Farrell et al. 1999; Lazio et al. 2004). More recent estimates, that attempt to take into account the stellar wind strength of τ Boo, predict flux densities as high as 300 mJy (Stevens 2005; Griessmeier et al. 2007). If iron-rich "super-Earths" exist, they may also have sufficiently strong magnetic fields to be detectable.

3 Scientific Context: Magnetic Fields and Planetary Characterization

The dynamo currents generating a planet's magnetic field arise from differential rotation, convection, compositional dynamics, or a combination of these in the planet's interior. Consequently, knowledge of the planetary magnetic field places constraints on the thermal state, composition, and dynamics of the planetary interior, all of which will be difficult to determine by other means.

Planetary Interiors: For the solar system planets, the composition of the conducting fluid ranges from liquid iron in the Earth's core to metallic hydrogen in Jupiter and Saturn to perhaps a salty ocean in Uranus and Neptune. Likewise, radio detection of an extrasolar planet would indicate the planet's internal composition, insofar as it would require the planet to have a conducting interior. Combined with an estimate of the planet's mass and radius, one could infer the interior composition by analogy to the solar system planets.

Planetary rotation: The rotation of a planet imposes a periodic modulation on the radio emission, as the emission is preferentially beamed in the direction of the local magnetic field and will change if the magnetic and spin axes of the planet are not aligned. For all of the gas giant planets in the solar system, this modulation *defines* the rotation periods. For instance, the rotation period of Neptune was determined initially by observations of differentially rotating cloud tops but then was redefined after detection of its radio emission (Lecacheux et al. 1993).

Planetary Satellites: In addition to being modulated by its rotation, Jupiter's radio emission is affected by the presence of its satellite Io, and more weakly by Callisto and Ganymede. As the Jovian magnetic field sweeps over a moon, a potential is established across the moon by its $\mathbf{v} \times \mathbf{B}$ motion in the Jovian magnetic field. This potential drives currents along the magnetic field lines, connecting the moon to the Jovian polar regions, where the currents modulate the radio emission. Modulations of planetary radio emission may thus reveal the presence of a satellite.

Atmospheric retention: A common and simple means of estimating whether a planet can retain its atmosphere is to compare the thermal velocity of atmospheric molecules with the planet's escape velocity. If the thermal velocity is a substantial fraction of the escape velocity, the planet will lose its atmosphere. For a planet immersed in a stellar wind, *nonthermal* atmospheric loss mechanisms can be important (Shizgal & Arkos 1996), as the typical stellar wind particle has a *supra-thermal* velocity. If directly exposed to a stellar wind, a planet's atmosphere can erode more quickly. Based on Mars Global Surveyor observations, this erosion process is thought to have been important for Mars' atmosphere and oceans (Lundin et al. 2001; Crider et al. 2005).

Habitability: A magnetic field may determine the habitability of a planet by deflecting cosmic rays or stellar wind particles, (e.g., Griessmeier et al. 2005b). In addition to its effect on the atmosphere, if the cosmic ray flux at the surface of an otherwise habitable planet is too large, it could cause cellular damage or frustrate the origin of life altogether.

4 Science Requirements and Key Advances for the Next Decade

The exploitation of extrasolar planetary magnetospheric emissions will require sensitive, low radio frequency observations.

Table 1 summarizes key requirements for detecting and exploiting extrasolar planetary magnetospheric emissions. In the rest of this section we motivate these requirements.

Table 1. Scientific Requirements for Extrasolar Planetary Magnetospheric Observations

Parameter	Value	Comment
Frequency (Wavelength)	<~ 100 MHz (>~ 3 m)	Determined by planetary magnetic field; Brown dwarf observations suggest higher frequencies possible.
Sensitivity	< 25 mJy	Extrapolations from solar system relations; Constrained by existing observations.

The characteristic frequency above which magnetospheric emissions are no longer generated is determined by the strength of the planet's magnetic field; Jupiter's magnetic field strength, at the cloud tops, is approximately 4 G, leading to a characteristic (cutoff) frequency of approximately 40 MHz. Allowing for the known extrasolar planets more massive than Jupiter, observations below 100 MHz (> 3 m) likely will be required.

The sensitivity requirement is motivated both by extrapolations of magnetospheric emissions from solar system planets as well as current observational limits (§4.1). Extrapolations from solar system planets suggest that at lower frequencies, magnetospheric emissions are stronger so that a somewhat less stringent requirement suffices. Also, because the emissions are driven by the stellar wind, planets closer to their host stars are likely to be brighter.

The performance of a low frequency radio telescope is determined by a combination of design parameters (notably its collecting area), the calibration of the telescope, and the identification and excision of radio frequency interference (RFI).

For low radio frequency observations of extrasolar planets, improved sensitivity generally requires increasing the effective collecting area of the telescope A_{eff}. The theoretical sensitivity of a radio telescope, as given by the radiometer equation, is determined by A_{eff}, its system temperature T_{sys}, the observation bandwidth, and integration time. The magnetospheric emission from solar system planets is broadband, and modern radio astronomical systems process bandwidths that are a large fraction of the observing frequency. Below 100 MHz, the Galactic nonthermal emission makes a significant contribution to the system temperature T_{sys}. Finally, magnetospheric emission from solar system planets can be "bursty," so that long integrations may not be as useful as for traditional observations, because a long integration may average together times when the planet is emitting with times when it is not.

Phase distortions induced by the Earth's ionosphere also impact the sensitivity of low frequency radio telescopes. In the past decade, new calibration schemes have been developed for the ionospheric phase corruptions at radio wavelengths, often bearing a conceptual similarity to adaptive optics techniques for tropospheric phase corruptions at visible wavelengths. Modern low frequency radio telescopes plan both to exploit these new algorithms as well as develop new algorithms in order to reach the theoretical noise limit implied by the radiometer equation.

The final impact on performance is RFI produced by civil or military transmitters operating at similar frequencies. These transmitters are often orders of magnitude stronger than the desired signal. Current RFI identification and excision techniques are relatively crude. More sophisticated algorithms are among the key development aspects for the next generation of radio telescopes.

4.1 Current Status

Magnetospheric emissions from the solar system planets motivated searches for analogous emissions prior to the discovery of extrasolar planets (Yantis et al. 1977; Winglee et al. 1986). Two significant changes in the past decade have been the discovery of extrasolar planets and the development of low frequency telescopes with an order of magnitude more sensitivity than historical telescopes.

There have been multiple searches toward a number of the known extrasolar planets, with many of these searches still in progress (***Table 2***). The premier instruments are the Very Large Array (VLA) with its 74 MHz system, the Giant Metrewave Radio Telescope (GMRT) with its 150 MHz system,

and the Ukranian T-shaped Radiotelescope (UTR-2), which observes at 8–31 MHz. The UTR-2 observes at low enough frequencies that it can detect Jovian emissions, and the 74-MHz VLA observes at a frequency that is within a factor of two of the highest frequency Jovian emissions (40 MHz).

Table 2. Limits on Extrasolar Planetary Magnetosphere Emission[2]

Frequency	Limit	Telescope	Reference
150 MHz	0.3–2 mJy	GMRT	Hallinan et al. 2009; Winterhalter et al. 2009
74 MHz	135–300 mJy	VLA	Lazio & Farrell 2008
25 MHz	100–1600 mJy[3]	UTR-2	Zarka 2007

The most effort has been directed toward the planet orbiting τ Boo. This planet's luminosity is $L < 10^{23}$ erg s^{-1}, unless its radiation is highly beamed into a solid angle $\Omega << 1$ sr, much smaller than that of any of the solar system planets. This luminosity limit is lower than some, but not all, recent predictions. Although higher sensitivity observations are likely required, the non-detection may also be hinting that the magnetic fields, and internal compositions, of extrasolar planets are as varied as the planets themselves.

4.2 Research and Analysis Advances for the Next Decade

There are a series of algorithmic and theoretical developments that could lead to the detection and exploitation of extrasolar planetary magnetospheric emissions with existing telescopes:

- **Key algorithm development** includes more sophisticated RFI excision and time-frequency processing of the signals, particularly, pattern-matching algorithms that could use Jovian radio emission as a template for higher significance detections.
- **Continued polarimetric observations of host stars** to characterize phenomena related to star-planet interactions (§1.2). These observations will allow fundamental planetary characteristics to be extracted, particularly when combined with modeling efforts (e.g., Lanza 2008; Kitiashvili & Gusev 2008).
- **Development of algorithms that may be deployed in firmware** to carry out active adaptive nulling of bright sources in antenna primary beams. Such techniques are currently computationally prohibitively expensive, but may be feasible with FPGA/GPU technologies.
- **Support for coordinated multi-wavelength observations with space missions**—some preliminary work has already been done with COROT and GMRT observations – similar arrangements may prove fruitful, particularly MW coordinated observations with Kepler.

4.3 New Telescopes in the Next Decade

In construction are the Long Wavelength Array (LWA, New Mexico) and the Low Frequency Array (LOFAR, the Netherlands). The LWA will operate in the 20–80 MHz band; LOFAR will operate in the 30–80 MHz and 110–240 MHz bands. Both instruments cover the frequency range expected for emission from Jovian-mass to several Jovian mass planets, with sensitivities expected to be 1–2 orders of magnitude better than existing facilities. LOFAR's initial operational phase is anticipated to be late 2009, with the LWA following within a few years.

Looking to the latter half of the next decade and beyond, the Square Kilometre Array (SKA) is a next-generation telescope in the design and development phase that is expected to operate above 70 MHz. Its design goals are such that it should be easily capable of detecting the radio emissions from the most massive extrasolar planets. However, a significant constraint to all ground-based telescopes

[2] We list telescopes rather than planets targeted, because multiple searches have been conducted, including the same planet at more than one telescope.

[3] The UTR-2 data acquisition system has been upgraded. The quoted limits are from the old system, as the analysis of current extrasolar planet searches is still in progress.

is the Earth's ionosphere, which is opaque below about 10 MHz. The Earth's magnetosphere emits auroral kilometric radiation (AKR) below 1 MHz. Thus, the detection of AKR from extrasolar terrestrial-mass planets—and assessments of their habitability—can only be accomplished from space. The most promising location for a telescope designed to detect AKR from extrasolar terrestrial-mass planets is the far side of the Moon, as it would always be shielded from the Earth.

References

Bastian, T. S., Dulk, G. A., & Leblanc, Y. 2000, ApJ, 545, 1058

Crider, D., Espley, J., Brain, D. A., et al. 2005, J. Geophys. Res., 110 (A9), 9

Farrell, W. M., Desch, M. D., & Zarka, P. 1999, J. Geophys. Res., 104, 14025

George, S. J., & Stevens, I. R. 2007, MNRAS, 382, 455

Griessmeier, J.-M., Zarka, P., & Spreeuw, H. 2007, A&A, 475, 359

Griessmeier, J.-M., Preusse, S., Khodachenko, M., et al. 2007, Planetary & Space Sci., 55, 618

Griessmeier, J.-M., Motschmann, U., Mann, G., & Rucker, H. O. 2005a, A&A, 437, 717

Griessmeier, J.-M., Stadelmann, A., Motschmann, U., et al. 2005b, Astrobiology, 5, 587

Hallinan, G., et al. 2009, in preparation

Kitiashvili, I. N., & Gusev, A, 2008, Celestial Mech. & Dynamical Astron., 100, 121

Lanza, A. F. 2008, A&A, 487, 1163

Lazio, T. J. W., & Farrell, W. M. 2008, ApJ, 668, 1182

Lazio, T. J. W., Farrell, W. M., et al. 2004, ApJ, 612, 511

Lecacheux, A., Zarka, P., Desch, M. D., & Evans, D. R. 1993, Geophys. Res. Lett., 20, 2711

Lundin, R., Barabash, S., Andersson, H., et al. 2004, Science, 305, 1933

Millon, M. A., & Goertz, C. K. 1988, Geophys. Res. Lett., 15, 111

Shkolnik, E., Bohlender, D. A., Walker, G. A. H., & Collier Cameron, A. 2008, ApJ, 676, 628

Shkolnik, E., Walker, G. A. H., Bohlender, D. A., Gu, P.-G., & Kuerster, M. 2005, ApJ, 622, 1075

Shizgal, B. D., & Arkos, G. G. 1996, Rev. Geophys., 34, 483

Stevens, I. R. 2005, MNRAS, 356, 1053

Svensmark, H. 2000, Space Sci. Rev., 93, 155

Walker, G.A.H., Croll, B., Matthews, J., et al. 2008, A&A, 482, 691

Winglee, R. M., Dulk, G. A., & Bastian, T. S. 1986, ApJ, 309, L59

Winterhalter, D., et al. 2009, in preparation

Wood, B. E., Müller, H.-R., Zank, G. P., Linsky, J. L., & Redfield, S. 2005, ApJ, 628, L143

Wood, B. E., Müller, H.-R., Zank, G. P., & Linsky, J. L. 2002, ApJ, 574, 412

Yantis, W. F., Sullivan, W. T., III, & Erickson, W. C. 1977, Bull. Amer. Astron. Soc., 9, 453

Zarka, P. 2007, Planet. Space Sci., 55, 598

Zarka, P., Treumann, R. A., Ryabov, B. P., & Ryabov, V. B. 2001, Astrophys. & Space Sci., 277, 293

Zarka, P., Queinnec, J., Ryabov, B. P., et al. 1997, in Planetary Radio Emissions IV, eds. H. O. Rucker et al., p. 101

Solar Wind and Suprathermal Ion Composition Measurements: An Essential Element of Current and Future Space Missions

A white paper for the 2010 Solar and Space Physics Decadal Survey

Principal Author: Susan Lepri (The University of Michigan)
Co-Authors: Jason Gilbert, Enrico Landi (The University of Michigan), Harald Kucharak (The University of New Hampshire), Mihir Desai (Southwest Research Institute), Devrie Intrilligator (Space Plasma Laboratory, Carmel Research Center)

1. Introduction

The first direct observations of the solar wind in the 1960s revealed the presence of protons and enabled the estimation of the particle flux and bulk properties (e.g. see review by Hundhausen et al., 1968). At the time, the solar wind was known to consist of mostly H+ and ^{4}He++ (and electrons), traveling along at the same bulk velocity. In 1968, observations from the electrostatic plasma analyzer on the Vela 3 satellite confirmed the existence of ions heavier than helium in the solar wind (Bame et al., 1968). Although these measurements were rudimentary, as they could not unambiguously resolve many ion species we now know exist today, they provided the foundation for future measurements.

Heavy ion composition measurements have provided not only a sensitive probe of coronal temperatures (e.g., Hundhausen, 1968; von Steiger et al., 1997; Zurbuchen et al., 1999), but also a measure of elemental composition of the solar wind; the corona; and pickup ions from planets, comets, interplanetary dust, and the local interstellar medium. Heavy ion measurements allow in-situ observations to be traced back to the original source as they reflect the environment in which they were formed. Thus, composition serves as a very sensitive tracer for the origin of heliospheric plasma. Highly resolved composition measurements have only been available for the last 20 years from solar wind experiments including the Solar Wind Ion Composition Spectrometers (SWICS). ACE SWICS has spent 13 years in the ecliptic, and the now decommissioned Ulysses SWICS has covered a range of radii, latitudes and distances over 19 years. Only recently have we gained multipoint measurements with STEREO PLASTIC. These measurements have resulted in transformative discoveries in the areas of solar and heliospheric physics as well as astrophysics.

Some of the major breakthrough results obtained by the study of heavy ion distribution functions, ionic charge states, and elemental abundances include the following:

1. The distinctly different origin of fast and slow wind, with an intriguing link between corona and chromosphere (e.g. Geiss et al. 1995).
2. The ubiquitous presence of suprathermal tails with a gamma = -5 power law spectrum indicating a common acceleration mechanism throughout the heliosphere (e.g., Gloeckler et al. 2003).
3. The direct observation of interstellar pickup ions, constraining the composition of the local interstellar medium (e.g., Gloeckler et al. 1998).
4. The still ill understood slow solar wind fills a larger portion of the heliosphere

than previously thought and therefore is key in the development of accurate solar and heliospheric models (e.g., Zhao, Zurbuchen and Fisk 2009).
5. Inner Source pickup ions whose origin is still hotly debated (e.g., Gloecker, Fisk and Geiss, 2010)

Some of these findings can be directly compared to remote sensing observations of the source regions of the solar wind and of coronal mass ejections and facilitate major advances in our understanding of the solar particle input into the interplanetary medium. This new understanding will boost our ability to predict the solar radiative and particle output and the effects on the heliosphere, the Earth's atmosphere and space based human activities. ***Knowledge of the heavy ion composition in the solar wind is essential for seeking answers to the following major science questions facing the solar and heliospheric community today:***

1. What is the role of reconnection in the heliosphere?
2. What governs acceleration processes in the heliosphere and beyond?
3. How can we improve our space weather prediction capabilities that affect conditions on Earth?

2. The Role of Heavy Ion Composition

Reconnection

Reconnection is a phenomenon that is at work all over the heliosphere, in the solar corona, in the Earth's magnetosphere, and at the outer boundary of the heliosphere. There are many science topics that have their roots in the problem of reconnection. One of these relevant topics is the origin of the slow solar wind. One of the most profound discoveries the past 15 years was that the solar wind is bimodal, and that its two components have very disparate chemical composition indicating different source regions on the Sun. The speed of the solar wind was found to be inversely correlated with the coronal electron temperature as determined from in-situ solar wind charge states (Geiss et al. 1995). As charge states provide a measure of the electron temperature in the region in which they were frozen-in, charge states in the slow wind tend to be higher than they are in the fast wind, indicating an origin on warmer, confined structures in the corona, like large coronal loops observed outside of coronal hole regions. Fast wind exhibits charge states reflecting the conditions in cooler coronal hole regions and of much smaller cool coronal loops. While fast wind is known to originate in coronal holes (although the precise location is still not well constrained), the source of the slow solar wind is still unknown. The slow solar wind has been found to contribute significantly to the amount of plasma in the heliosphere and therefore plays a vital role in heliospheric physics. The main question is how does material trapped on large coronal loops escape to become part of the solar wind. The most likely way for this to occur is via reconnection. The exact nature of this escape, however, is not well understood.

Solar wind ion composition measurements will play an important role in identifying the sources of the slow wind and in characterizing the small scale structures and features in the fast wind that relate to coronal features. Only solar wind composition measurements in concert with in-situ observations of the bulk plasma on micro and macro scales, the heliospheric magnetic field, UV and other remote observations of the

Sun and other observables will enable identification of the source of the slow solar wind and the nature of the boundary between fast and slow wind.

Fully understanding the global structure of the heliosphere requires knowledge of the connection between coronal magnetic fields and in-situ observed heliospheric magnetic field. Determining what controls and constrains the amount and evolution of the magnetic flux in the heliosphere is of crucial importance. Understanding the issues regarding the nature of the open heliospheric flux requires knowledge of the origin of the slow solar wind as the heliospheric current sheet is known to be embedded in the slow solar wind; the structure in the magnetic field is related to the generation of the slow wind. The evolution of the solar corona drives the heliospheric magnetic field and is reflected in changes in solar wind composition. In order to constrain and understand the global structure of the heliosphere, solar wind composition is vital.

Another outstanding topic that involves the process of reconnection is the initiation and evolution of Coronal Mass Ejections (CMEs). A third type of solar wind observed inside of coronal mass ejections and shows an even different composition linked to eruptive and intensely heated or even cooled regions on the Sun. Interplanetary CMEs (ICMEs) very often exhibit unusually high ionization states for heavy ions, indicating that they are heated to high temperatures through reconnection processes like those at work in solar flares (Gruesbeck et al. 2010, Lepri et al. 2001, Lepri and Zurbuchen 2004 and references therein). The eruption process of the CME close to the Sun supplies the plasma with hot, fast electrons, which increase the coronal electron temperatures and therefore increase the ionization states of heavy ions. These ions then expand rapidly and are not able to cool down to normal solar wind ionization levels before they freeze-in. The exact nature of the initiation and eruptive release and subsequent removal of the CME from the corona is not well understood. There are a variety of competing models and theories for CME genesis and evolution and solar wind ion composition measurements will play a key role in constraining these theories and models.

Future work may be able to link heliospheric magnetic topology with solar wind composition, which would provide a measure of the magnetic connectivity that is independent of scattering processes in the heliosphere. These measurements would enable a detailed study of the contributions of ICMEs and heat flux dropout to the heliospheric magnetic flux budget.

Delineation of solar wind type by only solar wind speed will not allow the separation of ICMEs from fast wind, or accelerated slow wind from decelerated fast wind, nor does it get at the root of the generation of the wind in the corona. In order to trace in-situ observations of solar wind to their associated sources and understand the heating and acceleration mechanisms, knowledge of the thermal environment in which the wind was generated is vital. Heavy ion composition is the most important observable that can link in-situ plasma conditions and acceleration mechanisms to those in the source region.

Particle Acceleration

The chemical composition of the Sun is considered a reference for astrophysics, and it is used to study a great variety of astrophysical objects. The importance of measuring it accurately cannot be overstated. Just as a few examples, element abundances

influence the structure of stellar interiors by determining the opacity of stellar plasmas - changes in the abundance of a single element can alter the sound speed and density profiles of the star and the position of the base of its convection zone. Abundances determine the radiative output of stellar coronae, and thus influence both their energy equation as well as the radiative environment of the interplanetary space that surrounds them and the planets that it hosts. Abundances can be used as tracers for nucleosynthesis processes in astrophysical plasmas; they are also very important for the evolution of galaxies. The firm prediction of the primordial He abundance has been considered one of the outstanding successes of big bang cosmology.

The Sun provides us with a unique opportunity to measure element abundances because composition data can be obtained using a variety of independent methods, almost all of which are unavailable to astrophysics. These methods range from spectroscopy (available for other astrophysical processes) to in-situ measurements, meteorite determinations, and lunar and planetary samples. Thus, not only can the solar composition be determined with accuracy unavailable to other astrophysical objects, but measurements of abundance anomalies or element fractionation obtained with different methods or in different plasmas provide us with an opportunity to investigate physical processes that do not emit radiation, and thus are not accessible in other astrophysical objects. In order to be able to constrain these measures of the chemical composition, we need to better understand how different ions and elements are accelerated out of the solar corona.

Particle acceleration occurs in a wide range of environments throughout the heliosphere, from flares on the Sun, interplanetary shocks, CIRs, planetary bow shocks, out to the termination shock. The exact mechanisms for particle acceleration are debated and therefore there remain many related unresolved science topics where solar wind composition can play a key role.

In addition to reflecting different electron temperatures in the low corona, heavy ions can also tell us about the acceleration mechanisms that eject them out of the low corona into interplanetary space. The elemental composition of heavy ions is different in the fast and slow solar wind, with the slow solar wind being enhanced in elements with low first ionization potential (FIP) (Geiss et al. 1995). The differences in the elemental abundances indicate a different origin, fractionation and/or acceleration mechanism for the two types of solar wind. The fast and slow wind are turbulent, with the degree of fluctuations varying between the two, potentially revealing the nature of the heating and acceleration differences between the types of solar wind.

One of the more striking findings about distribution functions of heavy ions observed on ACE and Ulysses is the presence of ubiquitous suprathermal tails, almost always with a spectral index of -5 in velocity space (Gloeckler 2003, Fisk and Gloeckler 2006, 2008). This striking finding is observed from 1AU and within the orbit of Jupiter with Ulysses and ACE SWICS up to 100 keV.e. Also, this spectral shape has been verified at the heliosheath from Voyager observations (Stone et al. 2005). Inner heliosphere (within 1AU) measurements will help us determine where this acceleration is occurring and pin down the process responsible for creating this unique shape. It may be that this spectrum is common throughout the universe and therefore finding the mechanism responsible for it is vital to many fields of astrophysics. The composition and spectrum of suprathermal tails play an important role in determining the seed population

for energetic particles. This is especially important in the inner heliosphere where energetic particles are accelerated.

While ACE and Ulysses SWICS provide 1D distribution functions, STEREO PLASTIC and future instruments will provide 3D distribution functions, enabling anisotropies to be quantified. Observations of heavy ion distribution functions also allow the mass and charge dependence of acceleration mechanisms in the heliosphere and in the corona to be examined.

Solar wind experiments have detected the presence of pickup ions from a variety of sources both external and internal to the heliosphere since 1985 (Mobius et al 1985). These singly charged ions start off as neutral particles and become ionized either via photo-ionization or charge exchange. Once ionized, they are picked up by the solar wind and flow away from the Sun with the rest of the solar wind. Beyond 1 AU, these pickup ions become more and more important until they begin to make up the large fraction of ions in the outer heliosphere. Pickup ions contain information about their source, revealing the composition of the local interstellar medium, the inner source, comets, planetary atmosphere, and therefore information on how the heliosphere and universe are evolving. These ions actually dominate the solar wind pressure in the outer heliosphere. As they have low q/m ratios, they also tell us about the behavior of low rigidity ions. They provide information on the composition and source of anomalous cosmic rays (ACRs).

Prediction

The solar wind and ICMEs determine the interplanetary space environment and have direct impacts on the Earth's atmosphere. Hence, the ability to predict ICME occurrence and solar wind characteristics is key to ensuring the safety of sophisticated ground systems affected by geomagnetic storms, satellites and their components, and human and robotic exploration of space. Improved predictions of space weather will be realized when there is a more complete understanding of processes like reconnection and particle acceleration achieved utilizing solar wind composition measurements along with other observations to constrain models for prediction and propagation of space weather events.

3. Recommendations

What next

The SWICS instruments on ACE and Ulysses and STEREO PLASTIC have allowed us to take the first steps towards exploring the Sun and the solar wind with a new and very powerful tool: heavy ion composition. However, these instruments had several limitations that must be overcome to reach further major breakthrough discoveries. Future missions should include composition measurements that provide the following:

1) **Higher resolution measurements at a range of solar radii:** In order to resolve small scale structure of the solar wind, details of acceleration and generation mechanisms, tracers of interchange reconnection processes, interfaces at CIRs, and other variability, higher time resolution measurements are necessary closer to the Sun. ACE SWICS and Ulysses SWICS measurements provide(d) reliable measurements of heavy ions in the solar wind at ~ hr cadence. STEREO PLASTIC has higher time resolution but is also at 1AU. In order to resolve

substructures and interface regions, ~ min resolution is needed inside the orbit of Earth. It is vital also to have high mass, mass per charge and energy resolution.

2) **High latitude observations:** While Ulysses was able to obtain a good deal of high latitude observations, its mission ended in what turns out to be one of the most interesting periods in the current space age. We observed an unusual behavior of solar wind composition that was not predicted during the unusually long solar minimum. The ability to observe and characterize polar coronal holes and the high latitude boundary regions between fast and slow solar wind was lost when Ulysses was decommissioned. A mission like Solar Polar Orbiter, with a heavy ion composition instrument as part of the payload, would facilitate breakthrough measurements of the global heliosphere.

3) **Long term observations:** We have yet to observe solar wind composition for an entire solar cycle (22 years). We had 20 continuous years of observation at a range of distances and latitudes with Ulysses, however, it was decommissioned before the solar cycle completed. It is impossible to gain a comprehensive understanding of long term evolution of the Sun, the heliosphere, and transient phenomena by observing less than one complete cycle of a periodic behavior. Hence current missions like ACE and STEREO should continue to be supported to develop a historic record of solar wind composition measurements.

A composition instrument that provides high mass, energy, mass per charge and time resolution would likely cost in the range $15-20M $FY10. This instrument would likely have a mass less than 8 kg. This instrument would consist of an energy per charge analyzer paired with a time of flight telescope with post acceleration and energy sensing solid state detectors. To cover the solar wind and sample the suprathermal energy range, the sensor should observe in the energy range from 0.5 – 100 keV/e.

The charge to the community:

Future heliospheric missions, including Solar Orbiter and Solar Probe Plus, must make composition measurements. There are several missions in the pipeline that will attain some of these improvements. These missions must be supported in full, especially since it is risky to assume that ACE and STEREO will be fully funded in 2018.

- The Heavy Ion Sensor on the Solar Orbiter must be fully supported by the community. Solar Orbiter, with its complimentary suite of in-situ instruments which cover a broad range of energies from the solar wind up to energetic particles, waves instrument and magnetometer combined with remote-sensing instruments that provide observations of the photospheric magnetic field, spectroscopic observations from the photosphere to the corona over a large part of the electromagnetic spectrum and a heliospheric imager, is poised to enable major advances in the topics listed above. Solar orbiter will go in to 0.3 AU and extend eventually to +/- 30 degrees in latitude and HIS will provide unprecedented high resolution and high time cadence measurements of the solar wind and suprathermal particles in the inner heliosphere. In addition, the unique opportunity afforded by Solar Orbiter to co-rotate with the Sun will allow fundamental new advancements in the ability to discern the composition specific plasma sources (e.g., loops of varying sizes and states of evolution, the streamer belt, coronal holes, quiet Sun).

- Solar Probe must have a solar wind composition spectrometer as part of the payload with capabilities to observe suprathermal tails of major ions. It is a tragedy that the currently planned mission that will go in to 9.5 Rs, where much of the acceleration of the solar wind and suprathermal particles is likely to occur, and where much of the interplanetary dust is located, will not observe composition and distribution functions of heavy ions. The lack of composition measurements on Solar Probe will prevent our ability to untangle variability due to solar wind acceleration from variability caused by traversing different solar wind sources. The uncertainty concerning solar wind sources fundamentally undermines the central objective of Solar Probe to discover once and for all how solar wind is accelerated. Many have argued that a Solar Probe mission without these easily implemented solar wind and suprathermal tail composition measurements is simply not worth the considerable expenditures for this mission.

Equally as important to the support of new missions, is the support of the NASA Research and Development Program, Instrument Development Program, and support for continuation of current missions, current and future data analysis, model, and theory efforts in heliospheric physics. In particular the support of data analysis is vital to making the most of our observations from current and past. Individual grants of up to $1-2M over 3-5 years would support these activities.

4. References

Bame, S. J., A. J. Hundhausen, J. R. Asbridge, and I. B. 1968, *Astronomical Journal*, *73*, 55

Fisk, L. A., & Gloeckler, G. 2006, ApJ, 640, L79

Fisk, L. A., & Gloeckler, G. 2008, ApJ, 686, 1466

Geiss, J., Gloeckler, G., & von Steiger, R. 1995, Space Sci. Rev., 72, 49

Gloeckler, G. 2003, Solar Wind Ten, 679, 583

Gloeckler, G., et al. 1998, Space Sci. Rev., 86, 497

Gloeckler, G., Fisk, L. A., & J. Geiss, 2010, AIP Conf. Proc. 1216, 514, DOI:10.1063/1.3395915

Gruesbeck, J. R., et al., 2010, submitted to the Astrophysical Journal.

Hundhausen, A. J. 1968, Space Sci. Rev., 8, 690

Lepri, S. T., & Zurbuchen, T. H. 2004, Journal of Geophysical Research (Space Physics), 109, 1112

Lepri, S. T., T. H. Zurbuchen, L. A. Fisk, I. G. Richardson, H. V. Cane, and G. Gloeckler, 2001, *J. Geophys. Res.*, *106*, 29, 231

Moebius, E., Hovestadt, D., Klecker, B., Scholer, M., & Gloeckler, G. 1985, Nature, 318, 426

Stone, E. C., Cummings, A. C., McDonald, F. B., Heikkila, B. C., Lal, N., & Webber, W. R. 2005, Science, 309, 2017

von Steiger, R., et al. 2000, J. Geophys. Res., 105, 27217

Von Steiger, R., J. Geiss, G. Gloeckler, 1997, eds. J. R. Jokipii, C. P., Sonett, M. S. Giampapa, The University of Arizona Press, Tucson

Zhao, L., Zurbuchen, T. H., & Fisk, L. A. 2009, Geophys. Res. Lett., 36, 14104

Zurbuchen, T. H., Hefti, S., Fisk, L. A., Gloeckler, G., & von Steiger, R. 1999, Space Sci. Rev., 87, 353

A “swarm” Mission to Study Particle Acceleration at Interplanetary Shocks

G. Li[1, 2], J. Adams[3], M. Christl[3], D. Gallagher[3], G. Zank[1, 2]

[1]Department of Physics, Univ. of Alabama in Huntsville, AL
[2]CSPAR, Univ. of Alabama in Huntsville, Huntsville, AL
[3]NSSTC, Marshall Space Flight Center, Huntsville, AL

To better understand particle acceleration at CME-driven shock and interplanetary traveling shock, there is an urgent need for 1) accurate and high cadence particle measurements that can measure particle energies up to ~800 MeV/nucleon and 2) simultaneous multi-point plasma and magnetic field measurements at an interplanetary shock.

Understanding the acceleration and transport of Solar Energetic Particles (SEPs) in the inner heliosphere is one of the outstanding problems in heliospheric plasma physics. Solar Energetic Particles pose major radiation hazards for spacecraft and astronauts and when penetrating from the cusp region, they can affect the health of airline crews and passengers on polar flights. Over the past several decades, a tremendous amount of data on SEPs has been obtained. It is now generally believed that our Sun accelerates particles to high energies mainly via two processes: solar flares and coronal mass ejections (CMEs). The observed temporal profiles of particles at 1 AU for these two processes are often different, with flare events being more “impulsive” and CME shock events being more “gradual” [Reames, 1995, 1999]. Besides the observed temporal profiles, spectra, abundances, and ionization states of energetic particles in these two classes also differ substantially (see e.g. recent surveys on CME-driven shocks and associated SEP events, [Desai et al. 2003, 2004, Ho et al. 2003], and surveys on ion composition and spectra during SEP events [Tylka et al. 1995, 2005b, Mewaldt et al., 2001, Cohen et al., 2000, 2003]. In this white paper, we address particle acceleration at CME-driven and interplanetary shocks.

The problem and our current understanding: It is generally accepted that the underlying mechanism of particle acceleration at CME-driven and planetary shocks is diffusive shock acceleration, also known as first-order Fermi acceleration [Axford et al., 1977; Bell, 1978a,b; Krymsky et al., 1977; Blandford and Ostriker, 1978]. Here particles scatter back and forth between the upstream and downstream medium due to magnetic turbulence. Upon completing each traversal, a (non-relativistic) particle will gain momentum proportional to m v, where m is the particle mass and $v = |v2 - v1|$ is the background velocity difference between the upstream and downstream. For a quasi-parallel CME-driven shock, the upstream and downstream turbulence is generally assumed to be Alfven waves that are thought to be created by protons streaming away from the shock front [Lee, 1983]. As such, the wave intensity is coupled to the anisotropic part of the particle distribution function, *f*. Therefore, a self-consistent model of particle acceleration and transport requires evaluating the wave intensity simultaneously with energetic particles. Earlier theoretical work calculating the power of the upstream wave intensity, based on a quasi-linear theory, can be found in Lee [1983]. This work was later extended by Gordon et al. [1999] and the effects of proton excited waves on the transport of various ions have been discussed in Ng et al. [1999] and Ng et al. [2003]. Attempts of modeling particle

acceleration at a propagating quasi-parallel CME-driven shock has been put forth, for example, by Zank et al. [2000], Rice et al. [2003], Li et al. [2003, 2005]. In these works, the power of upstream excited Alfven waves is obtained self-consistently with the particle energy spectra a the shock, and the power of upstream waves (in resonance with ~ 1MeV – 800 MeV protons) can be over 100 times stronger than the corresponding pristine solar wind MHD turbulence power. Therefore, observationally detecting these upstream waves allow a way to examine the validity of the Lee [1983] theory. Note that these upstream Alfven waves will be largely suppressed at a quasi-perpendicular shock. This means particle acceleration at a quasi-perpendicular shock cannot be explained by the Lee [1983] theory. Observationally, it has been shown (e.g. Zank et al [2006]) that particle acceleration does occur at quasi-perpendicular shocks where no enhancement of upstream Alfven waves was identified. Furthermore, the acceleration efficiency at the quasi-perpendicular shock studied in Zank et al [2006] was comparable to quasi-parallel shocks. Clearly, a different mechanism is in operation in regulating the diffusion coefficient at a quasi-perpendicular shock.

The traditional theory of $K_{\perp}$ is the quasi-linear theory, in which particle perpendicular diffusion is primarily described by the Field Line Random Walk model (FLRW) [Jokipii, 1966]. However, it has been shown that FLRW did not agree with numerical simulations when particle energies are low. Matthaeus et al. [2003], assuming that magnetic field lines separate diffusively, introduced the Non-Linear-Guiding-Center (NLGC) model which appears to be consistent with numerical simulations in both the high- and low-energy particle regime. Attempts to use NLGC theory to model particle acceleration at a perpendicular shock have been first discussed by Zank et al. [2004, 2006], who introduced an explicit model for solar wind turbulence (2D + slab). They concluded that fast particle acceleration can occur at perpendicular shock and the injection energy at a perpendicular shock is much higher than at a quasi-parallel shock. This implies that the seed population at a quasi-perpendicular shock is of supra-thermal particles, which can have different composition from a quasi-parallel shock. Indeed, based on compositional analysis, Tylka et al. [2005] argued that quasi-perpendicular shocks are more efficient accelerators than quasi-parallel shocks.

What is clear from the above is that, besides the in-situ observation of enhanced Alfven waves correlating with energetic particles at a quasi-parallel shock, the other critical measurement of particle acceleration at CME-driven or interplanetary shocks is shock geometry. To obtain shock geometry, simultaneous multi-point measurement of magnetic field is necessary. It has been known that [Bale et al 1999] interplanetary shock surfaces have ripples. Indeed, there have been pioneering studies of particle acceleration at a rippling Termination Shock (TS) [Li and Zank, 2006, Jokipii and Kota, 2008] in which the idea of “local acceleration” has been put forth. In this scenario, the observed accelerated particle spectrum at one location is an ensemble of particles from multiple “local sources” [Li and Zank 2006]. To verify this proposal, it is critical to understand the micro structures of a propagating interplanetary shock. Recently, using multiple spacecraft including ACE and WIND, Giacalone and Neugebauer [2008] suggested that interplanetary shocks have spatial structures with scales of the correlation length of the solar wind MHD turbulence. At 1 AU, this yields a scale size of 0.01-0.03 AU.

To improve our theoretical understand of particle acceleration at an interplanetary shock and/or CME-driven shock, the following key questions has to be answered by future missions.

1) The coupled behavior of accelerated particle and upstream waves stimulated by streaming protons at quasi-parallel shock.
2) What is the geometry of an interplanetary shock; what are the micro-structures, in particular, ripples, of interplanetary shocks?
3) What is the composition (and perhaps the charge states) of the seed population?

Required and desired measurements: We summarize the required measurements to answer the above key questions include,

1) Time intensity and spectra of energetic particles from ~100 keV to 800 MeV/nucleon.
2) Composition of solar wind material before the shock arrival as well as energetic particles.
3) Magnetic field measurement from 0.001 Hz – 0.5 Hz.
4) Simultaneous measurement of plasma parameters (V_{sw}, P_{sw} and T_{sw}) at multiple points with a separation of 0.01 to 0.05 AU.

Mission concept: The multi-point measurements are crucial to obtain the micro-structure of an interplanetary shock and the measurement of shock geometry and because of this special requirement, our proposed mission is best described by a "swarm" composed of a mother spacecraft surrounded by a swarm of small sub-spacecraft, where all measurements of energetic particles (containing perhaps most of the payload) will be performed on the mother spacecraft only. Each surrounding sub-spacecraft contains a relatively simple payload consisting of a magnetometer and plasma instruments to measure solar wind parameters. These sub-spacecraft should be distributed on a scale that enables a search for possible correlations in the solar wind MHD turbulence, which are unknown, but likely to be found in the range of 0.01 AU to 0.05 AU at 1 AU from the Sun.

The concept for deploying for deploying the "swarm" is based on the CubeSat concept (http://www.cubesatkit.com/). Each CubeSat must be equipped with plasma and magnetic field instrumentation, a small transmitter, a momentum wheel for three-axis stabilization and a solid micro-thruster. The idea is to place the mother spacecraft (hereafter called the mother) in the same orbit around the Sun as the Earth, but leading or trailing the Earth at a distance that is well beyond the Earth's magnetosphere. Once the orbit of the mother has been trimmed to match that of the Earth, the CubeSats will be dispensed in pairs, ejected in opposite directions with the same Δv, one forward along the trajectory of the mother and the other backward. Six months later the forward-ejected (backward ejected) CubeSat will reach aphelion (perihelion). At that point timers, started by ejection from the mother will run out, causing the solid micro-thrusters to fire. The thruster will be directed into (opposite to) the ram direction on the forward-ejected (backward ejected) CubeSat. The resulting thruster burns will place the CubeSats in orbits with the same period at the mother but with different eccentricities so that aphelion of each CubeSat will be 0.025 AU farther from the Sun than the mother and 0.025 AU closer at perihelion. Because the CubeSats have the same orbital period as the mother, the swarm will fly in formation, the CubeSats running ahead of the mother around its orbit as when they are nearer the Sun and falling behind the mother when they are farther from the Sun, thus sampling not only in the radial direction but in the azimuthal direction as well.

Four such pairs must be deployed at different times during the first 4.5 months of the mission in

order to generate a swarm that will always have CubeSats positioned to sample plasma and magnetic field conditions over the spatial range from 0.01 to 0.05 AU.

Mother Spacecraft Instrumentation: The Mother Spacecraft must carry instrumentation to make all four of the measurements described above. The particle energy spectral measurements will require several instruments. The IMPACT suite of seven instruments (see Luhmann et al., 2008) on the STEREO mission covers the energy range from ~100 KeV/nuc up to 89 to 167 MeV/nuc depending on the ion species. To extend the measurements up to 800 MeV/nuc and additional instrument is required such as the Solar Energetic Particle Spectrometer (SEPS), (Christl et al., 2009). The solar wind composition measurements can be made using an instrument like the Plasma and Supra-Thermal Ion Composition instrument on STEREO (Galvin et al., 2008).

The magnetic field measurements could be made with an instrument such as the ACE Magnetic Fields Experiment (Smith et al., 1998). This instrument uses triaxial fluxgate magnetometers to cover the frequency range from 0 to 12 Hz and the field strength range from 4 nT to 65 μT.

The required plasma measurements could be performed by ESA's Solar Wind Plasma Analyzer (SWA) which is under development by Dr. Christopher Owen, Mullard Space Science Laboratory, United Kingdom. The Solar Wind Plasma Analyzer, SWA, consists of a suite of sensors that will measure the density, velocity, and temperature of solar wind ions and electrons, thereby characterizing the solar wind between 0.28 and 1.4 AU from the Sun. In addition to determining the bulk properties of the wind, SWA will provide measurements of solar wind ion composition for key elements (e.g. the C, N, O group and Fe, Si or Mg). A combination of a Faraday cup, a vector spectrometer and a Strahl spectrometer could be used to make the required measurements as was done on the Wind spacecraft (Ogilvie et al., 1995).

CubeSat Instrumentation: Each CubeSat must make three-axis magnetic field measurements. This can be accomplished with one tri-axial fluxgate magnetometer which weighs 225g and consumes 1.4 Watts. Each must also be equipped with a Faraday cup, a vector spectrometer and a Strahl spectrometer or equivalent instrumentation to make the required solar wind measurements.

Estimated Cost: The estimate was prepared by MSFC's Engineering Cost Office. The estimated cost for the mission described in this white paper is shown in the table below:

Interplanetary Shock Mission	**FY10$M**	
Mothership		152.3
STEREO Spaceceraft (Dev)	108.5	
STEREO Spaceceraft (Flight Unit)	38.4	
Plasma Instrument (Dev)	1.0	
Plasma Instrument (Flight Unit)	0.3	
HE Solar Particle Detector (Dev)	2.9	
HE Solar Particle Detector (Flight Unit)	1.1	
CubeSat (Total)		2.3
CubeSat (each)	0.048	
Plasma Instrument (Flight Unit)	0.3	
Spacecraft Integration--PeaPods		5.9
Mission Total (FY10$M)		**160.5**

This cost does not include launch cost or ground operations. This estimate was made using the NASA Air Force Cost Model (NAFCOM). The estimate was based on one of the spacecraft form the STEREO mission. The STEREO cost data was normalized for NAFCOM from The Project Office CADRe (Cost Analysis Data Requirements Document). These are the actual costs for STEREO development and spacecraft bus fabrication. The instrument cost (development and flight unit) were estimated in NAFCOM using the Scientific Instruments Cost Model (SICM) module. The CubeSat costs were obtained from open sources, available on the WWW.

Importance of the proposed Mission: Understanding particle acceleration is identified as one of the four science objectives in the "The New Science of the Sun-Solar System Connection Roadmap for Science and Technology 2005-2035". "*Understand the plasma processes that accelerate and transport particles*" is listed as the second area of the "Open the Frontier Research Focused Area (RFA)".

Interplanetary shocks generate the largest solar energetic particle events, the ones most hazardous to both space crews and electronics in space. Providing predictions of the evolution of solar particle events that are accurate enough to be actionable requires much more accurate modeling of particle acceleration at CME-driven shocks.

The recently funded LWS mission Solar Probe Plus has as one of its primary research goals (taken from "The Sun to the Earth and Beyond Panel Reports 2003"),

> *Identify the acceleration mechanisms and locate the source regions of solar energetic particles, and use these particles as remote probes of physical conditions*.

Our proposed mission – to study particle acceleration at interplanetary shocks via multi-point observation will complement the research goal of the Solar Probe Plus mission.

The mission proposed here will make use of mature technology and is conceptually an extension of existing successful missions such as ACE, WIND, SAMPEX and STEREO. It makes use of the mature CubeSat technology, and is therefore at a high TRL level.

References:

Axford, W. I., Leer, E., & Skadron, G., ICRC, Vol. 11, 1977;
Bale, S.et al, Geophysical Research Letters, 26, 1573-1576, 1999
Bell, A.R., Mon. Not. Roy. Astron. Soc., 182, 147-156, 1978a.
Bell, A.R., Mon. Not. Roy. Astron. Soc., 182, 443-455, 1978b.
Blandford and Ostriker, Astrophys. J. 221, L29 1978.
Cohen C. M. S., et al., AIP Conference Proceedings vol. 528, ed. R. A. Mewaldt, J. R. Jokipii, M. A. Lee, E. Mbius, & T. H. Zurbuchen (New York: AIP), pp55, 2000.
M. J. Christl et al., Proceedings of the 31st ICRC (Lodz) (2009).
Cohen C. M. S. et al., Adv. in Space Res., 32, Issue 12, p.2649-2654, 2003
Desai, M. I., et al., Astrophys. J., vol. 558, 1149, 2003.
Desai, M. I., et al., Astrophys. J., 2004.
Gordon, B., M. Lee, E. Mobius, & K. Trattner, J. Geophys. Res., 104, 28,263, 1999.
Galvin et al., Space Sci Rev (2008) 136: 437–486
Giacalone J. and Neugebauer, , ApJ, 673, 629, 2008.
Ho, G. C., et al., Proceedings of the 28th ICRC Conf., SH2.3, vol. 6/7, pp 3689, 2003.
Jokipii, J. R., Astrophys. J., 146, pp. 480-487, 1966.
Jokipii, J. R. and Kota, J., , AIP Conf. Proc., 1039, pp. 390-396, 2008.,
Krymsky et al., 1977;
Lee, M., J. Geophys. Res., 88, 6109–6119, 1983.
Li, G., G. P. Zank, and W. K. M. Rice, J. Geophys. Res., 108, 1082, 2003.
Li, Gang; Zank, Gary P.; Rice, W. K. M., J. Geophy. Res., 110, A06104.
Li, G. and G. P. Zank, AIP Conference Proceedings, 858, pp. 183-189, 2006.
Luhmann, J.G. et al., in *The STEREO Mission*, C. Russell (ed.), 111,184, DOI: 10.1007/978-0-387-09649-0_6, 2008.
Mewaldt et al., 2001,
Ng, C., D. Reames, and A. Tylka, Geophys. Res. Lett., 26, 2145, 1999.
Ogilvie, K.W. et al., Sp. Sci. Rev., **71**: 77-77, 1995.
Reames, D.V., Adv. Space Res., **15**, 41-51, 1995.
Reames, D.V., Space Sci. Rev., **90**, 413-491.
Rice, W.K.M., Zank, G.P. and Li, G., J. Geophy. Res., 108, pp. SSH 5-1, 2003.
Smith, C.W. et al., Space Science Reviews **86**: 613–632, 1998.
Tylka, A., et al., Astrophys. J., **444**, L109–L113, 1995.
Tylka et al., Astrophysical Journal, 625, 474, 2005.
Zank, G., W. Rice, and C. Wu, J. Geophys. Res., 105, 25,079–25,095, 2000.
G.P. Zank, G. Li, V. Florinski, W.H. Matthaeus, G.M. Webb, J.A. le Roux, J. Geophys. Res. 109, A04107, 2004.
G.P. Zank, . G. Li, V. Florinski, Q. Hu, D. Lario, C.W. Smith, J. Geophys. Res. A06108, 2006.

White Paper for: **The Multi-spacecraft Inner Heliosphere Explorer (HELIX)**

R.P. Lin, S. Bale, S. Krucker, G. Hurford, D. Larson, M. Lampton, J. Luhmann, G. Fisher, (UCBerkeley); B. Dennis, A. Szabo (GSFC): G. Mason (APL); G. Gloeckler (U.Md).

Abstract: We propose a multi-spacecraft mission, HELIX, to study the dramatic, large-scale transient phenomena – Coronal Mass Ejections (CMEs) – that evolve rapidly during their passage from the Sun to 1 AU. The origin of CMEs, their relationship to flares, their evolution and interactions with the ambient solar wind are not well understood, but these processes determine the impact of these disturbances on the Earth's environment. An equally important and not understood phenomenon closely-related to these transient eruptions is the acceleration of Solar Energetic Particles (SEPs) in solar flares or CME-driven shock waves (or coronal processes). Both these large scale phenomena are best studied in the inner heliosphere with widely separated multi-point measurements, in particular, to separate propagation effects from acceleration for SEPs. HELIX combines comprehensive in situ measurements of the plasma, energetic particles, and fields, from multiple spacecraft at heliocentric distances of 0.46 to 0.8 AU, with remote sensing down to the lower corona/ chromosphere, to investigate these fundamental physical processes. HELIX is designed to be a low-cost mission utilizing the Explorer mission philosophy, patterned after the extremely successful 5-spacecraft THEMIS mission for the magnetosphere.

1. Science Objectives. The primary scientific objective of HELIX is to explore the global structure of the near-equatorial inner heliosphere, where the most energetic and dynamic phenomena in the solar system take place. HELIX will address the following key science questions:

- How and where are solar energetic particles (SEPs) accelerated?
- What is the structure and topology of the magnetic field from the corona to the inner heliosphere and how does it evolve?
- What is the origin of CMEs, and how do they evolve in transit?
- What are the conditions in the inner heliosphere that affect Space Weather at the Earth and how can we obtain reliable earlier warning?
- What are the physical processes that accelerate the slow solar wind and how does it evolve in transit?

HELIX utilizes five spacecraft in heliocentric orbits spanning ~0.46 to ~0.8 AU, to provide multi-point in situ and remote measurements of these complex, evolving phenomena. The payload is made up of relatively small and simple in situ and remote sensing instruments (Table1) that take advantage of comparisons between the multiple widely spaced spacecraft to significantly increase their capabilities. HELIX X-ray imaging, together with radio tracking of solar type III bursts, remotely track fast electrons along open magnetic field lines from the footpoint source at the Sun out through the inner heliosphere to **trace the magnetic structure of the solar corona and the heliosphere.** Measurements of the solar wind halo electron anisotropy provide a continuous probe of whether the interplanetary field is connected to the hot corona.

HELIX measurements of solar wind composition, which is determined in the corona, combined with determination of the magnetic structure of the corona, and relevant solar observations, **uniquely identify solar wind sources.**

The five HELIX spacecraft, spaced in radial distance, and separated in longitude, in combination with **Solar Orbiter (perihelion 0.3 AU) and Solar Probe Plus (perihelion 9.5 Rs)**, **will provide the first comprehensive, multi-point observations of the origin, propagation, and interactions of CMEs**, so all aspects, from their origin and evolution, through their generation of energetic particles can be studied.

HELIX will address the critical issues in SEP acceleration. Multiple observations in the inner heliosphere allow the injection of SEPs, especially at low energies (<0.1MeV) to

be separated from propagation and in situ shock-acceleration. Precise injection times determine whether flares or flare wave/shock, coronal processes, or CME shocks are the accelerator. **HELIX will detect solar ~0.5-20 MeV neutrons, and possibly positrons,** as well as gamma-rays produced by the accelerated ions interacting close to the Sun. HELIX's multiple viewpoints provide information on the directivity of these emissions, and on the height distribution of the source in the corona through solar limb occultation.

HELIX will probe solar wind plasma microphysics in the inner heliosphere, using multi-point comprehensive particle and fields, measurements. Wave-particle interactions involving low frequency Alfven waves, whistler and ion acoustic waves, weak double layers, Langmuir and radio waves, are important in the transfer of energy in the solar wind, in particle acceleration, and possibly in the interplanetary electric field.

HELIX will provide real-time inner heliosphere Space Weather observations through a continuously transmitted low-bit-rate beacon mode. This could, in principle, provide much earlier warnings for ICMEs and SEP events than an L1 monitor.

HELIX with Solar Orbiter and Solar Probe Plus together form a constellation of 7 in-situ measurements distributed from ~9.5 Rs to 1 AU. The combination of these observations will tremendously enhance the scientific payoff for HELIX.

2. Mission Overview The HELIX science objectives are achieved with a comprehensive suite of simple instruments – no imagers (Table 1) all with strong heritage from previous missions. HELIX is patterned closely after the extremely successful THEMIS MIDEX mission (total cost <~$200 million in 2002 dollars, including launcher) where for the first time a constellation of five spacecraft were launched into distant Earth orbits on a single launcher. The five HELIX spacecraft (Table 2), utilizing the same (or similar) carrier as THEMIS, will be launched together aboard a single launch vehicle. Although the HELIX spacecraft are more mass than THEMIS, even the cheapest currently available low end Atlas or Delta ELV would provide more than adequate performance to reach Venus. After separation and checkout, each spacecraft is separately targeted for a Venus gravity assist to insert it into its orbit.

Like THEMIS, the HELIX spacecraft (Table 3) have no pointing required for imaging (but they use multiple spacecraft occultation & direction finding to obtain detailed spatial information and directivity with simple instruments), and so they can be spin-stabilized with spin axis perpendicular to the orbit plane. This makes thermal control straightforward with present technology outside of ~0.45 AU, the closest HELIX perihelion.

The spacecraft will be single-string; only three working spacecraft is sufficient to achieve the minimum mission (note that the five THEMIS spacecraft are still operating nominally 4.75 years after launch, with two of them now in lunar orbit).

Data is taken continuously and stored at an average rate of ~1 kbps into a 10 Gbyte on-board memory. About once per week the data is dumped to a 40 m antenna at the Owens Valley Radio Astronomy facility. DSN is used for uplink and backup. The Mission Operations and planning will be done by the experienced and highly efficient THEMIS Operations team, using the UCB MOC/SOC.

The nominal mission is two years, but given the history for Explorer missions such as THEMIS, and the relatively benign environment, we fully expect a much long operating lifetime. As is the case for THEMIS, all of the science data will be made available immediately to the scientific and forecasting community.

3. Summary HELIX will provide outstanding science and lead to an understanding of the inner heliosphere and, in particular, of the large transient events and their acceleration of particles to high energies. Basing HELIX on the THEMIS MIDEX philosophy can hold the cost down to somewhat larger than MIDEX levels, plus the additional launch vehicle costs. HELIX would especially benefit from being operational during the high activity portion of the Solar Orbiter and Solar Probe Plus prime missions.

Table 1. HELIX Instruments

SWI SWE	Solar wind ions, 0.1-20 keV/q Solar wind electrons, 1 eV-5 keV
SET	Suprathermal electrons, 2-100 keV
SWICS	Solar wind & suprathermal ions, composition & charge state, ~ 2-160 keV/q
SIS	Suprathermal Z>~2 ion composition, ~5 keV-2MeV/nucleon
EPT	Energetic electrons, 20-400 keV Energetic protons, 20 keV – 7 MeV
EPS	High energy electrons, 0.4 – 20 MeV Energetic ions, elemental composition, 5- 100 MeV/nuc Neutrons, 0.5 – 20 MeV Positrons, 0.3-1 MeV Gamma-rays, 0.3 – 8 MeV
MAG	Vector magnetic fields
WAVES	Radio & plasma waves, few Hz - 16 MHz
XRI	X-ray imaging, ~3-100 keV

Table 2. Mission Summary

Launch	Chosen to reach Venus and to catch the solar maximum
Venus Encounter	few months later
Launch Vehicle	low end Atlas or Delta

Orbits (Heliocentric)	Perihelion (AU)	Aphelion (AU)	Inclination
Spacecraft 1	0.72	0.79	0.5°
Spacecraft 2	0.59	0.79	10.0°
Spacecraft 3	0.55	0.75	8.6°
Spacecraft 4	0.46	0.74	3.1°
Spacecraft 5 (spare)	0.46	0.74	3.1°

Table 3. HELIX Spacecraft Mass & Power

Subsystem	Mass (kg)	Power (W)
Structure	47.6	
Attitude Control	6.4	6.3 - 18.3
Electrical Power	38.3	8.5
Propulsion	15.1	12.3 - 40.3
C&DH	8.5	30.3 - 38.5
Thermal	3.2	9.0
TT&C	17.8	16.6 - 28.5
Max S/C Bus	136.9	83.0 - 149.4
Max Instruments	24.5	24.6
Dry Observatory	161.4	107.6 - 160.0
Propellant	40.0	N/A
Total Observatory	201.4	107.6 - 160.0
Total 5 Spacecraft	1007	N/A

Solar Eruptive Events (SEE) 2020 Mission Concept

R. P. Lin, A. Caspi, S. Krucker, H. Hudson, G. Hurford, (SSL/UCB)
S. Bandler, S. Christe, J. Davila, B. Dennis, G. Holman, R. Milligan, A. Y. Shih (GSFC)
S. Kahler (AFRL); E. Kontar (Glasgow); M. Wiedenbeck (JPL); J. Cirtain (MSFC)
G. Doschek, G. H. Share, A. Vourlidas (NRL); J. Raymond (SAO); D. M. Smith (UCSC)
M. McConnell (UNH); G. Emslie (WKU)

Abstract Major solar eruptive events (SEEs), consisting of both a large flare and a near simultaneous large fast coronal mass ejection (CME), are the most powerful explosions and also the most powerful and energetic particle accelerators in the solar system, producing solar energetic particles (SEPs) up to tens of GeV for ions and hundreds of MeV for electrons. The intense fluxes of escaping SEPs are a major hazard for humans in space and for spacecraft. Furthermore, the solar plasma ejected at high speed in the fast CME completely restructures the interplanetary medium (IPM) - major SEEs therefore produce the most extreme space weather in geospace, the interplanetary medium, and at other planets. Thus, understanding the flare/CME energy release process(es) and the related particle acceleration processes are major goals in Heliophysics. To make the next major breakthroughs, we propose a new mission concept, SEE 2020, a single spacecraft with a complement of advanced new instruments that focus directly on the coronal energy release and particle acceleration sites, and provide the detailed diagnostics of the magnetic fields, plasmas, mass motions, and energetic particles required to understand the fundamental physical processes involved.

1 Introduction

Solar eruptive events (SEEs) consist of both a large flare and a near simultaneous large, fast, coronal mass ejection (CME). The largest SEEs are the most powerful explosions in the solar system, releasing 10^{32}-10^{33} ergs on time scales as short as tens of minutes. They are also the most powerful and energetic particle accelerators in the solar system, producing ions up to tens of GeV and electrons up to hundreds of MeV. For flares, the accelerated particles often contain 10-50% of the total energy released in the flare, a remarkable efficiency that indicates the particle acceleration and energy release processes are intimately related. Similar processes appear to occur elsewhere in the universe, in stellar flares, magnetars, etc. Escaping solar energetic particles (SEPs) appear to be accelerated by the shock wave driven by the fast CME at altitudes of ~2-40 R_s, with an efficiency of ~10%, about what is required for supernova shock waves to produce galactic cosmic rays. Thus, SEEs are our most accessible laboratory for understanding the fundamental physics of transient energy release and particle acceleration in cosmic magnetized plasmas. The intense fluxes of SEPs are a major hazard for humans in space and for spacecraft. Furthermore, the ~10^{15}–10^{16} g of solar plasma ejected at high speed in a fast CME significantly restructures the interplanetary medium (IPM) - major SEEs therefore produce the most extreme space weather in geospace, the IPM, and at other planets. Thus, understanding the fundamental energy release process(es) in SEEs and the related particle acceleration processes are major goals in Heliophysics.

Our basic working model starts with the buildup of energy in the magnetic field of the solar corona over days to weeks. The free magnetic energy is released explosively through a process that is only poorly understood, but most probably involves magnetic reconnection to enable the magnetic configuration to relax to a lower energy state. The magnetic energy so released appears primarily in the form of the acceleration of particles (both electrons and ions), the acceleration (mass motion) and heating of the bulk material of the CME, and heating of the flare plasmas. Subsequently, further particle acceleration occurs at the CME-driven shock. The detailed partition of the released energy among these various products is shown in Fig. 1 for the SEE on 2002 July 23 (Emslie et al. 2004, 2005). In this case, and in a few other well-observed events, there is an approximate equipartition of energy between the flare and the CME.

Great progress has been made in developing this picture of SEEs, but fundamental science questions remain unanswered.

- How is magnetic energy suddenly released to produce both a flare and a CME?
- How are CMEs accelerated?
- How can flares accelerate electrons and ions so efficiently?
- How are escaping SEPs accelerated?
- How are the flare and the CME plasmas heated?
- How is magnetic energy built up and stored in the corona prior to an SEE?

The answers to these and other questions are needed before a comprehensive physical understanding can be claimed or reliable forecasting made of these events. This will require a coordinated program of theory and modeling of the fundamental physical processes combined with our proposed SEE 2020 space mission and supporting space- and ground-based observations.

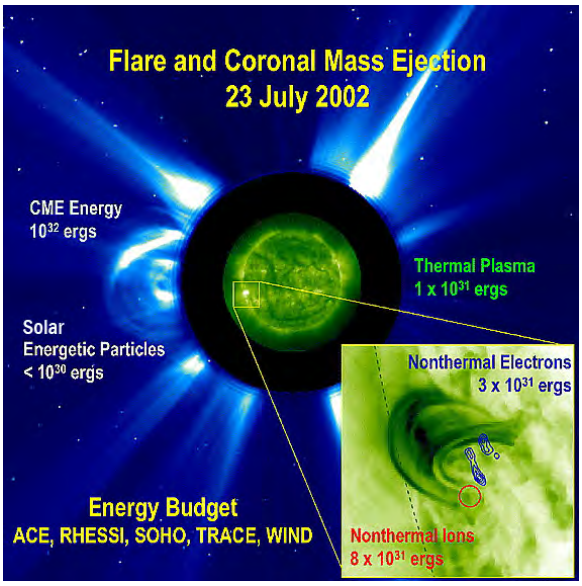


Fig. 1. Images and energy budget of the major components of an SEE (Emslie et al. 2004). This comprehensive picture required input from instruments on 5 individual spacecraft.

Given our current state of knowledge and the development of new instrumentation capable of advanced observations of the most intense SEEs, we believe that the next solar maximum in the 2020 time frame presents an excellent opportunity to make dramatic advances in this field. Since the largest events are of great scientific interest and the most important for space weather, they must be the focus of the effort. Consequently, all instruments must be optimized for observations of the largest M- and X-class events while maintaining high sensitivity, and they must have coordinated and complementary capabilities to maximize the science return from simultaneous observations. The instruments must have the resolutions and dynamic ranges in space, time, and energy/wavelength required to resolve the weakest features of interest in the presence of the strongest components, without saturation. The observations of the largest events will thus provide the greatest information possible while at the same time allowing sensitive and accurate observations of many weaker events.

The basic philosophy of the proposed SEE 2020 research program is to concentrate on the sites of the energy release and particle acceleration in the solar corona, and to characterize the most energetically important components as they evolve in time and space. Observations will be made of these sites before, during, and subsequent to the event to characterize the magnetic fields, the plasma density, temperature, and flow velocities, and plasma wave fields (where possible), as well as the accelerated particles. In addition, the accelerated particles, mass motions, and heated plasma will be observed as close to the release site as possible both in space and in time so that the important processes involved can be accurately inferred with minimal uncertainties due to propagation effects and temporal evolution.

1.1 Location and Nature of the Energy Release Site

Many lines of reasoning point to the corona as the location of the energy release site, and the coronal magnetic field as the source of the energy that powers these SEEs. Relatively weak coronal sources are sometimes detected. When two coronal X-ray sources with opposite energy gradients are observed one above the other, the energy release site can be inferred to lie between them (Fig. 2). In the first such observation, Sui and Holman (2003) found that the energy release site was at an altitude of between ~9 and 23 Mm. With RHESSI's limited dynamic range, weak coronal HXR sources are difficult to detect in the presence of the much brighter footpoint sources, but systematic studies of flares whose footpoints are occulted show that coronal HXR emission is commonly present. Yohkoh observations (e.g., the Masuda

flare) and more recent RHESSI observations of an occulted flare (Krucker et al. 2010) suggest that the particle acceleration/energy release region is located about 5-10 Mm above the flare loops.

The acceleration vs. time profiles for fast CMEs up to ~4 R_s are found to be closely synchronized with energy release of the associated large flares, as measured by the hard X-ray flux vs. time profile (Qui et al. 2004; Temmer et al. 2008, 2010). This is consistent with magnetic reconnection occurring in the current sheet behind the CME, and confirms that the energy release processes for CMEs and flares are intimately related.

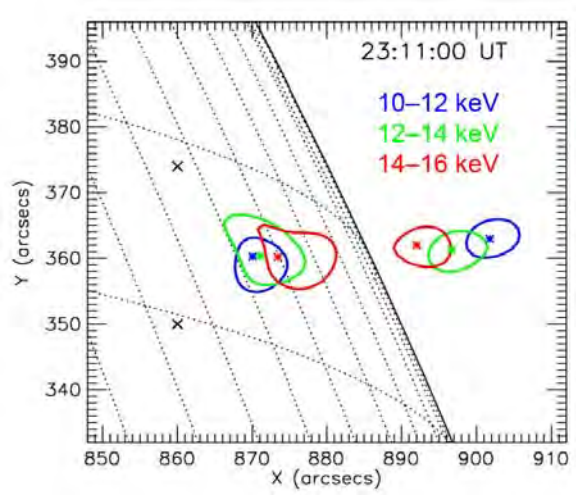


Fig. 2. RHESSI observation of two coronal thermal X-ray sources that are believed to bracket the flare energy release site. Each HXR footpoint is indicated with an ×.

2 SEE 2020 Required Measurements

2.1 Particle Acceleration

The large fraction of the released magnetic energy that goes into the acceleration of particles, both electrons and ions – perhaps as high as 50% – makes the determination of the physical processes by which this is achieved central to any attempt at understanding SEEs. Accelerated particles that interact at the Sun are most effectively explored by observing the X-rays and gamma-rays that they produce. Detection of outwardly directed particles has been limited to *in situ* observations, but new techniques may now be possible to image those particles using energetic neutral atoms (ENAs) resulting from charge exchange. Such images could provide spatial and temporal connections to the shock regions imaged by coronagraphs.

2.1.1 Flare-accelerated Electrons

The physical mechanisms that accelerate electrons and that affect their distribution as they propagate away from the acceleration site are still unknown. In large flares, the required rate of acceleration of electrons above 20 keV can exceed 10^{37} electrons s^{-1}. Such a large flux of charged particles imposes fundamental constraints on the global electrodynamics of the system and highlights the importance of collective effects. In the standard flare model, the nonthermal electrons are accelerated in the corona and stream down the field lines to the footpoints, where they interact collisionally leading to bright HXR emission. RHESSI observes electrons primarily where they are stopped in the thick target at the footpoints. Consequently, it is difficult to isolate properties of the acceleration mechanism because of unobserved transport effects such as collective plasma effects.

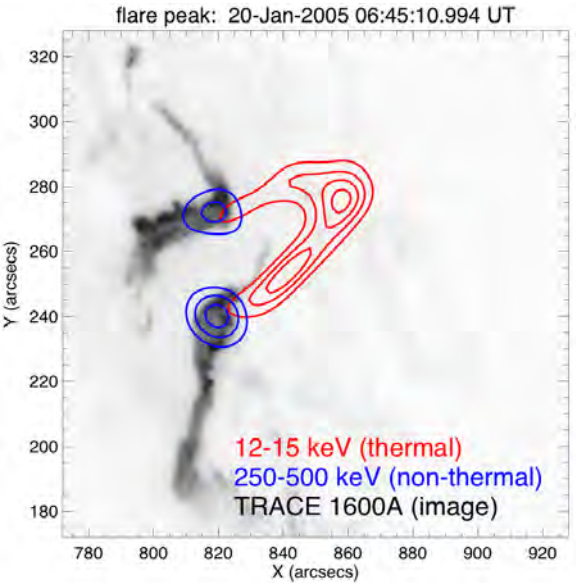


Fig. 3. Note the lack of HXR emission (blue) detected from the flare loop (red) at the peak time of the flare (Krucker et al. 2008).

As illustrated in Fig. 3, RHESSI's dynamic range limitations make it difficult to image X-rays from electrons in the tenuous corona in the presence of much brighter footpoints. Also, the low-surface-brightness albedo of HXRs back-scattered from the photosphere - crucial for electron anisotropy measurements - cannot be imaged with RHESSI. Both the instrument sensitivity and the dynamic range must be increased by a factor of at least 10, and preferably 100, to enable such observations to be made

routinely. These improvements can be achieved with HXR grazing-incidence focusing optics that operate up to ~100 keV. HXR observations of electrons in the corona with these improved capabilities will show how the electron spectrum evolves in space and time from the acceleration site to the footpoints. The high sensitivity and dynamic range will also allow the HXR albedo to be imaged.

The degree of electron anisotropy can also be determined from HXR and gamma-ray imaging polarimetry. New polarimeters, designed for the first time specifically for this purpose, can provide measurements of the magnitude and direction of the polarization that can distinguish, for example, between different particle acceleration models.

2.1.2 Flare-accelerated Ions

Because of their large energy content, no model of SEEs can be complete without addressing the acceleration of ions. In fact, there appears to be a rough equipartition in energy between electrons and ions that interact with the Sun (Ramaty & Mandzhavidze 2000), and the flare acceleration of ions and electrons is generally tightly correlated (e.g., Shih et al. 2009). However, the sources of ion-associated emission and electron-associated emission in the two flares for which both were well observed have shown surprising spatial separations, indicative of unexplained differences in the acceleration and/or transport processes (Hurford et al. 2003, 2006). Fig. 4 shows the best example of this separation. The relationship between the locations where ions and electrons interact can be studied in a systematic fashion using observations with the sensitivity and resolution to detect and resolve sources in medium to large flares. This requires high-resolution gamma-ray imaging with a factor of >~10 increase in sensitivity, and <10 arcsec FWHM resolution (with a clean point-spread function) compared to RHESSI's 35 arcsec resolution. Imaging is necessary not only in the neutron-capture line as achieved with RHESSI, but also in the nuclear de-excitation gamma-ray lines (both resolved and unresolved) and in the positron-annihilation line at 511 keV. Comparing the spectral, spatial, and temporal behavior of ions and electrons requires X-ray and gamma-ray imaging spectroscopy from <10 keV to >15 MeV. Moderate-resolution gamma-ray spectroscopy (~3% FWHM) can disentangle the various gamma-ray components that are produced by energetic ions. Besides information about the spectrum, directivity, and composition of the accelerated ions, these observations provide information about the ambient plasma in the interaction region, including abundances.

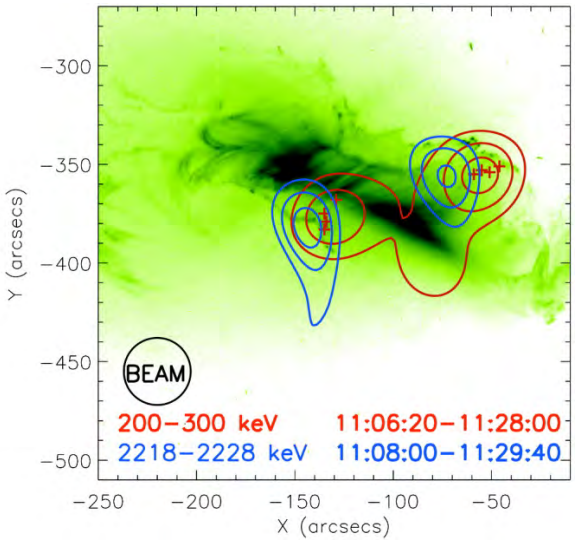


Fig. 4. Unexplained spatial separation of X-ray (red) and gamma-ray (blue) sources during 2003 Oct. 28 flare (Hurford et al. 2006).

Determinations of the total ion energy content have large uncertainties largely because the spectrum is completely unknown below a few MeV, where a major fraction of the total energy may reside. Three promising signatures of such low energy ions are, (1) radiative-capture gamma-ray lines (MacKinnon 1989; Share et al. 2001), (2) Doppler-shifted Lyman-alpha 1216 Å line emission produced by ~300 keV protons that charge exchange in flight, and (3) Doppler-shifted He II 304 Å emission produced by accelerated alpha particles.

2.1.3 Particle acceleration in CME-driven shocks

In situ observations of interplanetary SEPs will be made by inner-heliospheric missions such as Solar Orbiter, Solar Probe Plus, etc.. However, the broader understanding of the spatial and temporal origins of SEPs has been severely restricted by the limitation to such *in situ* measurements since SEPs are scattered in turbulent magnetic fields. *In situ* observations at only two or three separate locations from an evolving

3-D heliospheric shock cannot address such basic questions as the relative importance of perpendicular and parallel shocks (Tylka and Lee 2006), the radial dependence of acceleration efficiencies, and why some shocks fail to accelerate detectable SEPs (Cohen et al. 2005). These questions can only be addressed by remote imaging of SEP signatures produced by the interactions with ambient coronal and interplanetary particles (Kahler and Ragot 2009).

The serendipitous discovery of energetic neutral atoms (ENAs) at MeV energies associated with the 2006 December 4 eruptive event (Mewaldt et al. 2009, 2010) has introduced a new probe of SEP acceleration. ENAs are not deflected by the magnetic field, and the charge exchange that neutralizes the ions does not significantly affect their energies. It is unlikely that ENAs produced in the chromosphere can escape without charge exchanging again, but imaging of CME-associated ENAs in association with coronagraph CME observations can then address basic questions of shock characteristics and SEP acceleration. ENA measurements from ~10 keV to ~10 MeV include both higher-energy ENAs produced by charge exchange with heavy ions (e.g., oxygen) and lower-energy ENAs that are predicted to result from charge exchange with neutral hydrogen and hydrogen-like helium in the ambient medium.

High-sensitivity gamma-ray imaging will distinguish between interacting flare-accelerated ions that are localized to the flare loops and those ions accelerated by the CME shock that precipitate to form a larger, diffuse source. At particle energies above ~100 MeV, imaging of pion-decay gamma-ray emission may distinguish between flare-accelerated particles in the chromosphere and CME-shock-accelerated ions in the high corona or solar wind.

In addition to measuring the SEPs themselves, understanding the acceleration process requires measuring the shock characteristics, the plasma conditions, and the seed population from which they are accelerated. The coronagraph images described earlier track the latitudinal and radial development of the CME-driven shock at the crucial heights within ~4 R_{sun}, where the SEPs are believed to be accelerated. Spectroscopy with a UV coronagraph measures parameters of both the pre-shock and the post-shock coronal plasma, including the density, temperature, compression ratio, outflow speed, and the elemental and charge-state composition. Such observations with high sensitivity, cadence, and spectral range will also measure the seed population of suprathermal protons. The seed population may also consist of the flare-accelerated particles that would be simultaneously studied.

2.2 Evolution of SEE Magnetic Fields

Any attempt to understand SEEs must begin with knowledge of the coronal magnetic field, the configuration leading up to the energy release, and the evolution during and following the event. Since it is difficult to measure the coronal magnetic field ***strength***, estimates are traditionally made from force-free extrapolations of the measured photospheric vector magnetic field. Direct measurements of coronal magnetic fields with the Advanced Technology Solar Telescope (ATST) and from the Frequency Agile Solar Radiotelescope (FASR) should be available by 2020. These measurements can be coupled with measurements of the coronal magnetic field ***structure*** from space through high-spatial-resolution optical and EUV coronagraph observations. Images of the corona taken from the ground during recent eclipses (Fig. 5) demonstrate that fine-scale coronal structure in electron density is visible, and that it traces the magnetic field above the solar limb. An eclipse-like visible-light coronagraph would image the magnetic field without temperature bias, and can also observe the pre-eruption magnetic configuration as regularly done during eclipse observations. Such images taken in space every 10 s with ~5 arcsec angular resolution between 1.02 (14 Mm) to 4 R_{Sun} would reveal the changes to the magnetic configuration on the expected fine spatial and temporal scales This range covers the expected altitude of the initial energy release. Extending coronagraph observations out to 15 R_{Sun} will allow the energy release site to be observed in off-limb events and to follow the CME during its acceleration phase up to its full velocity. Such observations can provide strong constraints on theories of CME initiation and acceleration (see Forbes et al. 2006). These measurements can also determine the coronal magnetic field configuration during the build-up of energy before the event.

2.3 Evolution of SEE Plasmas

Measurements of plasma parameters in the SEE energy release/particle acceleration sites with appropriate spatial and temporal resolution are needed. It is necessary to obtain diagnostics of the current sheets where magnetic reconnection occurs. Spectra from an off-limb EUV imaging spectrometer can determine the density, temperature, inflow speeds (Yokoyama et al. 2001), outflow speeds (Wang et al. 2005), and turbulent velocities associated with large-scale current sheets that have been observed extending from the flare loops to the CME core.

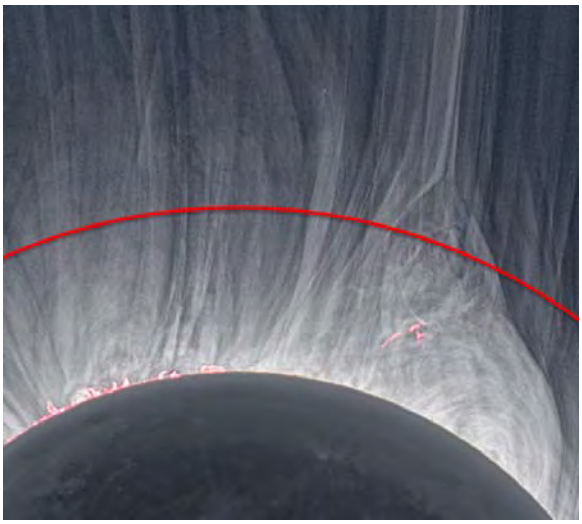


Fig. 5. Eclipse picture from Druckmüller et al. (2006) with ~5 arcsec effective resolution showing the complex magnetic structure not visible with STEREO/COR1 below 1.4 R_{Sun}.

2.3.1 Flare plasma heating

Observations of the co-evolution of energetic particles and thermal plasma are critical to obtaining an understanding of particle acceleration and plasma heating in solar flares. Direct plasma heating is expected to occur in the energy release region in the corona. The hottest flare plasma of one event has been located high in the corona, possibly at the location of particle acceleration (Caspi and Lin 2010), and could be evidence of such direct heating. It is well established, however, that accelerated electrons heat plasma away from the acceleration region, especially in the chromosphere and lower transition region. The chromospheric plasma indirectly heated in this way expands upward into the corona (Abbett & Hawley 1999; Allred et al. 2005), and this, in turn, changes the conditions in the acceleration region. Coincident imaging spectroscopy is required in the EUV, soft X-rays, and hard X-rays to obtain a clear picture of this co-evolution of the particle acceleration and plasma heating. The observations must cover a full active region with <10 s cadence and moderate spatial resolution so the same features are observed in the EUV, and soft and hard X-rays. High spectral resolution is desired to take advantage of computed line shapes, to better determine plasma velocities down to ~10 km s^{-1}, and determine the presence of line excitation by nonthermal particles. A high sensitivity EUV instrument is also critical for observing pre-impulsive phase energy release and plasma evolution.

2.3.2 CME plasma heating

Although the kinetic energy from the large-scale mass motion of the CME is the obvious manifestation of energy escaping from the Sun, the ejected plasma is strongly heated, as is apparent when erupting filaments change from absorption to emission features. Understanding the energetics of this aspect is crucial to understanding the total energy budget. The energy in the heated plasma is comparable to the kinetic energy of the CME in all the events studied so far (Landi et al. 2010). While heating from dissipation of the magnetic field is strongly suspected, the nature of the heating mechanism is not yet known. Detailed studies over many events of such implications can be made using UV and EUV spectroscopy of lines formed in the 10^4 to 2×10^6 K range at several heights between ~1.2 and ~3 R_{sun}. To measure the heating as a function of height, one requires the spectral range and sensitivity to obtain good density diagnostics, along with observations at multiple heights and high enough cadence to identify specific features at multiple heights. One also needs a spectral range that provides a large number of lines because the plasma can quickly depart from ionization equilibrium and more constraints are needed to determine plasma parameters.

3 Mission Concept

The proposed SEE 2020 mission will measure all the energetically and diagnostically important aspects of solar eruptive events. It will have the temporal, spatial, and spectral resolution and coverage necessary to

identify the coronal conditions needed for an eruption to take place and the physical processes that make it happen. The instruments necessary to make the required new measurements are listed in Table 1 and described in more detail in separate white papers. They will be combined together on a single spacecraft, likely in low-Earth orbit. Observations from ground-based observatories (such as ATST and FASR) and other space missions (such as Solar Orbiter, Solar Probe Plus, and Solar-C) will complement this mission concept by providing context information on magnetic fields, thermal plasma, and interplanetary SEPs. The technologies capable of achieving the instrument requirements that have not already been proven are currently under development and will be demonstrated on a variety of rocket and balloon missions over the next few years. We envision SEE 2020 to be similar in cost and schedule to the Solar Dynamics Observatory (SDO).

Table 1. Complement of instruments for the SEE 2020 mission to be flown on a single spacecraft, likely in low-Earth orbit.

Instrument	Key Instrument Characteristics	New Measurement	Primary Science Question
White-light coronagraph	Angular resolution: 5 arcsec Full Sun FOV: 1.02 to 12 R_{sun}	Observe the rapidly changing structure of the magnetic field at the energy release site.	How does evolution of the coronal magnetic field lead to eruptive events?
Off-limb EUV imaging spectrometer	Angular resolution: 1 arcsec Temperature: 10^4–10^7K Cadence: ~10 s	Observe CME acceleration and heating at very low altitudes.	How are CMEs accelerated?
HXR imaging spectrometer	Sensitivity: >100 × RHESSI Dynamic Range: >100:1	Image electrons in the corona as they are accelerated.	How are electrons accelerated?
Gamma-ray imaging spectrometer	Angular resolution: 5 arcsec Sensitivity: >10 × RHESSI	Image and separate all ion-associated sources.	How are ions accelerated?
ENA imaging spectrometer	Angular resolution: 0.1 degrees Sensitivity: 10-100× better than STEREO/LET		
EUV imaging spectrometer	Temperature: 10^4–10^7K 4D* imaging at <10 s cadence	Observe the coronal plasma where energy is released.	How is the plasma heated?
SXR imaging spectrometer	Energy res.: 1 eV (1 to 20 keV) 4D* imaging at <10 s cadence		

* "4D imaging" means imaging in two spatial dimensions over time and energy or wavelength.

References

Abbett, W. P., & Hawley, S. L. 1999, Astrophysical Journal, 521, 906

Allred, J. C., Hawley, S. L., Abbett, W. P., and Carlsson, M. 2005, Astrophysical Journal, 630, 573

Caspi, A. and Lin, R. P. 2010, Astrophysical Journal Letters, in press.

Cohen, E., Kessler, D. A., & Levine, H. 2005, Physical Review E, 72, 66126

Druckmüller, M., Ruˇsin, V., & Minarovjech, M. 2006, Contributions of the Astronomical Observatory Skalnat´e Pleso, 36, 131

Emslie, A. G., Kucharek, H., Dennis, B. R., Gopalswamy, N., Holman, G. D., Share, G. H., Vourlidas, A., Forbes, T. G., Gallagher, P. T., Mason, G. M., Metcalf, T. R., Mewaldt, R. A., Murphy, R. J., Schwartz, R. A., & Zurbuchen, T. H. 2004, Journal of Geophysical Research, 109, 10104

Emslie, A. G., Dennis, B. R., Holman, G. D., and Hudson, H. S., 2005, Journal of Geophysical Research, 110, 11103

Forbes, T. G., et al. 2006, Space Science Reviews, 123, 251

Holman, G. D. et al., 2010, Space Science Reviews, in press.

Hurford, G. J., Krucker, S., Lin, R. P., Schwartz, R. A., Share, G. H., & Smith, D. M. 2006, Astrophysical Journal, 644, L93

Hurford, G. J., Schwartz, R. A., Krucker, S., Lin, R. P., Smith, D. M., & Vilmer, N. 2003, Astrophysical Journal, 595, L77

Kahler, S. W., & Ragot, B. R. 2008, Astrophysical Journal, 675, 846

Kontar, E. P. et al., 2010, Space Science Reviews, in press.

Krucker, S., Hurford, G. J., MacKinnon, A. L., Shih, A. Y., & Lin, R. P. 2008, Astrophysical Journal, 678, L63

Krucker, S., Hudson, H. S., Glesener, L., White, S. M., Masuda, S., Wuesler, J.-P., and Lin, R. P. 2010, Astrophysical Journal, 714, 1108.

Landi, E., Raymond, J. C., Miralles, M. P., & Hara, H. 2010, Astrophysical Journal, 711, 75

MacKinnon, A. L. 1989, Astronomy and Astrophysics (ISSN 0004-6361), 226, 284

Mewaldt, R. A., Leske, R. A., Shih, A. Y., Stone, E. C., Barghouty, A. F., Cohen, C. M. S., Cummings, A. C., Labrador, A. W., von Rosenvinge, T. T., & Wiedenbeck, M. E. 2010, Twelfth International Solar Wind Conference. AIP Conference Proceedings, 1216, 592

Mewaldt, R. A., Leske, R. A., Stone, E. C., Barghouty, A. F., Labrador, A. W., Cohen, C. M. S., Cummings, A. C., Davis, A. J., von Rosenvinge, T. T., & Wiedenbeck, M. E. 2009, Astrophysical Journal Letters, 693, L11

Qui, J. et al. 2004, Astrophysical Journal, 604, 900

Ramaty, R., and N. Mandzhavidze (2000), Gamma-rays from solar flares, in Highly Energetic Physical Processes and Mechanisms for Emission from Astrophysical Plasmas, ASP Conf. Ser., IAU Symp. 195, edited by P. C. H. Martens, S. Tsuruta, and M. A. Weber, p. 123

Share, G. H., Murphy, R. J., and Newton, E. K. 2001, Solar Physics, 201, 191

Shih, A. Y., Lin, R. P., & Smith, D. M. 2009, Astrophysical Journal Letters, 698, L152

Sui, L., & Holman, G. D. 2003, Astrophysical Journal, 596, L251

Tylka, A. J., & Lee, M. A. 2006, Astrophysical Journal, 646, 1319

Wang, T., Sui, L., and Qiu, J. 2007, Astrophysical Journal Letters, 661, L207

Yokoyama, T, Akita, K., Morimoto, T., Inoue, K., & Newmark, J., 2001, Astrophysical Journal Letters, 546, L69.

Complementary White Papers

Bandler et al. "XMS"

Christe et al. "The Focusing Optics Solar X-ray Imager"

Davila et al. "Understanding Magnetic Storage, Reconnection, and CME Initiation"

Doschek et al. "A Concept White Paper for a New Solar Flare Instrument Designed to Determine the Plasma Parameters in the Reconnection Region of Solar Flares at Flare Onset"

Doschek et al. “The High Resolution Solar-C International Collaboration”

Gary et al. "The Frequency Agile Solar Radiotelescope"

Gopalswamy et al. “Earth-Affecting Solar Causes Observatory (EASCO): A New View from Sun-Earth L5”

Laming et al. "Science Objectives for an X-Ray Microcalorimeter Observing the Sun"

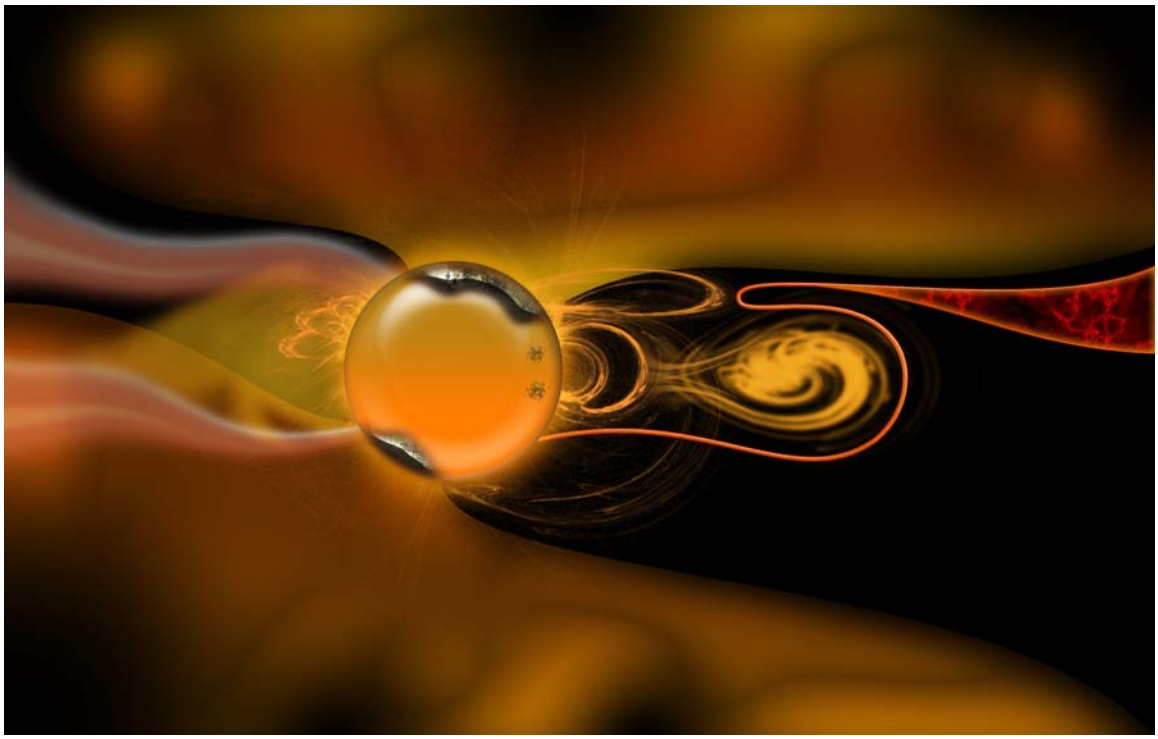
McConnell et al. "Hard X-ray and gamma-ray polarization measurements of solar flares during the next solar maximum"

Shih et al. "Solar Ion Acceleration and the Flaring Atmosphere"

St. Cyr et al. "Solar Orbiter: Exploring the Sun-Heliosphere Connection"

Vourlidas et al. "Mission to the Sun-Earth L5 Lagrangian Point: An Optimal Platform for Heliophysics & Space Weather Research"

Solar Wind Ion Composition Measurements: Direct Measurements of Properties of the Corona



White Paper Submitted to the
Decadal Survey Panel on Solar and Heliospheric Physics

November 12, 2010

Stefano A. Livi, StefanoLivi@swri.edu, 210-522-3310
Antoinette Galvin, University of New Hampshire
Lynn Kistler, University of New Hampshire
Susan Lepri, University of Michigan
Mark Popeki, University of New Hampshire
Michael Collier, Goddard Space Flight Center
Thomas Broiles, University of Texas in San Antonio

Summary

The chemical composition and charge state distribution of heavy ion elements in the solar wind will establish the physical links between the outward transport of solar energy and the solar wind, will provide direct measurements of solar eruption products in coronal mass ejections, and will discover the fundamental processes by which interplanetary shocks accelerate suprathermal ions and produce hazardous particle radiation events.

Introduction

The Sun's million-degree corona generates the solar wind, which supersonically expands into the solar system and permeates the whole heliosphere. This flow of coronal particles defines the space weather that impacts and controls Earth's space environment, the plasma environments of all the planets, and the protective boundaries that surround our solar system. Despite remote observations of the rapidly changing magnetic structure of the corona from missions such as Hinode, STEREO, SOHO, RHESSI, and TRACE, advances in understanding of the 3D solar wind structure provided by Ulysses, and the discovery of the solar wind's control over the geospace environment provided by ACE, Wind, and IMAGE, the heliophysics community is not yet able to definitively connect the morphology and dynamics of the solar atmosphere with the global structure of the solar wind, the dangerous shocks that propagate therein, and the sources and acceleration mechanisms of solar energetic particle radiation

Characterizing the heavy ion environment of the solar wind is a unique tool to determine how the Sun influences our environment:

- The measurement of the solar wind composition, in particular the heavy ion component is essential in tracing the sampled solar wind to its coronal source (Geiss et al, 1995):
  - Heavy ion composition (Figure 1), and bulk plasma parameters are correlated with solar source structures (loops, the legs or tips of the helmet streamer, spicules, polar plumes, and other solar magnetic morphologies), and are needed to unambiguously identify the source.
  - Statistical syntheses of these data-sets over the mission lifetime will yield the sources and distributions of open-flux over various parts of the solar cycle, leading to a deep understanding of the physical processes leading to solar wind, coronal hole and streamer associated solar wind.

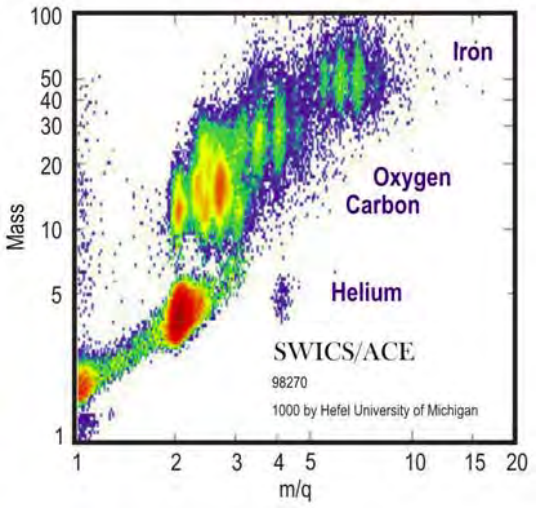


Figure 1: example of mass vs. mass-per-charge distributions from ACE-SWICS

- Heavy ion composition identifies and characterizes overall CME morphology, particularly the different structures within a CME that relate to different parts of the solar feature from which the CME originated (Li et al, 2005).

- Ion composition measurements relate in situ plasma to coronal conditions near the source regions;
  - The ionic charge states will be interpreted together with the remote observations of the coronal plasma. Elemental and compositional signatures will be related to ultraviolet diagnostics of the solar source regions.
  - Furthermore, the kinetic properties of the heavy ions promise important clues about the non-thermal processes heating the solar corona and accelerating the solar wind and particle acceleration processes.
  - The measurement of suprathermal particles from 0.5 to 100 keV/q address the fundamental properties of particle acceleration throughout the heliosphere.
  - By comparing observed suprathermal ions with energetic particles measurements (Figure 2), we expect to determine the mechanisms of particle injection and acceleration as functions of shock properties (Fisk and Gloeckler, 2007), and also stochastic sources.

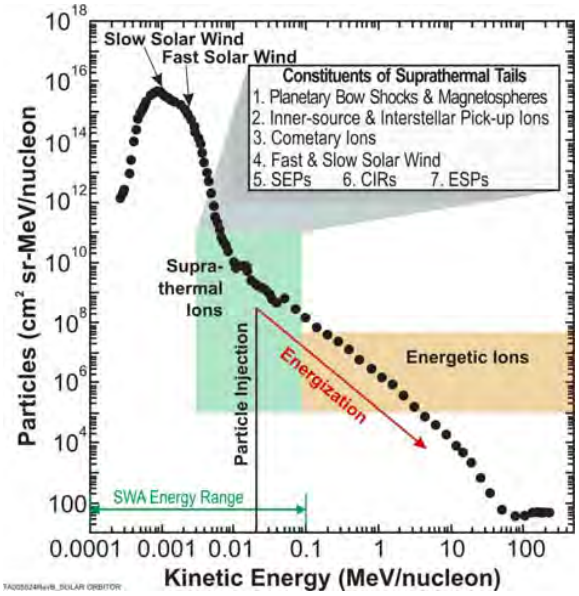


Figure 2: The composition of suprathermal tails varies with energy due to the different properties of suprathermal sources. The evolution and properties of shocks control the injection energy, acceleration rate, and, in turn, the composition of the SEPs energized from suprathermal ions [adapted from Mewaldt et al., 2006; shown here are oxygen fluences from ACE measured over a 3-year period].

Primary Scientific Goals

- Elucidating the relationship between evolving global solar wind structures and the ever changing solar morphologies of coronal holes, loop populations, helmet streamers, plumes, and other solar structures.
- Determining the global structure of evolving shocks, the spatial variability of the suprathermal ion populations, and how these properties affect particle injection and acceleration mechanisms.
- Characterizing the global topology and evolution of CMEs and determining their roles in constraining and distinguishing between CME models.

Instrument Concept

Solar wind ion composition instruments are of two basic types. The first type (i.e., Wind-MASS, SOHO-MTOF) is designed for isotopic and high resolution elemental composition measurements. The second type (i.e., ACE-SWICS, STEREO-PLASTIC) is designed to provide independent measurements of charge (q) and mass (m) for individual ions. Solar Wind characterization to link coronal properties to in-situ measured quantities demand a composition instrument with modest mass resolution obtained by the second type. State-of-the art instruments designed to this purpose (for an example, see Figure 3) comprise two major components: (1) an Electrostatic Analyzer (EA) with Ion Steering (IS) that is designed to cover the required FOV, and (2) a Time-Of-Flight (TOF) telescope with solid- state detectors (SSD) for total ion energy (E_{tot}) measurements, and microchannelplates (MCP) for detection of the timing signals generated by electrons coming from either the entrance foil or the face of the SSD . Such a sensor measures five key properties for all ions: mass (m), charge (q), speed (v) and direction of incidence (a,f). EA with IS provides E/q and elevation (f). TOF provides the azimuth (a) through the imaging of secondary electrons generated as the ion penetrates the entrance thin carbon foil.

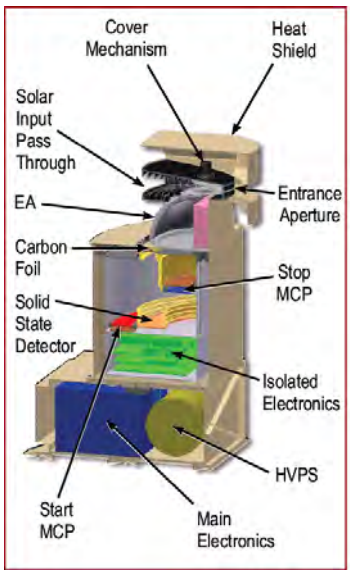


Figure 3: 3D-view of a heritage-based design to measure the 3D velocity distribution, charge state, and composition of solar wind heavy ions

To enhance secondary electron emission and push the solar wind ions above the SSD energy threshold, the TOF/SSD telescope floats at an appropriate potential, typically of the order of 20-30 kV. This potential accelerates the incoming solar wind ions so that their detection efficiencies are almost independent of the initial speed. Secondary electrons emitted from the front face of the SSDs are detected to complete the TOF measurement (t), and the SSD measures the total energy of the accelerated ions (E_{tot}). This completes the five independent measurements required for unique ion identification and characterization.

Heritage

All subsystems of this design are based on heritage designs. The top-hat electrostatic analyzers/deflectors have heritage from STEREO-SWEA, Cluster-HIA, Mars and Venus EXPRESS-ASPERA; the 20 kV power supplies have heritage from ACE-SWICS, Ulysses-SWICS, STEREO-PLASTIC; the TOF-SSD subsystems have heritage from ACE-SWICS, Ulysses-SWICS, STEREO-PLASTIC; and the TOF subsystem is currently flying on MESSENGER–FIPS.

Instrument Requirements

To answer the scientific question posed above, the instrument must comply with following minimum requirements. These requirements are derived from the existing body of observations of solar wind composition existing to-date.

Parameter	Range/Resolution	HIS
Sensors		1xEA, 1 TOF-SSD
Sensitivity	Geometric factor	10^{-4} cm^2 s sr
Mass	Resolution (m/Δm)	4
Charge state	He	+1, +2
	C	+1, (+4 : +6)
	O	+1, (+5 : +8)
	Ne	(+6 : +9)
	Mg	(+6 : +12)
	Si	(+6 : +12)
	Fe	(+6 : +20)
Energy	Range	0.5 – 60 keV/q
	Resolution Δ(E/q)/(E/q)	6%
Angle	Azimuth range	-30° : +66°
	Elevation range	-17° : +17°
Temporal	Resolution	30 sec

References

Fisk, L. A., & G. Gloeckler, Acceleration and Composition of Solar Wind Suprathermal Tails, Space Science Reviews, Volume 130, Issue 1-4, pp. 153-160, 2007.

Geiss, J., G. Gloeckler, R. von Steiger, H. Balsiger, L. A. Fisk, A. B. Galvin, F. M. Ipavich, S. Livi, J. F. McKenie, K. W. Ogilvie, and B. Wilken, The Southern High-Speed Stream – Results from the SWICS Instrument on ULYSSES, Science, 268, 1033, 1995

Li, G., G. P. Zank, and W. K. M. Rice, Acceleration and transport of heavy ions at coronal mass ejection-driven shocks, Journal of Geophysical Research, Volume 110, Issue A6, CiteID A06104, 2005

Mewaldt, R. A., C. S. Cohen, and G. M. Mason, “The source material for large solar energetic particle events” in Washington, DC, American Geophysical Union Monograph 165: Solar Eruptions and Energetic Particles, p. 115-126, 2006

White Paper: Turbulence and Nonlinear Dynamics and its many effects in Solar and Heliospheric Physics

Submitted to: NRC 2010 Decadal Survey in Solar and Heliospheric Physics

Authors: W. H. Matthaeus, M.Goldstein, M. Velli , G.P. Zank, and C. W. Smith

I. Preface & *Overview:* The complex nonlinear behavior of plasmas and fluids that we call turbulence is a subject of great importance due to its diverse and substantial impact in a variety of naturally occurring systems, ranging from the earth's dynamo and atmosphere affecting climate to many fundamental processes in the Heliosphere, and on larger scales to the behavior of galaxies and clusters of galaxies. Within the heliosphere, turbulence plays a major role in cross-scale couplings, self-organization, dissipation, and in mediating interactions involving energetic particles, the thermal plasma and self-consistent electromagnetic fields.

In many ways, turbulence and nonlinear dynamics as a basic physics element of Solar and Heliosphysics came into its own in the last ten to fifteen years. There has been an enormous increase in the number of papers studying both turbulence theory and the effects of turbulence on the state and dynamics of the coupled heliospheric dynamical systems. Applications of turbulence are no longer always considered esoteric, and many concepts such as cascade, similarity, and intermittency, have entered the mainstream vocabulary of space physicists. The near-ubiquitous importance of nonlinearity has become an accepted notion. The vocabulary based upon linear waves and instabilities, superposition of solutions, laminar flows and fields, and quasilinear theory, remains familiar, but no longer is presumed to provide a complete description. Instead these traditional simplifications are more generally regarded as pathways to addressing the full nonlinear and often strongly turbulent systems that are frequently encountered in solar and heliospheric dynamical systems.

Study of turbulence itself, as an element of fundamental physics, remains of great importance. Substantial progress has been made in the last decade, guided frequently by computer simulations and by observations of plasmas in space and in the laboratory. But there remains much to understand and the subject remains far from complete – consider that hydrodynamic turbulence, a simpler system, remains unsolved although it has been studied at least a century longer than plasma turbulence. The study of applications and implications of turbulence in solar and heliospheric physics is an even bigger subject.

The purpose of the present White Paper is to direct the attention of the Survey Committee to the broad implications of nonlinearity and turbulence within the complex, coupled solar and heliospheric system. This had been a major theme in the prior Survey, and we suggest that this emphasis be continued. This involves an awareness at several levels – instrumentation, specific missions, long term monitoring and prediction, turbulence aspects of studies of various

phenomena, interdisciplinary implications, and planning for required observations on an opportunistic basis. Here we summarize briefly several aspects of these issues that we suggest to be relevant to the Survey's deliberations and priorities.

[Note: For brevity, instead of including a very large number of references, we will rely mainly on a few review articles given in the References Section.]

II. Some Accomplishments . In the past decade or so, there has been a great increase in understanding the basic physics of heliospheric turbulence, and in understanding how turbulence enters into the complex, coupled solar–heliospheric dynamical system.

SOLAR WIND studies have often led the way, due to the availability of large amounts of data to carry out the statistical studies that so often characterize turbulence studies. Indeed, the state of the system at 1 AU has been and remains a testing ground for turbulence theories. Correlation functions and associated spectral distributions of energy, cross helicity, magnetic helicity, and Alfvén ratio are familiar descriptions of the state of the turbulence, and these have leading order implications for plasma heating and for scattering and transport of energetic particles. Searches for "universal" laws in turbulence continue, but are moderated by the prevalent observations of a dynamic and highly variable medium, driven ultimately by solar variability itself. Generally one finds broad (and frequently log-normal) distributions of turbulence properties such as energy densities and correlation scales, while powerlaw indices and anisotropies are also broadly distributed from parcel-to-parcel. It is not just one type of turbulence, but rather a family of turbulence types, reflecting the variability of the coupled components of the heliospheric system. Understanding and monitoring this variability is a problem of great fundamental and practical importance.

Understanding the radial and latitudinal dependence of turbulence throughout the solar cycle is an essential element of both space weather (inner heliosphere) and the boundaries of the heliosphere. This major ongoing area of inquiry is essential for prediction of the Geospace environment, for support of potential human exploration, for understanding galactic cosmic rays, and finally, as a prototype for many astrophysical problems.

In the inner solar wind, a major problem is to understand the interaction of CMEs and shocks with both pre-existing fluctuations and with possible admixtures of self-generated turbulence. The turbulence interactions can be expected to have a major influence on the forces, such as drag, that are imposed on the CME. The back reaction on the ambient medium is also complex, and directly affects particle acceleration and transport. These are major factors entering Space Weather and prediction, and need to be understood in the coming decade.

In SOLAR physics, turbulence is probably ubiquitous. Its presence and impact in the photosphere and in the convection zone is unquestioned. Furthermore the complexities of

turbulence enter into the fundamental physics of the dynamo. There seems to be no doubt, for example, that the stochasticity of dynamo reversals is intimately related to the nonlinear dynamics and self-organization properties of the MHD dynamo, even if the details are still being elucidated. Turbulence and nonlinear dynamics in the corona, in both open field regions and closed loops, has been recognized as a factor in heating and possibly in the origin of the solar wind itself. Meanwhile very low frequency variations, probably originating with self-organization properties of the dynamo, appear to affect the corona and even the observed magnetic and density variability in the solar wind at 1AU. Far from being only an issue of an inertial range "Kolmogorov cascade" the nonlinear properties of the system influence the behavior across a very wide range of space and time scales. The landmark imaging of instruments such as TRACE, Lasco, and EIT have defined graphically the degree of structure and variability in the solar dynamical system.

Turbulence is also potentially major factor in MAGNETOSPHERIC dynamics and in SOLAR WIND-MAGNETOSPHERIC couplings.

III. Turbulence related problems for the coming decade: Relation of turbulence to other fundamental processes.

Reconnection. While the relation between turbulence and reconnection has been studied for several decades, there has recently been increased attention to this topic. Part of this is due to recognition that reconnection is an ongoing and statistical aspect of a turbulent medium, and turbulent fluctuations can drive reconnection to higher rates. Another aspect is the recognition that fluctuations (preexisting or due to plasmoid instabilities) can generate multiple reconnection sites, and these may substantially modify the associated reconnection rates.

Dynamo In various formulations of the dynamo problem, its efficiency is related to coupling across scales and behavior of fluctuations and small scale reconnection activity. Both the regularity and the stochasticity of reversals must be related to properties of the turbulence and nonlinear dynamics, and this remains an active area or research. Ultimately these are the drivers of the magnetic atmosphere of the sun that control the properties of the heliosphere.

Magnetic self-organization This term refers to the dynamical generation of long lived nonlinear structure in a turbulent plasma or MHD medium. It is well known in plasma confined devices such as Reversed Field Pinch or Spheromak experiments. It also has antecedents in pure electron plasmas and in two dimensional or rotating fluids, and in certain MHD systems with a large scale conserved quantity such as magnetic helicity. Analogous physics is present in the dynamo, and in magnetic flux tubes such as coronal loops and even potentially in open field regions of coronal holes. Not surprisingly the statistical dynamics of turbulence plays a key role in understanding the dynamical path to self-organized or relaxed states. Interestingly, there is

recent evidence from simulations, and from laboratory experiments, that self-organization can lead to generation of very long time tails on correlation functions. Some solar observations seem to support this idea as well. This generation of very low frequency "noise" directly influences predictability. This is a "large scale" effect of turbulence, in contrast to the direct cascade smaller scale effects. This brings home the point that study of turbulence also involves large scale features. [See white paper by Brown et al.]

Magnetospheric Dynamics. The magnetosphere in itself is a complex nonlinear system, and driven by interactions with the turbulent solar wind. This is a system that is highly nonlinear, but also inhomogenous, involving close coupling of kinetic scales. A very difficult turbulence problem at moderate to high Reynolds number, magnetospheric studies have recognized the turbulent nature of busty bulk flows, the highly dynamic magnetosheath, nonlinear pulsations, and large excursions from the time averaged configuration.

CME dynamics The problem of the interaction of CMEs with background interplanetary medium is analogous to high Reynolds number turbulent flow past an obstacle, but more complex because of expected kinetic effects. Turbulent drag, self consistently generated fluctuations, particle acceleration, and leading edge shock formation are aspects of the problem that depend on turbulence. This is a very important class of problems for Space Weather.

Dissipation, heating and Intermittency. Termination of the turbulent cascade through kinetic processes, commonly described as "dissipation," is a fundamental problem in space physics and in plasma astrophysics in general . Irreversible conversion of collective fluid or kinetic scale motions into internal energy, or ``heating'' is a closely related concept. Neither is well understood in plasmas such as the corona and solar wind, yet the operation of such mechanisms is responsible for the presence of the wind and the overall structure of the heliosphere . This is a major fundamental issue in which turbulence ranging from fluid to kinetic scales clearly plays a central role. [The White Paper by Podesta et al discusses this further.] Recent observational and theoretical evidence connect dissipation and small scale non-Gaussian inhomogeneities such as current sheets. Far from being only an academic problem, the intermittency of turbulence provide a dynamical connection between coherent structures (e.g., current sheets), patchiness of dissipation, and the entire inertial range of turbulence. Understanding this truly cross scale relationship is a challenge for the next decade.

Shock Structure. Shocks are widespread in space and astrophysics and represent important boundaries between plasmas and sources of energetic particles. There is a growing indication, from both theoretical and observational directions, that turbulence plays an important role in shock formation and dynamics, in addition to being an essential element in diffusive shock acceleration. Further research in shock-turbulence interactions including observational study is

needed to fully understand this topic, which also has important implications for astrophysics, such as in SNR shock that propagate in a turbulent medium as well.

Particle Acceleration The two most of prominent models of charged particle acceleration are diffusive shock acceleration, and stochastic acceleration. Both depend in an essential way on the nature of the turbulent fluctuations, which are often viewed either as self-generated by the energetic particles, or as random phase waves. Treating the fluctuations more realistically as turbulence allows for coherent structures -- it then becomes possible to have trapping and other effects that further enhance acceleration. The more general interaction of turbulent dissipation electric fields with the self-consistent particle distributions in a plasma environment with few collisions, and the resulting tails in particle distribution functions, have only begun to be explored. The possibility that the ubiquitous 1/v^5 spectrum is due to such types of effect is an important subject to be studied.

Particle scattering and transport Scattering theory depends explicitly on turbulence properties including the energy spectrum, the time variations, and in principle, the higher order statistics. The attainment of the diffusive limit can be directly influenced by certain details of the turbulence properties. Furthermore, coherent structures such as persistent flux tubes can cause trapping effects and delayed attainment of diffusion. Evidently particle transport cannot be understood without a fairly sophisticated knowledge of the turbulence.

Solar Modulation of Galactic Cosmic Rays. This classic problem in Heliospheric Physics is a grand challenge problem in understanding the effects of turbulence. It involves understanding large scale flows over great distances, and it involves knowing how the particle transport properties depend on the turbulence properties (i.e. the diffusion coefficients). Finally it involves understanding how the turbulence properties vary in space and time throughout the Heliosphere. Every aspect of this problem depends upon understanding the turbulence.

Prediction and Space Weather A variety of effects influence our ability to predict magnetic fields and particle distributions at an particular point in the heliosphere, such as decorrelation of fluctuations in time, field line random walk, turbulent drifts, trapping and stochastic orbits. Each of these depends in detail on turbulence properties. Consequently, prediction for Space Weather purposes will be limited as a matter of principle by these effects. It will be important to develop a quantitative understanding of these influences, beginning from theory, and supported by extensive observational studies. .

Conclusion: Turbulence enters into many crucial Heliophysical dynamical processes, influencing cross scale couplings in an essential way. Therefore beyond its justification as basic physics (which may be a matter of aesthetic appeal), study of turbulence emerges as an essential component of Solar and Heliospheric research.

IV. Implications and Recommendations for Missions:

a) Cruise modes. Heliospheric and magnetspheric turbulence studies have often proceeded as a valuable "add-on" to other missions. It is important that missions designed for other goals (e.g., distant exploration, planetary studies, etc,) have "cruise modes" that will add to our knowledge of the physical and statistical properties of turbulence whenever it is feasible to do so. [See white paper by C. Smith]

b) Support of planned discovery missions. Solar Probe Plus and Solar Orbiter are milestone missions of exploration and will make breakthrough discoveries. It is likely that SPP and SO will uncover the nature of fluctuations and turbulence in the inner solar system, and in particular will reveal how turbulence plays role in the origin and variability of the solar wind. The scientific benefit of these missions will be greatly enhanced by going forward with supporting space and ground based missions to the extent that this is feasible. [See white paper by Cranmer et al.]

c) Current Multispacecraft Missions: Cluster has led the way in understanding nonlinear behavior of the magnetsphere and in uncovering the space time fabric of fluctuations, including solar wind fluctuations (along with ACE and Wind). MMS will continue this, with an emphasis on magnetopause and magnetotail reconnection, using high resolution instrumentation. Furthermore Cluster and MMS have motivated intensive development of multispacecraft analysis methodologies. In view of the recent recognitions of the intimate relationship between turbulence, reconnection, and particle acceleration, MMS science benefit can be maximized by restoring the original emphasis on all three of these goals in its implementation phases.

d) New Multispacecraft missions Roadmap missions such as SEPAT, HMAG and HELIX, in multispacecraft implementations, can provide vital information about the structure and variability of Heliospheric turbulence. (Older mission concepts such as of "Sentinels" missions, and multispacecraft Plasma Turbulence "explorer" missions had been discussed having similar goals.) If properly configured and designed such missions can serve multiple purposes of Space Weather effects and turbulence studies.

e) Continuation of a 1 AU monitor satellite at L1 is greatly needed, as was emphasized by the prior Survey. To include turbulence studies as well as Geospace and geoeffectiveness parameters, adequate time resolution and accuracy are required. Even better scientifically would be at least a pair of s/c to provide some separation of space and time effects.

f) Possible new mission: Combining elements of the above ideas, the scientific interests in Space Weather, 1AU monitoring, inputs to magnetospheric modeling, and long term studies of solar wind turbulence, including anisotropies and dissipation mechanisms, could be

simultaneously satisfied by a new long lifetime Cluster/MMS style mission at L1. We urge the survey Committee to consider such a mission.

g) Theory and Modeling Support. A deal of intellectual effort will be required to continue advances in understanding both the fundamental physics and the practical implications of the complex, multiscale nonlinear dynamics of the coupled Heliosphysics systems. This mandates a strong and continued support of efforts in a broad range of theory and modeling efforts.

REFERENCES

Goldstein, M. L.; Roberts, D. A.; Matthaeus, W. H., Magnetohydrodynamic Turbulence In The Solar Wind, Annual Review of Astronomy and Astrophysics, Volume 33, 1995, pp. 283-326.

C.-Y. Tu & E. Marsch, MHD Structures, Waves and Turbulence in the Solar Wind: Observations and Theories, , Space Sci Rev., 73, 1 (1995)

Roberto Bruno and Vincenzo Carbone, "The Solar Wind as a Turbulence Laboratory", Living Rev. Solar Phys. 2, (2005), 4. URL: http://www.livingreviews.org/lrsp-2005-4

Turbulence in the Solar Atmosphere and Solar Wind, A. Petrosyan, A. Balogh, M.L. Goldstein, J. Léorat, E. Marsch, K. Petrovay, B. Roberts, R. von Steiger, J.C. Vial, Space Sci. Rev., DOI 10.1007/s11214-010-9694-3; http://www.springerlink.com/content/102996/

D. Biskamp, Magnetohydrodynamic Turbulence, Cambridge U. Press, New York, 2003

U. Frisch, Turbulence: The Legacy of A. N. Kolmogorov, Cambridge U. Press, New York, 1995

For an up to date and more complete bibliography on this subject, see the review article "Who Needs Turbulence", by W. H. Matthaeus and M. Velli, to appear in the Proceedings of the 2010 Yosemite Workshop on Magnetic Reconnection, Special Issue of Space Science Reviews, ed. by W. Lewis (to be published.)

Ultra-heavy nuclei in solar flare: the rarest elements in the Sun

Summary

The ACE and Wind missions revealed that some particle events with 3He and other more abundant ion enrichments also have tremendous enrichments if ultra-heavy nuclei. This heliospheric source of ultra-heavy (UH) energetic particles offers a unique view of not only the abundances of these ions but also how they become preferentially enriched in the flare process. UH abundances are important in understanding their nucleosynthesis and further processing in flare acceleration, interplanetary acceleration, and transport. There have been efforts to measure them in the galactic cosmic ray population thereby understanding the cosmic ray source in general and stellar nucleosynthesis in particular. The cosmic rays require large area detectors because of the relatively low abundances and thus intensity. The solar source of enriched UH is more accessible because of the relative intensities involved in the 3He-rich particle events. In addition, how the sun creates these enrichments is unknown.

Past achievements

The Wind and ACE measurements of UH enrichments in solar particle events quantified their abundances with much better resolution and event diagnostics than possible with earlier nuclear emulsion experiments [e.g. Price & Shirk 1975]. As a result we know that:

- In events associated with large coronal mass ejections, UH nuclei up to Z~82 are similar to coronal abundances [Reames, 2000]
- The average enhancement of heavy nuclei increases with mass in 3He-rich event with values of ~7 for Fe, ~40 for mass 78–100 amu, ~120 for mass 125–150 amu, and ~215 for 180–220 amu [Mason et al. 2004]
- The enhancements in 3He-rich events are approximately proportional to the particle charge-to-mass ratio raised to a power, as seen previously in large, shock-associated SEP events [Mason et al. 2004]

Comments on implementation

Progress beyond the Wind/EPACT and ACE/ULEIS measurement set requires established charged particle sensor techniques, but with at least an order of magnitude increased geometry factor. In this sense progress means an order of magnitude better counting statistics. Measurements at lower-energies (less than 1 MeV/nucleon) are of course relatively easier because of the typical SEP energy spectrum.

A dedicated mission to study solar UH enrichments seems out of the question because of its limited scope, and the appearance of duplicating ACE objectives could be a problem in gaining popular support. Perhaps making the solar UH

measurements in conjunction with an astrophysical mission that targets galactic cosmic ray composition along the lines of potential TIGER spaceflight mission [e.g. Waddington 2007] would be a viable approach. It has the added selling feature of using energetic particles from the sun and the cosmos to understand the cosmic origins of heavy elements.

References

Price, P. B. & Shirk, E. K., 14th Internat. Cosmic Ray Conf. 1, 268, 1975

Waddington, C. J., The Ultra Heavy Elements in the Cosmic Radiation, Space Sci. Rev. 130, 457-464, 2007.

Mason, G. M. et al., Abundances of heavy and ultra-heavy ions in 3He-rich solar flares, ApJ 606, 555-564, 2004

Reames, D. V., Abundances of trans-iron elements in solar energetic particle events, ApJ Lett. 540, L111-L114, 2000

Need to measure solar energetic particle ionization states from ~1 to above 100 MeV/nucleon

Summary

The ionization state of solar energetic particles is a critical parameter for understanding their acceleration and propagation as particle rigidity depends on the charge state. However the measurement is resource-intensive using electrostatic deflection techniques and currently cannot be performed at high (greater than a few MeV/n) energies, especially at 100's of MeV/nucleon where there is continued uncertainty over whether the ions originate in a flare or a lower-temperature environment. We propose the implementation of one or more charged particle sensors in low-Earth polar orbit to infer the ionization state of solar energetic particles using the geomagnetic cutoff technique. The cutoff technique is established and not new; neither is the scope or technology of the sensors themselves which require only elemental composition, nor are there needs for magnetic field measurements because the technique is self-calibrating. We imagine the hosting of these sensors on either a dedicated Heliophysics mission that requires the use of polar LEO, or the hosting of one or more sensors as missions of opportunity. One benefit of the sensors themselves is that they measure the intensity and abundances of the events while inferring charge states, thus negating the need for additional sensors to measure the events elsewhere in near-Earth space.

Role of SEP ionization states

The importance of SEP charge states in understanding the source, acceleration, and propagation processes has been discussed and established elsewhere. The insights provided by the first measurements using ISEE, and the later clues about the sources of ^{3}He-rich events and CME-related events provided by ACE and SAMPEX (e.g. Mazur et al. 1998; Klecker et al. 2006 & references therein) make it clear that there is not an average charge state one can apply to a given SEP event. There are energy-dependences that could arise from the source, and there are reasons to suspect that the particles that we see at 1 AU in one energy interval actually reflect a history of several energies at the Sun once adiabatic deceleration and other propagation effects are considered. Thus the existing databases of ACE and SAMPEX solar particle charge states are useful for the studies of those particular events, but in general are not applicable to the study of new events. In the end we do not sufficiently understand SEP processes, and we cannot make progress without charge state measurements because they are as critical as the basic quantities of particle abundances and energy distributions.

Benefits and limitations of the cutoff technique

The geomagnetic cutoff technique has several advantages and unique capabilities:

- Energy range over which charge state information is available limited only by the geometry factor of the LEO sensors
- Variations of the cutoff in local time and due to geomagnetic activity can be tracked and accounted for in the analysis by measuring protons (with known ionization state)
- No requirements for any in-situ magnetic field measurements
- Limited impact of pointing requirements on host vehicle (3-axis stabilized with zenith-pointing sensors over the poles and down to ~50 degrees latitude)
- Ion sensors provide accurate information about SEP energy input to the atmosphere whereas models of geomagnetic cutoffs do not adequately reproduce the changing area and extent of polar cap open field lines
- Potential for higher data rates than interplanetary missions because of the proximity of LEO
- Near-real-time capability for space weather applications including high-latitude aircraft and other LEO satellites

The limitations are:

- SEP charge state time dependence limited by orbital period (4 cutoff passes every orbit of ~90 minutes)
- The collection of the time history of the SEP event is also limited by the orbital period unless the sensors are hosted on multiple LEO vehicles

Options for implementation

The SAMPEX payload offers a useful example of what the sensor requirements might be [Baker et al. 1993]. In particular, the geometry factor plays a large role in which events can be studied, wherein instruments on the order of a few to 10's of cm^2sr would be capable of measuring the most intense SEP events. Of course larger apertures would move the sensitivity floor lower to more frequent but less intense events. Elemental composition would be sufficient.

One could implement the technique on a dedicated LEO mission to study the inner magnetosphere charged particle environment, (in which case the sensors have other mission requirements in addition to these), or on a mission of opportunity in polar LEO where the charge state data would be combined with other concurrent diagnostics of the event including the solar and interplanetary contexts.

References

Baker, D.N., Mason, G.M., Figueroa, O., Colon, G.,-Watzin, J., and Aleman, R., An overview of the SAMPEX mission, ***IEEE Trans. Geosc. Elec., 31,531*** (1993).

Mazur, J. E., G. M. Mason, M. D. Looper, R. A. Leske, and R. A. Mewaldt (1999), Charge states of solar energetic particles using the geomagnetic cutoff technique: SAMPEX measurements in the 6 November 1997 solar particle event, *Geophys. Res. Lett.*, *26*(2), 173–176, doi:10.1029/1998GL900075.

Klecker, B. et al., Advances in Space Research,Volume 38, Issue 3, 2006, Pages 493-497

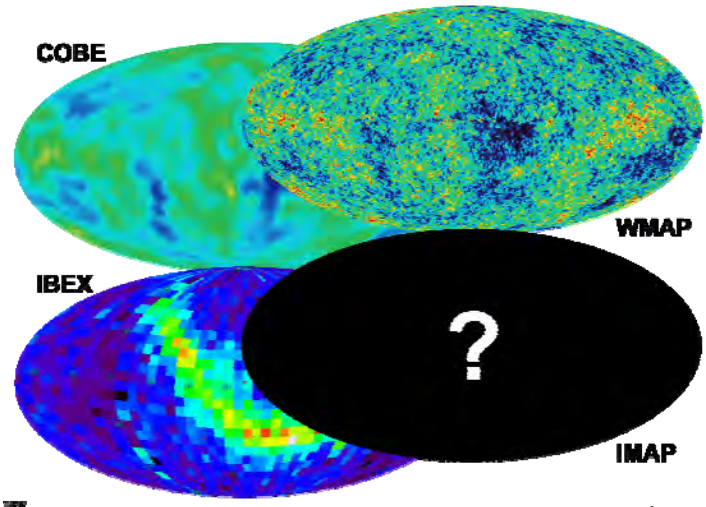
Interstellar MApping Probe (IMAP) mission concept: Illuminating the dark boundaries at the edge of our solar system

Principal Author: D.J. McComas (SwRI)
Co-Authors/Supporters (alphabetical): F. Allegrini (SwRI), M. Bzowski (PAS/SRC), A. Cummings (CIT), M. Desai (SwRI), E. Christian (NASA/GSFC), P. Frisch (UChicago), H. Funsten (LANL), E. Gruen (CU), M. Horanyi (CU), J.U. Kozyra (UMich), H. Kucharek (UNH), S. Lepri (UMich), D.G. Mitchell (APL), E. Möbius (UNH), M. Opher (BU), C.T. Russell (UCLA), N.A. Schwadron (UNH), R. Srama (MPI), M. Wiedenbeck (JPL), P. Wurz (UBern), and G. Zank (UAH)

A. Introduction

Recent groundbreaking all-sky images of the heliospheric boundaries by the Interstellar Boundary Explorer (IBEX) mission and Cassini INCA instrument, in concert with dual point *in situ* observations of the inner heliosheath from both Voyager spacecraft, have made outer heliospheric science one of the most exciting and fastest developing areas in Heliophysics. The **Interstellar MApping Probe (IMAP)** is a mission concept that carries on from the highly successful first heliospheric mapping to enable the discovery of the detailed processes and interaction between the heliosphere and the local interstellar medium (LISM)[1]. IBEX discovered a completely unpredicted, narrow (~20º) ribbon of Energetic Neutral Atom (ENA) emissions coming from the outer heliosphere, which appear to be ordered by the magnetic field just outside in the LISM[2-5]. Even more surprisingly these emissions appear to be evolving on timescales as short as six months[6], demonstrating that the heliosphere's galactic interaction is more dynamic than anticipated. IBEX also made the first *in situ* detections of interstellar H and O, including a secondary O component, which stems from the outer heliosheath and contains information about deflection of the interstellar flow around the heliosphere[7].

Together, the IBEX and INCA ENA images and Voyager observations have proven that our current ideas are truly inadequate to understand the global interaction of our heliosphere. In addition, detailed analyses of interstellar H and O, along with He and Ne, from IBEX are forcing a reexamination of even what we thought we knew about the characteristics of the LISM. Thus, the scientific excitement and broad range of fundamental open issues exposed by these recent observations call out for an advanced and more capable mission – IMAP – which will provide ENA images over a broader energy range and with one to two decades better sensitivity/resolution. In addition, IMAP includes high resolution observations of all the other forms of interstellar matter that can be sampled directly in the inner heliosphere: pickup ions (PUIs), produced by ionization of incoming interstellar neutrals; Galactic Cosmic Rays; Anomalous Cosmic Rays, which are highly energized interstellar neutrals; and cosmic dust, which delivers grains whose chemical and isotopic composition exposes processes and their origins beyond our solar system. Finally, the IMAP concept includes observations of Ly-α backscatter, the interplanetary magnetic field, solar wind ions and electrons, and solar energetic particles, which are used to produce background free ENA observations – the last three of these time-critical space weather observations are provided through a real-time downlink.



What WMAP was to COBE, IMAP will be to IBEX - enabling a "quantum leap" forward in understanding, analogous to that which occurred from COBE to WMAP observations of the cosmic microwave background. With a factor of 10-100 enhanced sensitivity, higher angular and energy resolution and a broader energy range, IMAP will revolutionize our understanding of the heliosphere's interaction with, and our place in, the Galaxy.

All of the great IMAP science can be accomplished in a properly instrumented small-class (250-500M $FY10) mission based on a simple Sun-pointed spinning spacecraft, much like the Advanced Composition Explorer (ACE). As summarized in this white paper, IMAP promises to cost effectively 1) advance critical scientific objectives of understanding our home in the galaxy, 2) contribute to the fundamental understanding of the Sun-Earth system, and 3) produce secondary observations critical to space-weather related societal needs.

B. IMAP Science Concept

IMAP combines ten state-of-the-art measurement capabilities (Table 1) to produce a) critical new observations of ENAs from the heliospheric boundary region over an extended energy range and with significantly enhanced sensitivity, and spatial and energy resolution, compared to prior observations (1-2); b) high sensitivity observations of the other direct samples of interstellar matter available in the inner solar system – interstellar neutral atoms (3), PUIs (4), GCRs and ACRs (5), and dust (6); and c) other supporting measurements required to monitor the solar wind input to the outer heliosheath, as well as understand and mitigate all backgrounds from the local environment around the spacecraft (7-10). This latter element of the IMAP payload will also provide societally important real-time solar wind (RTSW) monitoring.

Table 1. Nominal Payload for IMAP Mission	[kg]	[W]
High Sensitivity, High Resolution ENAs		
1. Solar Wind & PUI ENAs	20	18
2. High Energy ENAs	7	7
Other Samples of Interstellar Matter		
3. Interstellar Neutrals (w Low-E ENAs)	15	15
4. Pickup Ions	10	8
5. ACRs/GCRs	5	5
6. Interstellar Dust	8	16
In Situ/Background Payload with RTSW		
7. Ly-α Photometry	4	4
8. Interplanetary Magnetic Field	3	3
9. Solar Wind Ions & Electrons (two sensors)	7	7
10. Solar/Interplanetary Energetic Particles	3	3
Full Science Payload (CBE)	**80**	**86**
30% Contingency	24	26
TOTAL PAYLOAD ALLOCATION	**104**	**112**

1. Solar Wind & PUI ENAs - The critical energy range from ~0.3-20 keV covers hydrogen ENAs produced from the solar wind and interstellar PUIs, which mass load the solar wind as it travels out through the heliosphere. This energy range also spans the emissions from the IBEX ribbon. Observations of these ENAs with a factor of ~100x the combined sensitivity/duty cycle compared to IBEX-Hi will enable the extreme spatial and energy resolution needed to track the detailed processes and interactions at work in the inner heliosheath and outer heliospheric boundaries as well as fully resolve the time varying structure in the ribbon.

2. High Energy ENAs - The energy range ~3-200 keV covers ENAs that produce the majority of the particle pressure in the heliosheath, and are needed to understand the processes that accelerate solar wind and interstellar pickup ions to higher energies as a function of direction in the sky, where integration path lengths, termination shock strength, and angle to the interstellar field and flow vectors may all play important roles. All-sky ENA images at these energies with a factor of ~10x the angular resolution and ~100x the combined sensitivity/duty cycle of INCA, will resolve time variations on time scales of months, appropriate to the transit times of these particles through regions of interest.

3. Interstellar Neutrals/Low Energy ENAs - IMAP will make direct measurements (5-860 eV) of neutral matter (H, D, He, O, and Ne) from the ISM to determine their abundances, flow patterns (bulk velocity and

temperature), and the filtration of the ISM neutrals in the outer heliosheath. By stepping the viewing direction, both direct trajectories of neutral atoms (before or at the atom's perigee) and indirect trajectories (after their perigee) are observed, providing direct information about the ionization rate that the atoms are exposed to. IMAP will have significantly higher sensitivity for O, Ne, and He than IBEX and high accuracy pointing knowledge (<0.05°) so that it can well separate the primary and secondary neutral populations, illuminating the interactions in the outer heliosphere. Finally, IMAP data will make fundamental measurements the LISM D/H, which requires ~10x the sensitivity of IBEX[8].

4. Pickup Ions - Velocity distributions of PUIs created from atoms at rest, range from zero to twice the solar wind speed. IMAP covers the energy range from 50 eV to 80 keV, which provides good coverage of light PUIs and sufficient coverage of heavier PUIs. In addition, supra-thermal tails in the PUI distribution of H and He will be studied. To infer the complete angular distributions a significant swath of the sky will be viewed (at least 150º x 60º). Finally, this sensor also measures the composition of solar wind ions, which "label" the solar wind sources.

5. Anomalous and Galactic Cosmic Rays – Anomalous cosmic rays (ACRs) are pickup ions that are accelerated in the heliosheath region where most of the ENAs measured by IMAP are formed. Understanding the physics of solar modulation depends on obtaining a long-term record of the temporal variations of both GCRs and ACRs near 1 AU. IMAP measurements utilize stacks of silicon solid-state detectors to measure ACRs and GCRs over the energy range from a few to >100 MeV/nuc. Particle trajectory measurements allow isotopes of He and Ne to be measured. The geometric factor is large enough to provide reasonably accurate monthly energy spectra of the rarer ACR species (e.g., Ar).

6. Interstellar Dust - Interstellar dust particles (ISD) flowing through the heliosphere are expected in the mass range (m) of $10^{-13}<m<10^{-10}$ g with a range of impact speeds (v) $20<v<70$ km/s. Their unique identification requires <10% uncertainty in speed and <5° angular resolution. A dust telescope combining a trajectory sensor and an impact dust instrument will enable the high-fidelity identification of ISD, and the measurement of their chemical and isotopic compositions over $1<M<200$ AMU with a mass resolution of $M/\Delta M>200$. Statistically meaningful observations are provided from an integrated exposure (target area x duration) of >500 cm^2year.

7. Ly-α Backscatter – Ly-α photometry provides a level of global solar wind imaging, including the out-of-ecliptic conditions, which are needed to assess the survival probabilities of both ENAs and interstellar neutrals measured by IMAP. Measurements similar to those from SOHO/SWAN image the neutral interstellar hydrogen distribution within a few AU from the Sun[9-10]. However, IMAP will employ a narrower (~2 Å) filter around the Ly-α line to eliminate the extraheliospheric "chaff" that plagues the SWAN observations. With this, IMAP will be able to provide remote-sensing information on neutral interstellar H supplemental and complementary to those obtained *in situ* via direct detection.

8. Interplanetary Magnetic Field - The local Interplanetary Magnetic Field (IMF) vector direction and magnitude will be measured at a cadence of at least 10 Hz by a magnetometer mounted on a short boom. Measurements cover a wide dynamic range up to several hundred nT (the largest fields observed in the solar wind at 1 AU) with an accuracy better than 0.1 nT.

9. Solar Wind Ions & Electrons - Energy per charge (E/q) ion spectra covering the solar wind beam direction, over the energy range of ~0.1-20 keV/q, are required with a one minute cadence to resolve solar wind structures. Electron energy spectra covering the majority of 4π sr (using the spacecraft spin), over the energy range of ~0.005-2 keV, are required with a cadence of one minute to resolve solar wind structures and suprathermal halo and strahl electron distributions. Such observations are routinely made with electrostatic analyzer based instruments, such as those flown on Ulysses[11] and ACE[12].

10. Solar/Interplanetary Energetic Particles – IMAP makes high sensitivity and temporal resolution (~1 minute) measurements of the differential fluxes of ions and electrons from ~2 keV up to ~1-2 MeV energy covering as much of 4π sr over a large dynamic range (>10^5) to identify the sources of background in the ENA detectors. Similar energetic particle observations are routinely made with solid-state detector based instruments[13-15].

Integrated Theory and Modeling

In addition to the revolutionary new IMAP observations, this mission concept includes an integrated and dedicated theory and modeling effort, which allows the "big picture" of our heliosphere's interaction with our galaxy to be synthesized. One of the possibly outstanding implications of the IBEX discovery of the ribbon is the first accurate estimate of the very local interstellar magnetic field orientation and strength, this derived from large-scale simulations of the global heliosphere coupled to detailed microphysical plasma processes[2,5,16,17]. The confrontation of IBEX observations with prior theory and modeling is forcing an extensive revision of current approaches, reinforcing the critical need to couple kinetic and macroscopic processes in the modeling of complex physical systems.

IMAP, by making measurements over a broader range of energies and with significantly higher sensitivity and resolution, will propel theory and modeling of the heliospheric boundary regions to address the critical microphysical/ kinetic processes and their coupling to the large-scale heliosphere that appear to define so much of the physics of this complex region. Finally, the extension of what we learn from IMAP will expand our knowledge of other astrospheres and the search for habitable zones of extra solar planets.

Real Time Solar Wind – The Magnetospheric Connection

IMAP makes upstream solar wind observations to provide multi-point viewing of the "magnetosphere-scale" features in vast solar wind disturbances impacting Earth and serves as a critical supplier of observations in the era beyond ACE. The ever-changing solar wind and IMF observed by IMAP drive space weather disturbances at Earth. In contrast to tropospheric weather, these drivers are more important than initial conditions in determining the state of the geospace-atmosphere system making IMAP observations critical to holistic studies that promise breakthroughs in understanding the interconnected Sun-Earth system.

Even during quiet solar times, the conditions in the solar wind can conspire to produce space weather that can affect our space systems. An example of this is the recent failure of the Galaxy 15 Communication satellite that occurred when a powerful susbstorm was triggered by a strong solar wind pressure increase when the IMF was already southward. IMAP will observe the solar wind and IMF disturbances that trigger catastrophic reconfigurations of magnetospheric structures with consequences throughout geospace and the upper atmosphere, and feed models, which rely on real-time knowledge of upstream solar wind conditions for prediction. Developing a space weather prediction capability is a national priority [*National Space Weather Program Strategic Plan, FCM-P30-2010, Washington, DC, June 2010*] made more urgent by our growing awareness that complex interdependencies increase the vulnerability of societal infrastructures that are affected by space weather [*NAS, Severe Space Weather Events—Understanding Societal and Economic Impacts Workshop Report, 2008*] and knowledge of the changing upstream environment is an integral component of space situational awareness and treaty verification crucial to our national security.

C. How IMAP Will Advance Solar and Space Physics Science

Understanding the myriad of interactions at the outer boundaries of the solar system is critical for understanding the radiation and space environment throughout our inner solar system. The interstellar interaction leads to the compression of magnetic fields and plasmas in the inner and outer heliosheath. The compressed magnetic fields, in turn, lead to the deflection of the majority of GCRs (~90% at 10-100 MeV) around the heliosheath. These complex interactions cause:

- charge-exchange between incoming interstellar neutral atoms with the plasma of the outer heliosheath. This leads to the depletion (or filtration) of neutral atoms entering the solar system, which, in turn, modifies the composition of the interstellar gas flow entering the heliosphere and thus that of pickup ions born from interstellar neutral atoms. Studies using interstellar pickup ions and/or interstellar neutral atoms to determine the isotopic and elemental composition of the LISM utilize filtration models based on our understanding of the global heliosphere and charge-exchange interactions within the heliosheath. These filtration factors will reach a new level of accuracy once the interstellar boundary conditions of the MHD heliosphere models are better constrained. A precise understanding of the elemental and isotopic composition of the LISM also has fundamental implications for Big Bang cosmology, stellar nucleosynthesis and galactic evolution.

- interstellar pickup ions to form a source of anomalous cosmic rays that are accelerated at the termination shock, through stochastic processes in the heliosheath, or through magnetic reconnection. The study of interstellar pickup ion acceleration to anomalous cosmic ray energies provides a progenitor for understanding particle acceleration in diverse astrophysical environments ranging from the shocks, turbulence and dynamic magnetic structures that surround black holes and supernova remnants.
- modification of the population of interstellar dust grains from the interstellar medium into the heliosphere. The strong magnetic fields of the inner heliosheath cause filtration of interstellar grains, allowing only large grains (>0.2 μm) in. More accurate information about the interstellar interaction will allow us to unfold physical and compositional properties of the interstellar grains in the solar system to understand their sources in the interstellar medium. Finally, the large interstellar dust grains that reach the inner heliosphere create a gravitational focusing tail that links the LISM with dust grains of molecular cloud cores.
- the formation of a neutral population that generates the IBEX ribbon. One of the great discoveries of IBEX is the ENA ribbon nearly circling the heliosphere. While the heliospheric community is still grappling with the source of the ribbon, it is clear that it must be controlled, in part, by the local interstellar magnetic field. Both time-variations in the ribbon and more accurate measurements of fine-structure will yield critically needed insights into its origin, and ultimately help probe the magnetic and plasma properties of the LISM.
- the modulation of Galactic Cosmic Rays. The compression of plasma in the inner heliosheath causes intense magnetic fields that modulate as much as 90% of GCRs at tens of MeV. Understanding how this modulation changes with time is important for assessing how our solar system's radiation environment has changed in the past and how it will change in the future. GCRs pose one of the most stringent factors affecting possibilities of long-term human exploration and may have other effects on life on Earth and even, possibly, terrestrial climate.

Heliophysics has yielded a wealth of information about our magnetosphere, its associated shocks, the similar boundaries that surround other planets, comets, asteroids, the shocks and boundaries surrounding Coronal Mass Ejections (CMEs) as they plow through the heliosphere, and the shocks associated with Co-rotating Interaction Regions (CIRs) in the solar wind. These are all forms of magnetohydrodynamic (MHD) shocks and plasma boundaries that benefit greatly from comparative studies as they inform our understanding of MHD structures that pervade our cosmos and have critically important effects for the formation of energetic particles, the modification of the composition of astrophysical plasmas, the radiation environments of astrospheres and the evolution of planetary environments and atmospheres. Advanced understanding of the plasma boundaries of our heliosphere, the least understood boundaries in Heliophysics, enables generalized understanding of plasma boundaries in the cosmos.

Understanding the time variation of our heliosphere on short and long time-scales has fundamental implications for the history and future of our home in space. By comparing our heliosphere to astrospheres, we may learn how GCR radiation, stellar winds and astrospheres play into the habitability zones surrounding other stars. Our heliospheric boundaries help control, regulate and change the particle and dust populations through the heliosphere: neutral atoms, pickup ions, grains, GCRs and ACRs are all regulated and controlled, in part, by these boundaries. A thorough understanding of the changing Sun and its effects cannot be complete until we develop a comprehensive understanding of how the solar wind and interstellar medium interact to create the complex boundaries and the myriad of associated interactions to form our heliosphere and regulate our space environment. IMAP is a critical next step in the exploration and discovery of our global space environment.

D. Estimated IMAP Cost (FY10 Dollars) – Small Mission

The full IMAP science can be carried out in a straightforward way on a small spacecraft stationed in a halo orbit about the Sun-Earth Lagrangian - L1 point, ~1.5×10^6 km sunward of the Earth. IMAP fits readily into the small-class (250-500M $FY10) mission size because it is based on a simple Sun-pointed spinning spacecraft, similar to the highly successful ACE spacecraft, which has remained in such an orbit for over a decade. Sun-pointed, deploy-once solar arrays provide ample power for all spacecraft subsystems and for

continuous observations by all instruments. The roughly Sun-pointed configuration also provides an optimum configuration for a single high-gain antenna mounted on the aft side of the spacecraft to provide continuous low power, low rate RTSW data in addition to periodic high-rate telemetry of the stored science observations.

From a cost standpoint, the cost for IMAP can also be estimated from the extremely analogous ACE mission, which was launched in August 1997: IMAP has 10 instruments with a nominal payload mass of 104 kg and power of 112 W, ACE had a payload of nine largely similar type instruments of comparable complexity and resources, with resources of 156 kg and 102 W. The IMAP spacecraft is very similar to the ACE spacecraft with the same orbit and essentially identical technical and environmental requirements. ACE was launched on a Delta II which by the time of the IMAP launches should be replaced with a Taurus II. As such, we developed a Rough Order of Magnitude (ROM) cost estimate for IMAP using actual costs from the ACE mission plus the current Launch Vehicle (LV) costs. The cost data are partially reproduced in the Table below. Phase E operations, science, and theory and modeling costs will be highest in the first two years of operations (~$18M/year) and decrease annually (down to ~$6M/year) by the end of the decade of operations to achieve an average ROM cost of $12M/year over 10 years for a well funded, high science return mission.

Table 2. ROM Costs	ACE Pre-Launch Actuals (FY95 M$)[18]	ACE Actuals Inflated to FY10 M$	Estimated IMAP Costs FY10 M$
Project Management	5.1	8.1	10
Spacecraft	47.0	74.8	90
Science Payload	50.2	79.9	100
Ground Systems	1.5	2.4	4
Performance Assurance	1.2	1.9	5
Flight Operations	1.8	2.9	4
Total w/o LV&Phase E	106.8	169.9	213
Phase E Estimate			120
Taurus II Cost		65.0[19]	75
TOTAL IMAP cost w/ Phase E & Taurus LV			**408**
Note that the ACE actual costs[18] are real year and include civil servant cost. We assumed them to be at the mid-year of the development and production or FY95 M$. Using the NASA inflation factor of 1.591, we obtained the cost in FY10 M$ shown in the third column.			

E. How IMAP Meets the Evaluation Criteria

a. IMAP observations are high priority/requirements in previous studies or roadmaps. The fundamental questions addressed by IMAP are of central importance for Heliophysics in our continued quest to understand how the solar wind interacts with the LISM to form our heliospheric boundaries and the plasma and radiation environments contained by them. The related science questions have been highlighted in the previous NRC decadal survey:

The Sun to the Earth – and Beyond: A Decadal Research Strategy in Solar and Space Physics (2002). *Challenge 2: Understanding heliospheric structure, the distribution of magnetic field and matter throughout the solar system, and the interaction of the solar atmosphere with the local interstellar medium.* "The boundary between the solar wind and the local interstellar medium (LISM) is one of the last unexplored regions of the heliosphere. Very little is currently known about this boundary or the nature of the LISM that lies beyond it. … certain aspects of these regions can be studied by a combination of remote sensing and in situ sampling techniques."

b. IMAP makes a significant contribution to more than one of the Panel themes. IMAP is poised to make major advances in our understanding of the heliosphere and its particle, dust and plasma environments. IMAP is equipped with instruments to measure the solar wind, suprathermal populations, energetic particles, magnetic fields, and cosmic rays. With its RTSW data, the mission will also provide a required element for magnetopsheric, ionospheric, thermospheric and mesospheric science, serving as the future L1 monitor.

c. IMAP contributes to important scientific questions facing solar and space physics today. Understanding the global heliosphere and its complex interactions will lead to major advances in understanding particle acceleration at the termination shock, at traveling interplanetary shocks and disturbances, the evolution of MHD structures on global scales, the kinetic physics measured through *in situ* plasma, field, suprathermal and energetic particle measurements. The measurements will continue to provide insights into traveling plasma waves and shocks, the incident structures at the magnetosphere, and reconnection exhausts measured in the solar wind.

d. IMAP contributes to applications and/or policy making related to operations, applications, and societal benefits. IMAP's comprehensive measurements will untangle the influence of changes in the heliosphere on the hazard posed by GCRs to astronauts and provide direct measurements of space weather hazards at L1. The societal relevance of delivering both RTSW and comprehensive measurements is demonstrated by ACE, which is a critical asset in space weather networks.

e. IMAP complements other observational systems or programs available. IMAP complements the operational missions, IBEX, Cassini/INCA and Voyager, by taking the next steps in our exploration of the global heliosphere. IMAP advances the measurement of interstellar neutral atoms and PUI composition, complementing the ongoing ACE and earlier Ulysses missions, by measuring the velocities of interstellar neutrals so that heliospheric data on the ISM can be compared with interstellar data from the Hubble Space Telescope. Finally, IMAP provides a cornerstone to the Heliophysics Great Observatory by providing a critically needed L1 replacement mission.

f. IMAP is affordable, providing great benefit for modest costs. IMAP is highly affordable, and fits well within the small-class mission size. The science per dollar ratio is extremely high. The mission takes fundamental steps in the exploration and discovery of our heliosphere, and provides extensive practical benefits.

g. IMAP has an appropriate degree of technical, resource, and expertise readiness. IMAP requires no significant technical developments. All instruments are natural extensions of existing instruments flown on previous successful 1 AU missions and the spacecraft requirements are very similar to the cost effective ACE spacecraft. The L1 orbit and associated mission operations are well-known and readily achievable within budgeted resources; the mission is categorically low risk.

h. IMAP fits with other national and international plans and activities. The call to understand the complex and interconnected system from the Sun to the interstellar medium is also clearly articulated in the 2010 Science Plan for NASA's Science Mission Directorate:

"An effective plan requires viewing the Sun, heliosphere, and planetary environments as elements of a single interconnected system—one that… evolves in response to solar… and interstellar conditions."

"We know that the Sun, the solar system, and the region of the galaxy just outside the solar system present us with a complex, interacting set of physical processes. But it is also the part of the cosmos accessible to *in situ* scientific investigation; it is a hands-on astrophysical laboratory."

F. Conclusion

Our piece of cosmic real-estate, the heliosphere, is the domain of all human existence. Its history and future in the galaxy is key to understanding the conditions on our evolving planet and future expansion across the solar system. As we ask about the habitability of other planets surrounding other stars, we grapple with understanding the complex environments and interactions in the local parts of the galaxy where these stars exist. Our own heliosphere is an astrophysical case-history of the successful evolution of life in a habitable system. By exploring our global heliosphere and its myriad interactions, we develop key physical knowledge of the interstellar interactions that influence exoplanetary habitability as well as the distant history and destiny of our solar system and world.

G. Footnotes and References

1. Frisch, P.C., et al., Dust in the local interstellar wind, *The Astrophysical Journal, 525,* 492-516, 1999
2. McComas, D.J., et al., Global observations of the interstellar interaction from the Interstellar Boundary Explorer (IBEX), *Science, 326,* doi: 10.1126/science.1180906, 959-962, 2009.
3. Fuselier, S.A., et al., Width and variation of the ENA flux ribbon observed by the Interstellar Boundary Explorer, *Science, 326,* doi: 10.1126/science.1180981, 962-964, 2009.
4. Funsten, H.O., et al., Structures and spectral variations of the outer heliosphere in IBEX energetic neutral atom maps, *Science, 326,* doi: 10.1126/science.1180927, 964-966, 2009.
5. Schwadron, N.A., et al., Comparison of Interstellar Boundary Explorer observations with 3D global heliospheric models, *Science, 326,* doi: 10.1126/science.1180986, 966-968, 2009.
6. McComas, D.J., et al., The evolving outer heliosphere: Large-scale stability and time variations observed by the Interstellar Boundary Explorer, *J. Geophys. Res., 115,* A09113, doi: 10.1029/2010JA015569, 2010.
7. Möbius, E., et al., Direct observations of interstellar H, He, and O by the Interstellar Boundary Explorer, *Science, 326,* doi: 10.1126/science.1180971, 969-971, 2009.
8. Tarnopolski and M. Bzowski, Detectability of neutral interstellar deuterium by a forthcoming SMEX mission IBEX, *Astr. Ap., 483, No 3,* pp L35-L38, 2008.
9. Bzowski et al., Latitudinal structure and north-south asymmetry of the solar wind from Lyman-α remote sensing by SWAN, *Astr. Ap., 408,* pp 1165-1177, 2003.
10. Quemerais et al., Interplanetary hydrogen absolute ionization rates: retrieving the solar wind mass flux latitude and cycle dependence with SWAN/SOHO maps, *J. Geophys. Res.,* 111, A09114, 10.1029/2006JA011711, 2006.
11. Bame, S.J., et al., The Ulysses Solar Wind Plasma Experiment, *Astron. Astrophys. Suppl. Ser. 92,* 237-265, 1992.
12. McComas, D.J., et al., Solar wind electron proton alpha monitor (SWEPAM) for the Advanced Composition Explorer, *Space Sci. Rev., 86,* 563-612, 1998.
13. Lanzerotti, L.J., et al., Heliosphere instrument for spectra, composition and anisotropy at low energies, *Astronomy and Astrophysics Supplement Series (ISSN 0365-0138), vol. 92, no. 2,* Jan. 1992, p. 349-363, 1992.
14. Gold, R.E., et al., Electron, proton, and alpha monitor on the Advanced Composition Explorer spacecraft, *Space Science Reviews, 86, Issue 1/4,* p. 541-562, 1998.
15. Lin, R. P., et al., The STEREO IMPACT Suprathermal Electron (STE) Instrument, *Space Science Reviews, 136, Issue 1-4,* pp. 241-255, DOI: 10.1007/s11214-008-9330-7, 2008.
16. Heerikhuisen, J., et al., Pick-up ions in the outer heliosheath: A possible mechanism for the Interstellar Boundary EXplorer Ribbon, *Astrophys. J. Letters, 708,* L126-L130, doi:10.1088/2041-8205/708/2/L126, 2010.
17. Chalov, S.V., et al., Scatter-free pickup ions beyond the heliopause as a model for the Interstellar Boundary Explorer (IBEX) ribbon, *Astrophys. J., 716,* L99-L102, doi:10.1088/2041-8205/716/2/L99, 2010.
18. "Advanced Composition Explorer (ACE) Lessons Learned and Final Report", July 1998, NASA/Goddard Space Flight Center.
19. "Taurus II Development Status of a Medium-Class Launch Vehicle for ISS Cargo and Satellite Delivery", David Steffy, Orbital Sciences Corporation, 8th NRO/AIAA Space Launch Integration Forum, 15 - 16 Jul 2008.

X-Ray and Gamma-Ray Polarimetry of Solar Flares

White paper Submitted to the
Solar and Heliospheric Physics Decadal Survey

November 12, 2010

prepared by
Mark L. McConnell (UNH)

with support from
Peter F. Bloser (UNH)
Brian Dennis (NASA/GSFC)
A. Gordon Emslie (WKU)
Robert P. Lin (UCB)
James M. Ryan (UNH)
Albert Y. Shih (NASA/GSFC)
David M. Smith (UCSC)

BRIEF DESCRIPTION

A determination of the extent to which flare accelerated electrons are beamed constitutes an essential step towards a greater understanding of particle acceleration in solar flares. Robust high energy polarization measurements offer the most effective means of measuring the electron beaming.

INTRODUCTION

Solar flares represent a process of explosive energy release in a magnetized plasma, a process which is believed to take place at many other sites in the universe. Solar flares accelerate ions up to tens of GeV and electrons to hundreds of MeV, releasing as much as 10^{33} ergs in the process (see, e.g., Lin & Hudson 1976). The accelerated 10–100 keV electrons appear to contain a significant fraction, perhaps the bulk, of this energy, indicating that the particle acceleration and energy release processes are intimately linked. The details of how the Sun releases this energy, presumably stored in the magnetic fields of the corona, and how it rapidly accelerates electrons and ions with such high efficiency, and to such high energies, is presently unknown. Furthermore, the electrical currents associated with the accelerated electrons impose formidable constraints on the global electrodynamics of the system (Emslie & Hénoux 1995). These problems can be alleviated somewhat if the particle acceleration is nearly isotropic, such as in stochastic acceleration models (e.g., Miller et al. 1996). Furthermore, solar flare radiation spectra are dependent not only on the energy distribution, but also on the angular distribution of the energetic particles. Consequently, independent measurements of the angular distribution of the energetic particles may be important for proper interpretation of the radiation spectra. *A determination of the extent to which the accelerated electrons are beamed (anisotropic) therefore constitutes an essential step towards a greater understanding of particle acceleration in solar flares and, more generally, throughout the cosmos.*

The angular distribution of accelerated *ions* can be studied by measuring the energies and widths of broad γ-ray lines. There are several ways in which the anisotropy of the accelerated *electrons* can be ascertained from the high-energy photon emissions. However, statistical directivity studies (e.g., Kašparová et al. 2007) are inconclusive and multi-spacecraft stereoscopic observations (e.g., Kane et al. 1980) are sparse. *High energy polarimetry is the most effective means of measuring the anisotropy of the accelerated electrons.* It provides the ability to derive the anisotropy of the accelerated electrons from measurements by a single instrument and measurements with sufficient temporal resolution could even measure variations of the electron anisotropy during the evolution of the flare. Unfortunately, *reliable measurements of the polarization of solar flare X-ray and gamma-ray radiation lag well behind theoretical predictions of this key diagnostic.*

We are finally seeing the onset of observations capable of testing the numerous theoretical predictions of high energy polarization that have been in the literature for decades. The few existing polarization measurements are intriguing and inconclusive, yet collectively they suggest that the magnitude of the polarization vector is of the order predicted by models that have a strong anisotropy of the emitting electrons. However, the measured orientation of this vector may be in a direction substantially different from the local solar radial, as predicted by most solar models. This raises the fascinating possibility that significant refinements of our models for particle acceleration and transport in solar flares may be required.

The importance of polarization measurements at high energies has often been recognized. For example, the High Energy from Space Panel of the Astronomy and Astrophysics Survey Committee (1991) pointed out the possibilities for polarization measurements and suggested that the design of "various types of polarimeters" be considered as one useful area for basic technology development. In April of 1997, NASA's Gamma Ray Astronomy Working Group (GRAPWG), in their report titled *Recommended Priorities for NASA's Gamma-Ray Astronomy Program, 1996-2010*, states that one "scientifically interesting" solar mission would be "to design an instrument to attack a specific problem such as the polarization of the electron bremsstrahlung radiation."

1.0 SCIENTIFIC MOTIVATION

Studies of γ-ray line data from the SMM Gamma Ray Spectrometer (GRS) suggest that protons and α-particles are likely being accelerated in a rather broad angular distribution (Share & Murphy, 1997; Share et al., 2002). There is no reason to expect, however, that electrons are being accelerated in a similar fashion. Here we review two possible means of measuring the accelerated electron angular distribution: 1) by measuring photon directivity at X-ray and gamma-ray energies; and 2) by measuring X-ray and gamma-ray polarization. We argue that high energy polarimetry is the preferred approach.

1.1 Photon Directivity Measurements as a Probe of Electron Beaming

An anisotropic ensemble of bremsstrahlung-producing electrons will produce a radiation field that is not only polarized, but anisotropic. Measurements of the high energy photon directivity can therefore provide a probe of the extent to which the accelerated electrons are beamed. One technique for studying directivity on a statistical basis is to look for center-to-limb variations. Correlations between flare longitude and flare intensity or spectrum reflect the anisotropy of the X-ray emission and hence any directivity of the energetic electrons. For example, if the radiation is preferably emitted in a direction parallel to the surface of the Sun, then a flare located near the limb will look brighter than the same flare near the disk center. Analysis of SMM GRS data collected during cycle 21 (for $E > 300$ keV) provided the first clear evidence for directed emission, with a tendency for the high energy events to be located near the limb (Vestrand et al., 1987; Bai, 1988). Observations from SMM GRS during cycle 22 provided further support for directivity (Vestrand et al., 1991). However, several high energy events were also observed near the disk center by a number of different experiments during cycle 22 (e.g., on GRANAT and CGRO; see Vilmer 1994 for a summary), perhaps suggesting a more complex pattern. *Significant differences from one flare to the next in terms of geometry, energy release, time variability, etc., make quantifying the magnitude of the directivity from these statistical measurements a difficult task.*

Another method for studying the directivity in individual flares is the stereoscopic technique (Catalano & van Allen 1973). This method compares simultaneous observations made on two spacecraft that view the same flare from different directions. Several efforts to make stereoscopic measurements (using, for example, simultaneous data from PVO and ISEE-3) have failed to produce clear evidence for directivity (Kane et al. 1988; Li et al. 1994; Kane et al. 1998). *A potential problem with these data is that stereoscopic observations tend to suffer from cross-calibration issues between different instruments.*

The difficulties of statistical and stereoscopic observations for measuring photon directivity suggest the need for an alternative technique that can measure time-dependent particle anisotropies in individual flares. Polarization is a diagnostic that can meet these requirements.

1.2 Polarization Measurements as a Probe of Electron Beaming

Using polarization measurements at high energies for determining the accelerated electron angular distribution is possible because the emission from any bremsstrahlung source (such as a solar flare) will be polarized if the phase-space distribution of the emitting electrons is anisotropic. Electrons accelerated in the flare spiral around guiding magnetic field lines, emitting high energy photons from collisions with ambient protons and heavier ions. In conventional models of a solar flare, the magnetic field forms a loop structure that penetrates the chromosphere in a vertical direction (Figure 1). Since most of the high energy emission is emitted in the dense chromospheric regions of the loop, the direction of the magnetic field in these layers represents a preferred direction in the source. Hence, one expects the emission to be linearly polarized either in, or perpendicular to, the plane defined by this preferred (vertical) direction and the direction to the observer. This plane intersects the visible solar disk in the radial direction from the center of the disk to the source. Extensive modeling by many authors (e.g., Brown 1972; Hénoux 1975; Langer & Petrosian 1977; Bai & Ramaty 1978; Leach & Petrosian 1983; Zharkova et al., 1995; Charikov et al., 1996) shows that, for the generally-accepted "thick-target" model, in which the accelerated electrons are preferentially accelerated downward, the polarization vector is parallel to this reference plane, i.e., radial on the solar disk. Emslie & Brown (1980) and Emslie, Bradsher, & McConnell (2008) have shown, however, that non-radial polarization orientations are possible both for thermal models of the high energy emission (in which non-vertical source regions are involved), or for deviations of the chromospheric magnetic field direction from the local vertical. Model predictions for the polarization fraction range from a few percent (in cases where, for example, a rapidly converging magnetic field changes the pitch angle of the electrons to a more horizontal, "pancake-type," distribution, or where the accelerated electrons have a wide range of initial pitch angles), to some tens of percent (e.g., for highly-directed acceleration in near-uniform guiding fields). Thus, *measurement of the magnitude and direction of the polarization vector provide key diagnostic evidence to distinguish between particle acceleration scenarios.*

Some fraction of the flux at energies below 100 keV will be flux that is backscattered from the solar photosphere. For spatially-integrated (non-imaging) measurements, the reflected component will influence polarization results, since backscattering will tend to introduce polarization fractions of a few percent at energies below 100 keV (e.g., Langer & Petrosian, 1977; Bai & Ramaty, 1978). Even thermal models of the X-ray source predict a finite polarization of order a few percent (Emslie & Brown, 1980).

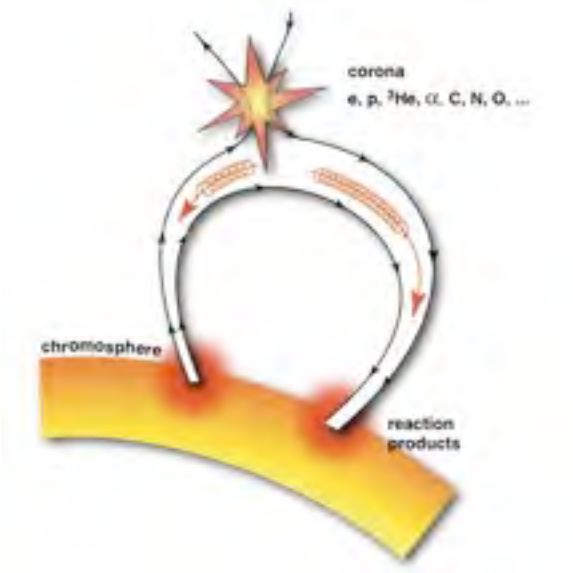


Figure 1 - High energy polarimetry offers a unique means of probing the angular distribution of the electrons that are accelerated in solar flares.

The thermal component, with its rather low polarization, tends to dominate the emission from all flares at energies below about ~25 keV. Imaging measurements will be required at these energies to ensure separation of the thermal and non-thermal components

At non-thermal energies, predictions for solar flare polarization are often criticized on the grounds that the modeling assumptions they contain are overly simplified. For example, most models typically assume a single, simple magnetic field structure. It could be argued that any real flare, particularly one sufficiently large to produce a signal of sufficient strength to enable a polarization measurement, will likely contain a mix of structures that would average out any polarization signal (see also Hudson et al., 2003), suggesting the need for imaging polarimetry. However, hard X-ray imaging observations in the impulsive phase generally show a fairly simple geometry, consisting of two footpoint sources and perhaps a loop-top source (e.g., Sakao et al., 1992; Masuda et al., 1995). These observations suggest that simple magnetic structures are responsible for the energetic emissions and give support to the possibility that a statistically significant polarization signal could be produced in a large event, even in the case of non-imaging measurements.

1.3 Efforts to Measure Polarization

The earliest efforts to measure X-ray polarization from solar flares were all obtained at energies below ~15 keV (Tindo et al. 1970, 1972a, 1972b; Nakada et al., 1974; Tindo et al. 1976; Tramiel et al., 1984). Some of these observations have been met with considerable skepticism (e.g., Brown, McClymont, & McLean 1974). Collectively, these measurements do not present a consistent picture of solar flare polarization.

Recent efforts to measure polarization at higher energies ($>$ 20 keV) have been made using data from the Ramaty High Energy Solar Spectroscopic Imager (RHESSI; Lin et al. 2002; Smith et al. 2002). Although designed primarily as a hard X-ray imager and spectrometer, RHESSI is also capable of measuring the polarization of hard X-rays and γ-rays from solar flares using two distinct methods. Polarimetry of low and medium energies (20-100 keV) is made possible by the inclusion of a small unobstructed Be scattering element that is strategically located within the array of nine segmented germanium detectors (McConnell et al. 2002). Solar flare photons below ~100 keV can reach a rear segment of a Ge detector only indirectly, by scattering. The azimuthal distribution of photons scattered from the Be carries with it a signature of the linear polarization of the incident flux. Sensitivity estimates suggest that a polarization sensitivity of less than a few percent can be achieved for X-class flares. Initial results from the X4.3 flare of 23-July-2003 show some significant modulation of the Be-scattered flux, but it is not yet clear whether this modulation results from polarization of the incident flux or from instrumental systematics (McConnell et al. 2003). A second approach to doing polarimetry with RHESSI (at higher energies, above ~200 keV) uses the scattering of photons between the Ge detectors. Boggs et al. (2005) reported marginal evidence for polarization in two X-class flares using this technique. Suarez-Garcia et al. (2006) also used Ge-Ge scatter events to study six X-class and one M-class flare in the energy range from 100-350 keV, but obtained no significant detection of polarization. To date, none of the RHESSI results provide unambiguous evidence for polarization.

Additional attempts to measure polarization have recently been made using data from the SPR-N instrument on the Coronas-F spacecraft (e.g., Veselovsky et al. 2004; Bogomolov et al. 2003a, 2003b; Zhitnik et al. 2006). The SPR-N instrument is sensitive to polarization

from 10-100 keV, using passive Be scattering elements, coupled with active scintillators, to detect and measure scattered photons. The effective area ranges from ~0.5 cm^2 (20 keV) to ~1.5 cm^2 (100 keV). Zhitnik et al. (2006) reported the detection of hard X-rays from more than 90 flares, only one of which (that of of October 29, 2003) showed significant levels of polarization, at levels in excess of 50% (with large uncertainties). Upper limits for 25 other events were in the range of 8% to 40%.

2.0 INSTRUMENTATION

Various techniques have been developed for measuring the polarization of hard X-rays and γ-rays. Here we review three different approaches that collectively cover the full range of X-ray and gamma-ray energies from a few keV up to energies above 1 MeV. All three approaches are currently being considered for use in solar flare measurements. In addition, all three techniques could be applied to the development of an imaging instrument. With sufficient angular resolution, spatially-resolved polarization measurements could study the particle distribution at different points along a flaring loop, providing further insight into the evolution of the particle distribution as it propagates along the field lines.

2.1 Low-Energy Techniques (< 50 keV)

Although Compton (Thomson) scattering has been used as a means to measure polarization at these energies, in recent years the photoelectric effect has been exploited for a number of polarimeter designs. In the photoelectric effect, the ejected photoelectron tends to be scattered parallel to the incident electric field (polarization). Photoelectric photometry relies on the ability to track the direction of the photoelectron and thus providing a measurement of the incident photon's polarization. Polarization measurements that exploit this process have recently become practical with the development of a highly sensitive measurement technique based on imaging the photoelectron track using a micropattern gas detector (MPGD). The sensitivity of MPGD polarimeters can be optimized by implementing a time projection readout scheme (Black et al. 2003). These time-projection chamber (TPC) polarimeters have been demonstrated as sensitive polarimeters in the 2–10 keV band and have also been optimized for higher energies. This technology is also the basis of the instrument on GEMS SMEX mission, currently under development for studying the 2-10 keV polarization of astrophysical sources. The Imaging X-ray Polarimeter for Solar flares (IXPS) represents one concept for performing polarization measurements in the 20-50 keV energy range, utilizing a TPC for tracking the photoelectron (Dennis et

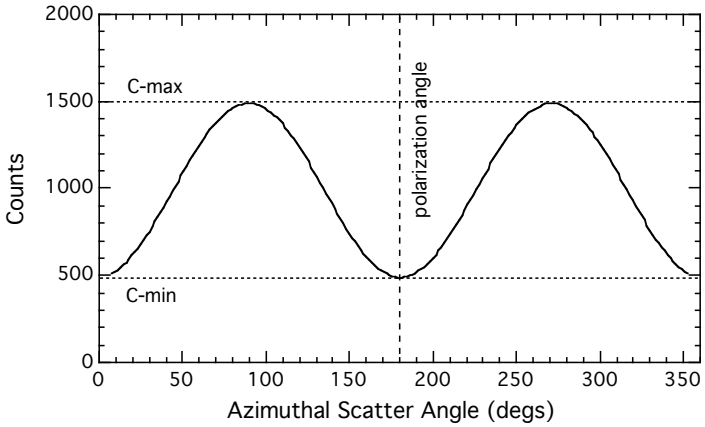


Figure 2 - The azimuthal scatter angle distribution of Compton scattered photons that are initially polarized. This represents the polarization signature of a Compton scatter polarimeter, where the photons tend to scatter perpendicular to the incident polarization direction. In a photoelectric polarimeter, a similar distribution is obtained for the distribution of emitted photoelectrons, except that the photoelectrons tend to scatter parallel to the incident polarization direction.

al. 2010). Imaging capability will be achieved by measuring the spatial Fourier components of the X-ray flare as is done with RHESSI.

2.2 Medium Energy Techniques (50-500 keV)

The physical process used to measure polarization in the 50-500 keV energy range is Compton scattering. Compton scattered photons tend to be scattered at right angles with respect to the incident electric field. In the case of an unpolarized beam of incident photons, there will be no preferred azimuthal scattering angle; the azimuthal distribution of scattered photon angles will therefore be uniform. However, in the polarized case, the incident photons will exhibit an asymmetric azimuthal distribution. Polarimetry in the 50-500 keV energy band requires low-Z scattering elements coupled with high-Z photon absorbers for achieving the best result. The Gamma-RAy Polarimeter Experiment (GRAPE; McConnell et al. 2009) is a scintillator-based Compton polarimeter designed to observe polarized astrophysical phenomena in the hard X-ray energy band (50 – 500 keV). The basic instrument concept has been validated in laboratory experiments and during a calibration campaign at a polarized X-ray beam (Bloser et al. 2006), and the hardware has been flight-tested on an engineering balloon flight. A large array of polarimeter modules is currently being prepared for a balloon flight in the fall of 2011, with the goal of measuring the polarization of the Crab (Connor et al. 2010). Subsequent long-duration balloon flights will permit observations of both solar flares and gamma-ray bursts. The GRAPE design lends itself readily to imaging with Rotation Modulation Collimators, as is currently done with RHESSI.

2.3 High Energy Techniques (>300 keV)

Compton scattering is also an effective way to measure polarization in the gamma-ray range (at energies extending up to 1 MeV or more). Newly developed 3D position-sensitive germanium detectors (3D-GeDs) combine the high spectral resolution of germanium with the capability of spatially resolving each energy deposition to within <0.1 mm^3. Using this information, it is possible to reconstruct the path of each gamma ray as it Compton scatters around the volume, and thus precisely measure the anisotropic scattering from a polarized source. Furthermore, the track reconstruction allows for the rejection of background events and the identification of incomplete depositions that both can reduce the polarization sensitivity. The Gamma-Ray Imager/Polarimeter for Solar flares (GRIPS) is a balloon-borne instrument that combines 3D-GeDs with a novel imaging technique using a multi-pitch rotating modulator (MRPM) to produce gamma-ray images with an unprecedented resolution of 12.5 arcsec (Shih et al. 2009). With the spectrometer acting as its own active scatterer, GRIPS thus has the capability for imaging polarimetry with sufficient fluxes. GRIPS is currently being built and is slated for a one-day test flight in the spring of 2012. Long-duration Antarctic balloon flights in following years are anticipated to fully test out the instrument and return science. Future instruments based on the GRIPS design can have increased angular resolution and sensitivity.

3.0 Summary

X-ray and gamma-ray emission from solar flares, like any other form of electromagnetic radiation, has four, and only four properties. Each photon can be completely characterized by its time of arrival, its energy, its direction of arrival, and its polarization state. To date, our understanding of high energy solar flare emissions has come from measurements of the

first three of these quantities. Efforts to study the fourth quantity (polarization) have so far met with only very limited success. The study of polarization at X-ray and gamma-ray energies is especially appealing in that the high-energy emission from any bremsstrahlung source (such as a solar flare) will be polarized if the phase-space distribution of the emitting electrons is anisotropic. Polarization measurements therefore provide a direct handle on the extent to which the accelerated electrons are beamed, which, in turn, has important implications for particle acceleration models.

Observations to date are highly inconclusive, in no small part because they have been made using instruments that were not optimized for the measurement of polarization. With the increasing realization that polarization measurements provide uniquely insightful information on the physical processes at work in particle acceleration sites throughout astrophysics, the time is propitious to develop instruments specifically tailored to the measurement of solar flare polarization at high energies. This will constitute a quantum leap in our capabilities to resolve many of the current ambiguities and outstanding questions in solar flare physics.

The primary goal of any future high energy polarimetry experiment must be to improve upon the sensitivity of current instrumentation. This can be achieved either by incorporating high energy polarimetry as one aspect of an instrument with a broad range of capabilities (such as GRIPS) and/or through the development of specialized instrumentation (like GRAPE or IXPS) that could be included as part of a suite of instruments on a major observatory. *In either case, high quality data collected during periods of maximum solar activity promise to provide important insights into the solar flare acceleration process.* Polarimetry would yield direct information on the directivity (i.e., anisotropy) of the emitting electron population and so provide critical constraints on the acceleration mechanism. Imaging polarimetry would reveal the characteristic polarimetric signature of the photospheric albedo patch, clearly distinguishing it from the direct X-ray flux, which, in general has a significantly different spatial distribution of the magnitude and orientation of the polarization vector.

REFERENCES

Bai, T. and R. Ramaty (1978). "Backscatter, anisotropy, and polarization of solar hard X-rays." Ap. J. **219**, 705-726.

Bai, T. (1988). "Directionality of continuum gamma rays from solar flares." Ap. J. **334**, 1049-1053.

Black, J. K., et al. (2003). “X-ray polarimetry with an active-matrix pixel proportional counter.” Nucl. Instr. Meth. A, 513, 639.

Bloser, P. F., et al. (2009). “Calibration of the Gamma-RAy Polarimeter Experiment (GRAPE) at a polarized hard X-ray beam.” Nuc. Instr. Meth. A 600, 424-433.

Boggs, S. E., et al. (2006). "Gamma-Ray Polarimetry of Two X-Class Solar Flares." Ap. J. **638**, 1129-1139.

Bogomolov, A. V., et al. (2003a). "Hard X-ray Radiation from Solar Flares in the Second Half of 2001: Preliminary Results of the SPR-N Experiment Onboard the Coronas-F Satellite." Solar System Research **37**, 112-120.

Bogomolov, A. V., et al. (2003b). "Parameters of the Intense X-ray and Gamma-ray Radiation from the Solar Flare of May 20, 2002, as Observed from the CORONAS-F Spacecraft." Astron. Lett. **29**, 199-204.

Brown, J. C. (1972). "The Directivity and Polarisation of Thick Target X-Ray Bremsstrahlung from Solar Flares." Solar Physics **26**, 441.

Brown, J. C., A. N. McClymont, et al. (1974). "Interpretation of solar hard X-ray burst polarization measurements." Nature 247: 448-449.

Catalano, C. P. and J. A. van Allen (1973). "Height Distribution and Directionality of 2-12 a X-Ray Flare Emission in the Solar Atmosphere." Ap. J. **185**, 335-350.

Charikov, J. E., et al. (1996). "Hard X-ray emission of solar flares and non-stationary kinetics of electron beams." Astron. Astrophys. **308**, 924-928.

Dennis, B., et al. (2010). “Imaging X-Ray Polarimeter for Solar Flares (IXPS).” In preparation.

Connor, T. P., et al. (2010). “Plans For The First Balloon Flight Of The Gamma-Ray Polarimeter Experiment (GRAPE).” Proc. SPIE, 7732, 139.

Emslie, A. G. and J. C. Brown (1980). "The polarization and directivity of solar-flare hard X-ray bremsstrahlung from a thermal source." Ap. J. **237**, 1015-1023.

Emslie, A. G. and J.-C. Henoux (1995). "The Electrical Current Structure Associated with Solar Flare Electrons Accelerated by Large-Scale Electric Fields." Astrophysical Journal 446: 371.

Emslie, A. G., Bradsher, H. L., and McConnell, M. L. (2008). “Hard X-Ray Polarization from Non-Vertical Solar Flare Loops.” Astrophysical Journal 674, in press.

Henoux, J. C. (1975). "Anisotropy and polarization of solar X-ray bursts." Solar Physics 42: 219-233.

Hudson, H. S., et al. (2003). "Coronal scattering as a source of flare-associated polarized hard X-rays." Solar Physics **214**, 171-175.

Kane, S. R., K. A. Anderson, et al. (1980). "Directivity of 50-100 keV X-ray emission from impulsive solar flares." Astrophysical Journal 239: L85-L88.

Kane, S. R., et al. (1988). "Directivity of 100 keV-1 MeV photon sources in solar flares." Ap. J. **326**, 1017-1031.

Kane, S. R., et al. (1998). "Stereoscopic Observations of Solar Hard X-Ray Flares Made by ULYSSES and YOHKOH." Ap. J. **500**, 1003.

Kasparov·, J., E. P. Kontar, et al. (2007). "Hard X-ray spectra and positions of solar flares observed by RHESSI: photospheric albedo, directivity and electron spectra." Astronomy and Astrophysics 466: 705-712.

Langer, S. H. and V. Petrosian (1977). "Impulsive solar X-ray bursts. III - Polarization, directivity, and spectrum of the reflected and total bremsstrahlung radiation from a beam of electrons directed toward the photosphere." Ap. J. **215**, 666-676.

Leach, J. and V. Petrosian (1983). "The impulsive phase of solar flares. II - Characteristics of the hard X-rays." Ap. J. **269**, 715-727.

Li, P., et al. (1994). "Directivity of 100-500 keV solar flare hard X-ray emission." Ap. J. **426**, 758-766.

Lin, R. P., et al. (2002). "The Reuven Ramaty High-Energy Solar Spectroscopic Imager (RHESSI)." Solar Physics **210**, 3-32.

Masuda, S., et al. (1995). "Hard X-Ray Sources and the Primary Energy-Release Site in Solar Flares." Publications of the Astronomical Society of Japan **47**, 677-689.

McConnell, M. L., et al. (2002). "RHESSI as a Hard X-Ray Polarimeter." Solar Physics **210**, 125-142.

McConnell, M. L., et al. (2003). "Evidence for Hard X-Ray Polarization from the Flare of 23 July 2002." Bulletin of the American Astronomical Society 34, 850.

McConnell, M. L., et al. (2009). “GRAPE: A Balloon-Borne Gamma-Ray Polarimeter.” Proc. SPIE, 7435, 13.

Miller, J. A., T. N. Larosa, et al. (1996). "Stochastic Electron Acceleration by Cascading Fast Mode Waves in Impulsive Solar Flares." Astrophysical Journal 461: 445.

Nakada, M. P., et al. (1974). "Polarization results of solar X-rays from OSO-7." Solar Physics **37**, 429-435.

Sakao, T., et al. (1992). "Hard X-ray imaging observations by YOHKOH of the 1991 November 15 solar flare." Publications of the Astronomical Society of Japan **44**, L83-L87.

Share, G. H. and R. J. Murphy (1997). "Intensity and Directionality of Flare-accelerated alpha -Particles at the Sun." Ap. J. **485**, 409.

Share, G. H., et al. (2002). "Directionality of Solar Flare-accelerated Protons and alpha-Particles from gamma-Ray Line Measurements." Ap. J. **573**, 464-470.

Shih, A. Y., et al. (2009). “The Gamma-Ray Imager/Polarimeter for Solar flares (GRIPS).” Bulletin of the American Astronomical Society 41, 846

Smith, D. M., et al. (2002). "The RHESSI Spectrometer." Solar Physics **210**, 33-60.

Suarez-Garcia, E., et al. (2006). "X-Ray Polarization of Solar Flares Measured with Rhessi." Solar Physics **239**, 149-172.

Tindo, I. P., et al. (1970). "On the Polarization of the Emission of X-Ray Solar Flares." Solar Physics **14**, 204.

Tindo, I. P., et al. (1972a). "Preliminary Interpretation of the Polarization Measurements Performed on 'Intercosmos-4' during Three X-Ray Solar Flares." Solar Physics **27**, 426.

Tindo, I. P., et al. (1972b). "New Measurements of the Polarization of X-Ray Solar Flares." Solar Physics **24**, 429.

Tindo, I. P., et al. (1976). "The polarization of X-ray emission of some solar flares in July 1974." Solar Physics **46**, 219-227.

Tramiel, L. J., et al. (1984). "Polarization evidence for the isotropy of electrons responsible for the production of 5-20 keV X-rays in solar flares." Ap. J. **280**, 440-447.

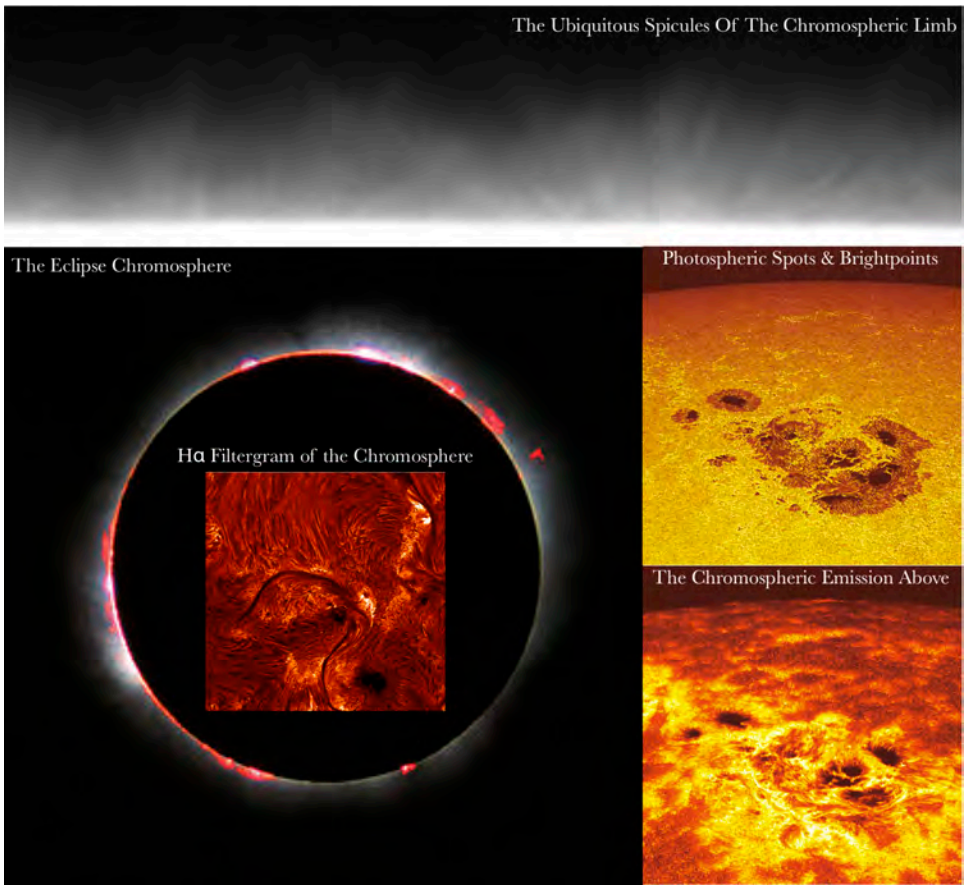
Veselovsky, I. S., et al. (2004). "Solar and Heliospheric Phenomena in October-November 2003: Causes and Effects." Cosmic Research **42**, 435-488.

Vestrand, W. T., et al. (1987). "The directivity of high-energy emission from solar flares - Solar Maximum Mission observations." Ap. J. **322**, 1010-1022.

Vestrand, W. T., et al. (1991). "New Measurements of the Anisotropy of Solar Flare Gamma-Rays." International Cosmic Ray Conference 3, 69.

Zharkova, V. V., et al. (1995). "Electron beam dynamics and hard X-ray bremsstrahlung polarization in a flaring loop with return current and converging magnetic field." Astronomy and Astrophysics **304**, 284.

Zhitnik, I. A., et al. (2006). "Polarization, temporal, and spectral parameters of solar flare hard X-rays as measured by the SPR-N instrument onboard the CORONAS-F satellite." Solar System Research **40**, 93-103.



The Solar Chromosphere: The Inner Frontier of the Heliospheric System

A White Paper Submitted to the Heliophysics 2012 NRC Decadal Survey

S. W. McIntosh (High Altitude Observatory; corresponding author)

T. Ayres (University of Colorado Boulder)
T. Bastian (National Radio Astronomy Observatory)
T. Berger, B. De Pontieu, C. Schrijver (Lockheed Martin Solar and Astrophysics Laboratory)
G. Cauzzi, K. Reardon (INAF - Arcetri Astrophysical Observatory)
M. Carlsson, V. Hansteen (University of Oslo)
C. DeForest (Southwest Research Institute)
D. Gary (New Jersey Institute of Technology)
S. Keil, A. Tritschler, H. Uitenbroek (National Solar Observatory)
R. J. Rutten, J. Leenaarts (University of Utrecht)
S. Solanki (Max Planck Institute for Solar System Research)
S. White (Space Vehicles Directorate, AFRL)

Executive Summary: The highly intricate, dynamic, phenomena-rich solar chromosphere results from the profound interaction of plasma motion and magnetic fields. The fruitful combination of high resolution multi-wavelength imaging and polarimetry with detailed numerical simulations offer the promise to unlock the mysteries of this crucial interface between the photosphere and corona - as it did so successfully in the study of photospheric granular convection. Understanding the solar chromosphere will open a new window onto mass and energy transport at the crucial innermost boundary of the heliospheric system, from which the corona and solar wind are powered, and the bulk of the Earth-impacting UV radiation is released, all key players in “space weather.” Recent momentum in the development of new instruments, theory, and facilities must be recognized, encouraged, and sustained to fully exploit ongoing four-dimensional observational dissection of the “magnetic complexity zone.” We are at a time when comprehensive modeling of chromospheric physics is tractable and reaching maturity. As a result, we may be able to slide this crucial, but currently missing, piece of the picture into the grand heliospheric puzzle.

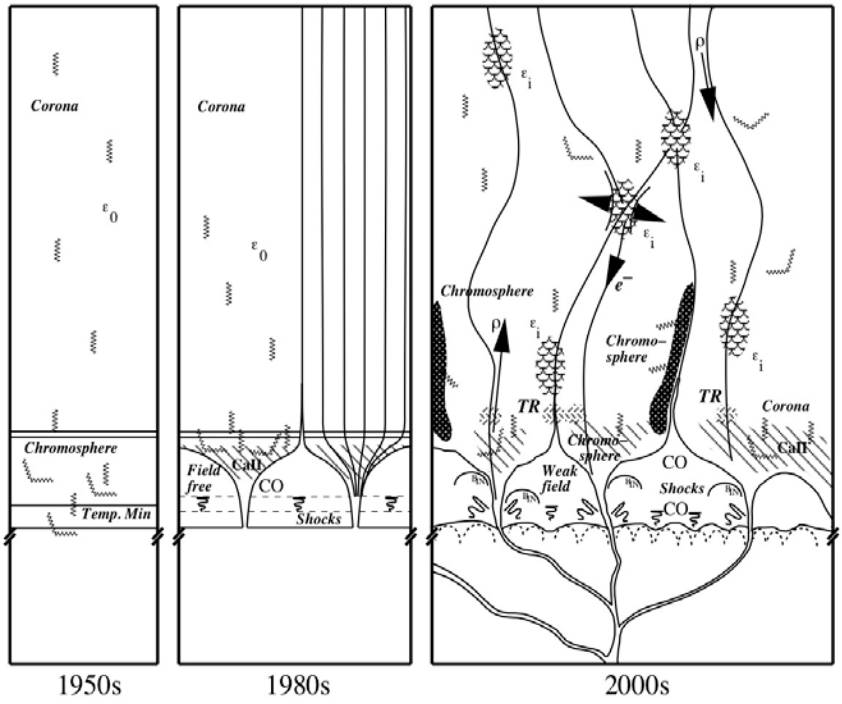


Figure 1: *Cartoon evolution of the chromosphere (Schrijver 2001). It has grown from a simple stationary "layer" of the 1950's; to topologically more diverse view of the 1980's, mixing cool ("CO") and hot gas at similar altitudes, pervaded by magnetic "flux tubes", albeit still largely static; to contemporary highly dynamic and spatially chaotic, "magnetic complexity zone." Recent observational evidence now points to the major influence of this highly complex region as the lower boundary of the heliosphere.*

Introduction: Observational Context

From ancient times, the Sun's chromosphere has been a subject of fascination. With the naked eye, it is seen only at eclipse, as a string of brilliant rosy beads above the solar limb, a bright and colorful contrast to the rather pale visage of the white light corona. In the modern era, the chromosphere retains its fascination, but now as a key interface region of the solar outer atmosphere, the de facto inner boundary of the heliosphere. While the photosphere above the convective overshoot layer is relatively quiescent (with the notable exception of the abundant acoustic and internal gravity wave flux) in the chromosphere this placidity gives way to the highly non-equilibrium, dynamic, mechanically heated, forced temperature inversion of the higher altitudes; a segue into the equally enigmatic super-hot (few MK) corona.

Like many boundaries, the transition from apparent order to disorder is intriguing and is determined by magnetism that permeates the outer solar atmosphere. The chromosphere at once hosts some of the coldest material in the whole solar atmosphere (T~3500K) as seen in off-limb emissions of 4660nm CO lines (Solanki et al. 1994) and very warm gas ($T>10^4$K) associated with ionizing shocks. The magnetic fields that pervade this region are highly twisted and tangled because of the shift from gas dominance in the deep photosphere (field lines frozen to the plasma, dragged about by surface flows) to field dominance in the corona (ionized plasma forced to flow along the field lines, bottled up in coherent large scale structures). At either extreme, the field takes on relatively simple forms: small scale intense flux concentrations in the high-β photosphere[1] and delicate near-potential loop-like structures in the low-β corona. Paradoxically, the field in the intervening β≈1 zone takes on a chaotic topology, reconnecting willy-nilly, and blossoming out into an overarching "canopy," the conspicuous source of the meandering mottling in Hα filtergrams (see cover page). In fact, the view presented in the rightmost panel of Fig. 1 would support a renaming of the chromosphere as the "magnetic complexity zone."

The signature property of the chromosphere is its enormous variability in vertical extent (Mms "thick" in internetwork regions to only a few 100km in hot, plage-rooted "mossy regions"). The prime suspect

1 The plasma β is the ratio of gas pressure to magnetic pressure.

driving chromospheric weight fluctuation is the local and temporal variation in the ionization of its dominant constituent, hydrogen. Given its large abundance and First Ionization Potential (FIP), hydrogen ionization acts as a sponge to soak up energy, buffering the gas somewhat from local heating spikes (like acoustic shocks), moderating the temperatures. Further, ionization frees electrons to fuel steady radiative cooling by collisional excitation of abundant species such as Fe, Mg, and Ca. This “ionization valve” works effectively because of the large dynamic range of the electron fraction, n_e/n_H - only 10^{-4} at the base of the “classical” inter-network chromosphere (where the electrons come from singly ionized metals); which increases dynamically (sometimes by a factor of 100; Leenaarts et al. 2007) as convection-triggered shocks (Rouppe van der Voort et al. 2007) propagate through the system. It approaches unity at the top where hydrogen is mostly ionized. These ionization effects permit the gas to balance heating even while the overall density falls outward (Ayres 1979). In the “non-classical” magnetic network and plage regions of the chromosphere a host of other phenomena are excited, as we will see below.

So, in an overall sense, the chromosphere is a creature of hydrogen ionization. Although this simple picture is appealing, it obscures the deep complexity of the region. In active regions the sunspot umbrae, penumbrae, and moats show a thick fibril web outlining the field topology as well as a host of dynamic phenomena such as umbral flashes, penumbral waves, penumbral microjets, Ellerman bombs and more. Filaments and prominences (on- and off-disk, respectively) represent enormous clouds of chromospheric gas embedded in the corona - that have the propensity to produce CMEs when erupting. A goal of contemporary solar physics, then, is to dig deeper into this intricate interface layer. Fundamental unresolved issues linking the chromosphere to the “Big Picture” are:

- ***How is the atmosphere heated and the solar wind accelerated?***
- ***How is energy transported into, and across, the chromospheric interface?***
- ***What is the source of the conspicuous chromospheric structure, both long-lived and transient?***
- ***What is the role of the chromosphere as a reservoir of mass and energy to feed the overlying corona (whose heating and mass loading requirements are but a small fraction of those of the chromosphere itself)?***
- ***What role does the ubiquitous magnetic field play in all of the above?***

The chromosphere was fertile ground in the 1950's and 1960's for the early development of specialized treatments of radiation transport and spectral line formation to confront the extreme departures from classical Local Thermodynamic Equilibrium (LTE) at the very low densities and high temperatures of the region. “Scattering” in strong resonance transitions on the one hand becomes a dominant radiation transport mechanism, but on the other hand, the induced non-locality of the radiation fields creates serious challenges for remote sensing and numerical simulation.

The Solar Optical Telescope of the *Hinode* satellite has motivated something of a renaissance in chromospheric physics and relevance. Underscoring the complexity of the region, Ca II and Hα imaging sequences at the solar limb recorded by the *Hinode* satellite reveal an incredibly rich, relentlessly dynamic, and highly structured environment. Indeed, the Hinode observations were of such quality that two distinct populations of “spicule” (Roberts 1945) could be discerned (De Pontieu et al. 2007). The first, evolving on timescales of a few minutes, typical of the underlying photospheric convection and internal p-mode envelope, showed material rising and falling along the same magnetically determined path at the speeds between 20 and 40km/s in a process described by Hansteen et al. (2006). The second, previously unresolved, class is taller (>1Mm taller), faster (50-150km/s), and evolves faster (10-100s) before quickly fading, a hint of heating to higher temperatures. *Hinode* also revealed that both of these spicule flavors undergo significant transverse (Alfvénic) motions. In coronal holes the estimated wave flux is enough to accelerate the fast solar wind (De Pontieu et al. 2007).

More recently, an observational picture of the chromospheric-coronal mass and energy transport has developed. De Pontieu et al. 2009, McIntosh & De Pontieu (2009a), and McIntosh & De Pontieu (2009b) have investigated the connection of these “type-II” spicules and the ubiquitous insertion of plasma, preheated to typical corona temperatures, in quiet, coronal hole, and active regions using a combination of spectral and imaging observations. These observations in some measure dispel the common belief that the coronal “tail” wags the chromospheric “dog” and ties directly into all of the issues raised above.

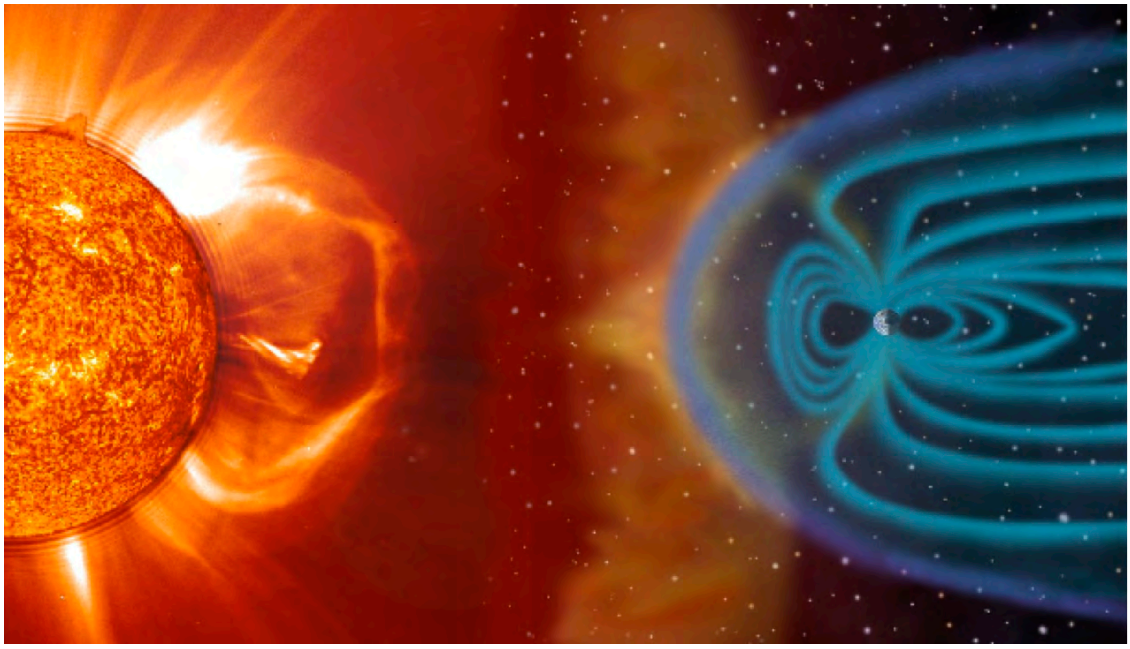


***Figure 2**: Popularized depiction of space weather; grossly out of scale, but nicely illustrating the impact of wind, UV radiation, and mass ejections on Earth.*

The Modern Relevance of the Chromosphere

The relentless emergence and convective forcing of magnetic flux through the solar surface determines the cycle of mass and energy inside the chromosphere and into the corona, producing the bulk of the Sun's UV and EUV emission. Understanding the mass-energy cycle and the formation of this emission provides essential physical insight into the radiative forcing of the terrestrial atmosphere, as embodied in state-of-the-art climate models (e.g., Marsh et al. 2007). Recent observational results discussed above have indicated that the release of mass and non-thermal energy at or near the base of the chromosphere is the seed of the solar wind, another key process whose physical origin remains elusive. Understanding the solar wind is relevant for society not only through direct space weather effects, but also because of more subtle indirect forcing of the terrestrial atmosphere, for example by modulation of Earth-impacting energetic particles.

Our models of coronal magnetism and eruptive phenomena rely almost exclusively on photospheric vector field measurements to extrapolate the magnetic force lines into the corona. These efforts typically fail to meet desired metrics on the model field (DeRosa et al. 2009), with the Lorentz forces acting on the field in the upper photosphere being the likely culprit. These forces can be so large that the inferred electrical currents are not field-aligned, and the magnetic force lines either never make it into the corona, or are deflected to very different positions before they enter the corona. The obvious remedy to this issue is to measure the vector field directly in the low-β regime of the upper chromosphere, and extrapolate from there. In principle, measurements of the three-dimensional, time varying, magnetic field at the heliospheric base can be made using imaging spectro-polarimetry or radio measurements. At the very least the line-of-sight magnetic fields can be accurately determined, and, under favorable conditions, so can the transverse components. Very recent observations also have demonstrated the presence of ubiquitous high speed (magnetized) flows that are closely aligned with the background magnetic field (McIntosh & De Pontieu 2009a). Thus, the combination of kinematic velocities and vector field polarimetry offers the highest potential of success for making routine measurements of the vector magnetic field at the heliospheric base (see also the white paper by Philip Judge).

Obstacles to Progress in Traversing the Heliospheric Frontier

Four effects mainly set the chromosphere apart from the much denser, near equilibrium photosphere. Each of these presents barriers to numerical simulations of chromospheric structure and physical processes; these also are potential obstacles to remote sensing.

1. The chromosphere encompasses the domain in which magnetic fields begin to gain dominance over purely hydrodynamic forces. It contains intermixed regions where the plasma β is greater than and less than unity, with substantial spatial and temporal fluctuations. Strong gradients can develop in the chromospheric field owing to the small magnetic diffusivity.
2. The chromosphere is optically thin at most visible wavelengths, apart from a few strong spectral lines. Because these lines also dominate the transport of radiative energy, they must be handled with so-called non-LTE formation: challenging to formulate, dependent on often poorly known atomic physics, and demanding to compute.
3. Recombination time scales of hydrogen - the heart of the “ionization valve” discussed earlier - can be much longer than dynamical time scales. Consequently, the ionization state at any instant depends on the previous history of the plasma. Chemical equilibrium time scales also are long compared with the dynamics. This can affect key species like CO, instigator of “molecular cooling catastrophes” in places where mechanical heating has temporarily abated (Ayres 1981). As with the magnetic field, strong gradients in ionization and chemistry can develop, with consequences for the plasma energy balance (e.g., Leenaarts et al. 2007).
4. There also is mounting evidence that the partial ionization leads to multi-fluid effects, very different from single-fluid MHD (e.g., Arber et al. 2007). This has clear consequences for wave propagation and mode couplings, but also might affect the outcomes when Lorentz forces act during, say, flux emergence, fibril dynamics, type-II spicule excitation, or filament disruption.

The severe departures from the equilibrium state make it much harder to model the chromosphere at a level of realism approaching what has been achieved for the photosphere.

The State of Modeling Efforts

In high-resolution modeling efforts, the near-equilibrium photosphere has served as an important laboratory to vet our understanding of kinetic transport, culminating in the sterling success of modern 3D convection models (e.g., Stein & Nordlund 2000; Vögler et al. 2005). The important lesson is the pivotal role played by small scale phenomena in controlling transport processes. Solar granular convection is a highly non-local phenomenon, driven from a thin surface thermal boundary layer. Here, radiative cooling produces low entropy gas that, in turn, forms strong downdrafts concentrated in the narrow intergranular lanes, smaller than 100km in size (less than a tenth the diameter of the parent granule). Most of the buoyancy work occurs in such tiny structures. Similar comparisons currently underway for the more complex physical situation of magneto-convection show that dynamical control by small scale elements holds generally: in magnetic fibrils of the quiet Sun (e.g., Stein & Nordlund 2006), sunspot umbrae (Schüssler & Vögler 2006; Rimmele 2008), and the surrounding penumbae (Scharmer et al. 2008; Rempel et al. 2009).

Now, it is the chromosphere that stands as the frontier for the palpably more challenging (Non-LTE) radiation-magnetohydrodynamic (R-MHD) simulations required when the classical transport processes begin to break down, and the magnetic field and non-linear acoustic disturbances take over (e.g., Wedemeyer-Böhm et al. 2004; Abbett 2007; Martinez-Sykora et al. 2008). Most of the mass and non-radiative energy that sustain the outer atmosphere resides in the chromosphere, and from here are channeled upwards, thanks mainly to the magnetic field (e.g., Gudiksen & Nordlund 2005). However, it still is far from clear exactly how the transport and heating are accomplished, and in particular the roles of the copious small and rapid phenomena, routinely observed in different guises as spicules, mottles, fibrils, jets, Ellerman bombs, and other events with similarly evocative names.

Some of these phenomena have physically consistent explanations, for example the creation of dynamic fibrils and mottles by means of magneto-acoustic shocks - driven themselves by wave leakage from the solar interior (De Pontieu et al. 2007; Rouppe van der Voort et al. 2007). However, for the most part, a complete description of the link between the photospheric drivers and the resulting hot outer atmosphere remains elusive. Fully 3D, R-MHD numerical simulations spanning from the upper convection zone through the corona - treated as a coherent system - are starting to appear, but cannot yet address (all) the salient physical ingredients on scales larger than a few granules or one super-granule. Thus, a close collaboration between observers and theorists, similar to the successful model for photospheric processes, is essential to make progress in the chromospheric arena.

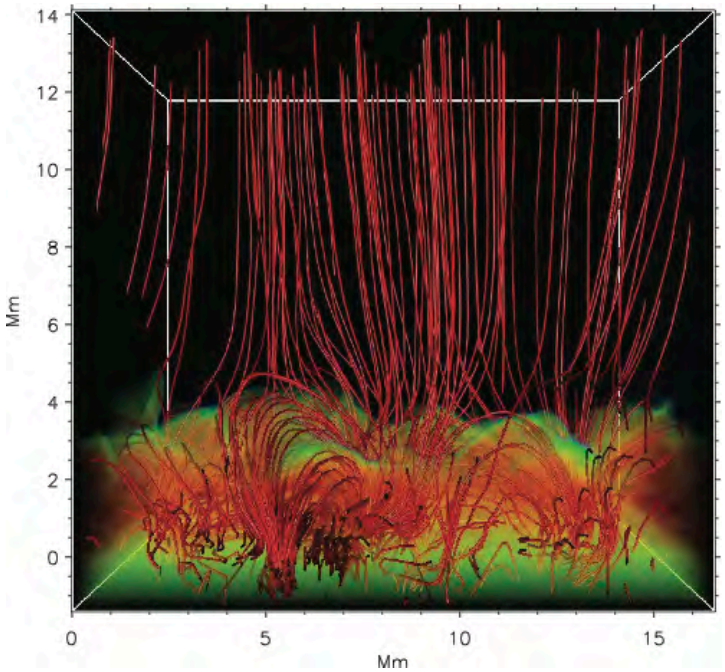


***Figure 3**: Snapshot of a 3-D R-MHD quiet Sun simulation from De Pontieu et al. (2007). The coloring illustrates the plasma temperature from the low chromosphere (red) to higher temperatures near the transition region (green).*

New Opportunities

To achieve a successful close comparison between chromospheric observations and theory, much work is needed on both sides. For example, from the observational perspective, the necessary polarimetric accuracy required to capture chromospheric vector fields is close to achievable at the desired spatial and temporal scales. With the present instrumentation and telescopes (dividing up the photons into space, time, frequency, and polarization states quickly exhausts the supply) although progress is very rapidly evolving in this area.

Innovations in both areas are actively being pursued, especially telescope aperture. Notably, the 4m Advanced Technology Solar Telescope (ATST) will be the flagship international facility in the coming decade to carry out pioneering high (spatial/temporal/spectral) resolution studies of the solar atmosphere. On the instrumental front, designers are attempting to address several fundamental, sometimes conflicting, requirements. First, recording multiple spectral lines is necessary to probe the many scale heights spanned by the chromosphere. Second, to investigate the discrete components of the system and adequately trace the 3D interconnectivity forced by the magnetic field, these detailed measurements must be obtained over a field of view capturing several cells of the supergranulation "network", a substantial portion of a prominence or filament. Third, it is essential to collect the multi-line diagnostics *simultaneously*, to avoid ambiguities introduced by fast structure evolution and solar surface kinematics. Prototype instruments like IBIS at NSO's Dunn Solar Telescope and CRISP at the Swedish Solar Telescope on La Palma already are driving significant breakthroughs in understanding (e.g., de la Cruz Rodriguez et al. 2010). We should stress, however, that the detailed ATST/CRISP/IBIS observations cannot exist in a vacuum, if we are to fully understand the transfer of mass and energy across the disparate scales of the chromosphere, assess its impact on the heliosphere, and support missions like SDO and IRIS that are blind to chromospheric magnetism. Here is where innovative synoptic instrumentation will play a key role, exploiting the same observational diagnostics, while pushing the limits of filter, detector, and image post-processing technology.

Deployment of high performance instruments with advanced capabilities, and operating at solar telescopes of large aperture, supported by next-generation "synoptic", or patrol, instruments is a clear priority for the coming decade.

A related issue is the need to achieve high spatial resolution, of order 0.1" (~70km) or better, over the extended fields of view, to resolve the fine scale structures in which much of the magnetic "action" likely takes place. Deployment of a large aperture solar telescope in space (e.g., Solar-C "Plan B") employing advanced filter imaging technology will ensure a seeing-free, consistently high signal-to-noise observations. A complementary option for ground-based observation are multi-conjugate and ground layer adaptive optics systems.

Development and deployment of wide-field Adaptive Optics systems should be strongly supported.

In terms of theory, efforts to simulate the structure and dynamics of the photosphere, chromosphere, and corona as one coherent system are underway. Prototype codes, such as Oslo's BIFROST, approach realism by incorporating time-dependent hydrogen ionization, non-equilibrium radiation transport, multi-fluid behavior, and diverse magnetic field configurations, to mention a few key components; all challenging to implement and very demanding to compute (e.g., Fig. 3 and Martinez-Sykora et al. 2009). Forward modeling not only is a key to sort out the most relevant physical effects, "joining the observational dots," but also a tool for vetting measurement techniques: model intensities can be subjected to a simulated observation to judge how much of the original physical information is recovered by that procedure.

State-of-the-art forward modeling should be strongly supported, not only to achieve physical insight, but also to validate and enhance observational strategies. This will require highly parallel high-end computing services with tens of thousands of cores and tens of millions of CPU hours. A necessary adjunct is access to diverse and accurate atomic physics data.

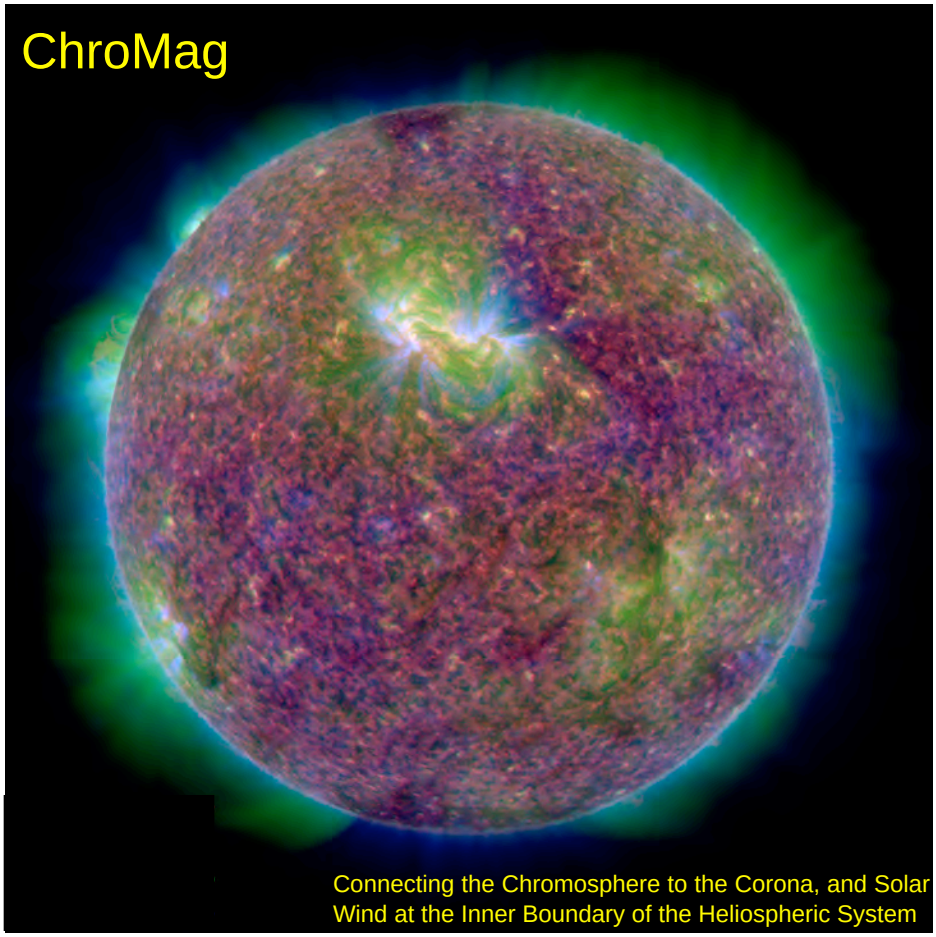
As a final remark, coordinating space-based UV, visible, and X-ray imaging (say, from Hinode) with ground-based visible and infrared spectroscopy continues to be a mainstay of contemporary solar astronomy. The chromosphere, however, presents a difficult remote-sensing target, because it is mostly transparent in the visible, aside from the cores of a few strong resonance lines. At the same time, the chromosphere is very opaque in the Lyman continuum below 91nm. This is a pity because many solar space imagers and spectrometers operate in the EUV around 20-30nm, and consequently cannot directly probe this region. Finding new windows into the state of the chromospheric gas thus is both a major challenge and an opportunity. For example, UV imaging spectroscopy longward of the LyC edge, as realized in the IRIS instrument currently under development, will allow us to "see" deep into the chromosphere, investigating the crucial temporal evolution of hydrogen ionization in the "clapotisphere" (Rutten 1995), and the upward extension of the explosive small-scale energy releases sprinkled across the magnetic network. Important additional views to the chromosphere lie in other regions of the electromagnetic spectrum. At millimeter wavelengths, for example, the Atacama Large Millimeter Array (ALMA; under construction in northern Chile) will probe the low-to-mid chromosphere with directly measured temperature contrast and high polarimetric accuracy at resolutions of 0.1" (or better). In addition, the Expanded Very Large Array (EVLA) will investigate explosive activity and particle acceleration at the top of the chromosphere in centimetric light, at the crucial level where magnetic complexity evolves into coronal heating. Unfortunately, these facilities mainly are dedicated to nighttime astronomy, and have relatively low solar duty cycles. In contrast, the Frequency Agile Solar Radiotelescope (FASR) will be a dedicated wide-band (1cm-6m) array that can probe the full-sun from the mid-chromosphere all the way to the corona, monitoring the thermal and magnetic conditions of the outer atmosphere in a "systems" context, providing complimentarity to missions like SDO.

Chromospheric physics is at the cutting edge of our understanding of the solar atmosphere. The chromosphere is a physical domain where current observations point to a convergence of complex effects - magnetic, kinematic, and radiative - whose impact on flows of mass and energy into the heliosphere is high and at the same time our potential to achieve physical insight is very promising. Consequently, there is clear need and drive to develop the observational diagnostics and interpretative tools vital to our understanding of this, the heliospheric system's crucial innermost boundary.

ChroMag: The Community Synoptic Chromospheric Magnetograph

A White Paper Submitted to the Heliophysics 2012 NRC Decadal Survey

S. W. McIntosh (High Altitude Observatory; corresponding author)
R. Casini, B. Lites, S. Tomczyk (HAO)
T. Berger (Lockheed Martin Solar and Astrophysics Laboratory)
G. Cauzzi, K. Reardon (INAF - Arcetri Astrophysical Observatory)
A. Oritz, L. Rouppe van der Voort (University of Oslo)



The Community Synoptic Chromospheric Magnetograph (ChroMag) is designed with the goal of capturing the relentlessly dynamic chromospheric environment through cutting-edge multi-wavelength imaging spectro-polarimetry of the full solar disk at a resolution approaching one arcsecond, and in a few seconds. The rapidly developing picture of chromospheric magnetism and dynamics is one that is growing in immediacy as breakthrough observations continue to identify the features that drive both the energetics and allow access to vector magnetic field measurements at the true lower boundary of the heliospheric system. Naturally, an essential part of the ChroMag program is a commitment to develop and provide community access to the "inversion" tools necessary to interpret the complex multi-component plasma environment. We contend that measurements of an instrument like ChroMag provide critical physical context for the Solar Dynamics Observatory (SDO) and Interface Region Imaging Spectrograph (IRIS) as well as the very high resolution observations of the Swedish Solar Telescope and soon-to-be-reality Advanced Technology Solar Telescope.

Introduction

The solar chromosphere is a creature of hydrogen ionization. Although this simple picture is appealing, it obscures the deep complexity of the region. In active regions the sunspot umbrae, penumbrae, and moats show a thick fibril web outlining the field topology as well as a host of dynamic phenomena such as umbral flashes, penumbral waves, penumbral microjets, Ellerman bombs and more (Fig 1). Filaments and prominences (on- and off-disk, respectively) represent enormous clouds of chromospheric gas embedded in the corona - that have the propensity to erupt and produce Coronal Mass Ejections (CMEs). A goal of contemporary solar physics, then, is to dig deeper into this intricate interface layer. Some of the fundamental unresolved issues in solar physics have their "roots" firmly embedded in the chromosphere:

- ***How is the atmosphere heated and the solar wind accelerated?***
- ***How is energy transported into and across the boundary?***
- ***What is the source of the conspicuous structure, both long-lived and transient?***
- ***What is the role of the chromosphere as a reservoir of mass and energy to feed the overlying corona?***
- ***What role does the ubiquitous magnetic field play in all of the above?***

Motivated in large part by the legacy of the *SOHO* mission and recent developments with the *Hinode* satellite we are investigating the impact of the relentless emergence and convective forcing of magnetic flux through the solar surface in determining the cycle of mass and energy through the chromosphere and into the corona (De Pontieu et al. 2009; McIntosh & De Pontieu 2009a; McIntosh & De Pontieu 2009b). It is becoming clear that *this mass cycle rooted in chromospheric activity is responsible for producing the bulk of the Sun's UV and EUV emission*. Understanding the cycle of mass and the formation of this emission provides an essential physical insight into the radiative forcing of the latest classes of terrestrial models (e.g., Marsh et al. 2007). Further, De Pontieu et al. (2007) have indicated that the *release of mass and non-thermal energy at or near the base of the chromosphere is the seed of the solar wind*, another key process whose physical origins remain elusive. Understanding the solar wind is not only relevant for society through the effects of space weather, but also because we need a clear basic understanding of the indirect forcing of the terrestrial atmosphere and modulation of Earth-impacting energetic particles.

Our models of coronal magnetism and eruptive phenomena rely almost exclusively on photospheric vector field measurements to extrapolate the magnetic lines of force into the corona. These efforts typically fail to meet reasonable requirements (DeRosa et al. 2009), with the Lorentz forces acting on the field in the upper photosphere the likely culprit. These forces can be so large that the inferred electrical currents are not field-aligned, and either never make it into the corona, or are deflected to very different positions before they enter the corona. The *immediate remedy to this issue is to measure the vector field directly in the low-β regime of the upper chromosphere, and extrapolate from there*. Fortunately, measurements of the three-dimensional, time varying, magnetic field at the heliospheric base can be made using imaging spectropolarimetry (e.g., de la Cruz Rodriguez et al. 2010). At the very least the line-of-sight magnetic fields can be accurately measured, and, under favorable conditions, so can the transverse components. The very recent observations have also demonstrated the presence of ubiquitous high speed (magnetized) flows that are highly aligned with the background magnetic field (McIntosh & De Pontieu 2009a). *Thus, the combination of kinematic velocities and vector field polarimetry give the highest possibility of success in making routine measurements of the vector magnetic field at the heliospheric base.*

Recording the spectro-polarimetric information from multiple spectral lines is necessary to probe the many scale heights spanned by the chromosphere. To investigate the discrete components of the system, adequately trace the 3D interconnectivity forced by the magnetic field, and *investigate the response of the atmosphere at the onset of Space Weather events, these measurements must be obtained over the full solar disk*. Finally, is essential to capture the multi-line diagnostics *as close as possible to simultaneously*, to avoid ambiguities introduced by fast structure evolution and solar surface kinematics. Prototype instruments like IBIS at NSO Sacramento Peak and CRISP at the Swedish Solar Telescope on La Palma and are already driving significant breakthroughs in understanding of the complex environment beyond the capabilities of slit spectroscopy (Judge et al. 2010). However, we should stress that these beautifully detailed CRISP/IBIS observations cannot are not sufficient if we seek to fully understand the transfer of mass and energy across the disparate scales of the chromosphere, assess, and monitor its impact on the heliosphere, and to complement in a unique and necessary way to missions like SDO and IRIS that are “blind” to chromospheric magnetism. They *must* be supported by innovative synoptic instrumentation using the same observational diagnostics as the high-resolution telescopes (e.g., SST/CRISP, DST/IBIS), pushing the limits of filter, detector, and image post-processing technology, as we will do with ChroMag.

A instrument like ChroMag can attack the five fundamental questions by providing the community with a comprehensive synoptic spectro-polarimetric observational dataset from spectral lines formed at multiple “heights” of the chromosphere and very base of the corona.

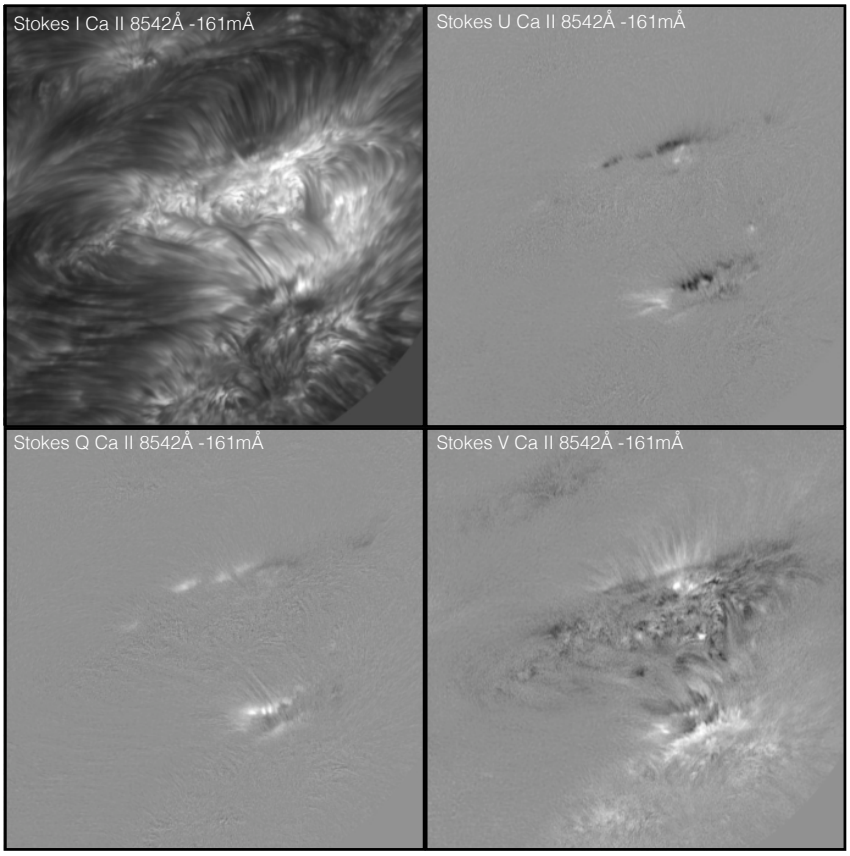


***Figure 1**: SST/CRISP observations near the limb in the blue wing of Ca II 8542 taken on June 12 2008. The columns show the progression of the Stokes vector (I, Q, U, and V), top to bottom, left to right. In these images the fine scale of the middle chromosphere is highly visible not only in I, but in the linear and circular polarization states (e.g., de la Cruz Rodriguez et al. 2010).*

Recent Motivational Science

The Solar Optical Telescope of the *Hinode* satellite has motivated something of a renaissance in chromospheric physics and its relevance to the fundamental problems of solar physics. Underscoring the complexity of the chromosphere, Ca II and Hα imaging sequences at the solar limb recorded by the *Hinode* satellite reveal an incredibly rich, relentlessly dynamic, and highly structured environment. Indeed, the Hinode observations were of such fidelity that two distinct populations of "spicule" (Roberts 1945) could be discerned (De Pontieu et al. 2007). The first, evolving on timescales of a few minutes, typical of the underlying photospheric convection and internal p-mode envelope, showed material rising and falling along the same magnetically determined path at the speeds between 20 and 40km/s in a process described by Hansteen et al. (2006). The second class, previously unresolved, is taller (>1Mm taller), faster (50-150km/s), and evolves much faster (10-100s) before rapidly fading, a hint of heating to higher temperatures. *Hinode* also revealed that both of these spicule flavors undergo significant transverse (Alfvénic) motions. In coronal holes the estimated wave flux is enough to accelerate the fast solar wind (De Pontieu et al. 2007).

More recently, an observational picture of the chromo-coronal mass transport has developed. De Pontieu et al. 2009, McIntosh & De Pontieu (2009a), and McIntosh & De Pontieu (2009b) have investigated the connection of these "type-II" spicules and the insertion preheated to typical corona temperatures in the quiet, coronal hole, and active regions identified in spectral and imaging observations. The recent launch and deployment of SDO has highlighted these transport processes. The significantly improved signal-to-noise, reduced scattered light, high cadence, multi-wavelength observations of AIA are revealing the critical connectivity of chromosphere and corona as never before observed (De Pontieu et al. 2010 - submitted Science; McIntosh et al. 2010 - submitted ApJL; Fig 2). *Unfortunately, SDO and the forthcoming IRIS mission do not have the on-board capability to perform chromospheric imaging spectroscopy, much less imaging spectro-polarimetry and so there will be a considerable and immediate need in the heliophysics community to provide these global observations of the chromosphere in a synoptic fashion.*

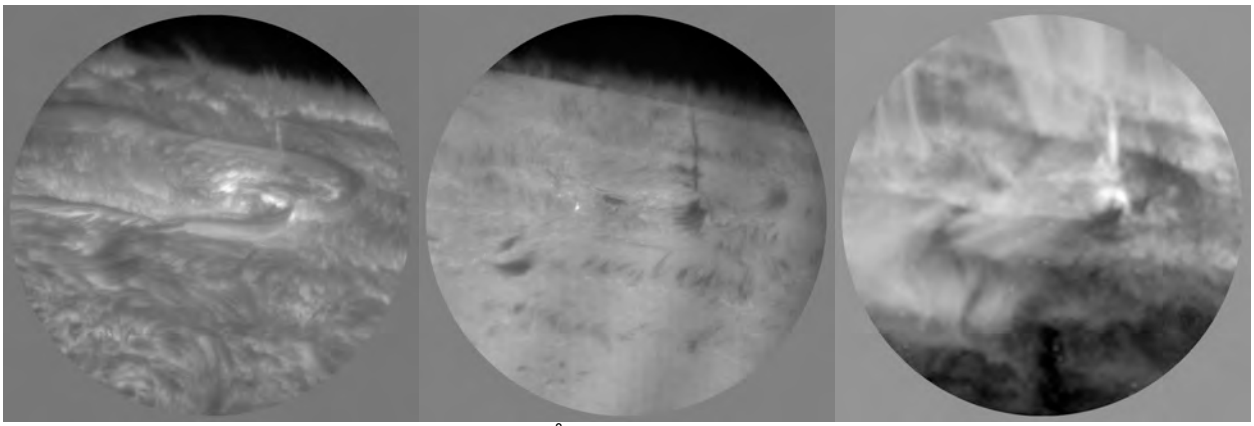


***Figure 2**: NSO/DST IBIS images of Ca II 8542Å line core (left) and 40-60km/s in the blue wing (center) in comparison with Fe IX 171Å coronal emission observed by SDO/AIA (right). Note the very strong visual correspondence of the chromospheric filter images and the coronal emission. The dark absorbing features in the center panel are "Rapid Blueshift Events" (RBEs) and are directly associated with type-II spicules (Rouppe van der Voort et al. 2009).*

Space Weather Drivers of ChroMag's Science Requirements

Besides the five fundamental unresolved issues mentioned above detailed scientific questions to drive the ChroMag concept were determined by the heliospheric community through the COSMO Science Advisory Panel[1]. They are:

1. *What is the nature of the changes in coronal magnetic structure that accompany the 11-year solar cycle?*
2. *What is the physics behind the storage and explosive release of magnetic energy in the solar atmosphere that leads to CMEs?*
3. *How are prominences formed, how do they evolve, and how are they related to CMEs?*
4. *What is the role of magnetic reconnection and flares in CME formation?*

In order to capture the dynamic and magnetic properties of these space weather scale events, as well as those of the small-scale mass and energy loading into the corona and solar wind, places tight scientific requirements on ChroMag. We require a 2.5 solar radii full field-of-view (FOV) to monitor the eruption of prominences/filaments and the initial phase of CMEs. The requirement to observe the small-scale energy release will require a spatial resolution of no more than two arcseconds[2]. Furthermore, the rapid temporal evolution of these phenomena over 10-100s require that a set of spectral samples need to be acquired on commensurate timescales. Finally, simulations of the inversion of magnetic field data in prominences and filaments indicate that a tunable filter instrument can provide data of sufficient quality to achieve 5G sensitivity as long as noise can be kept at a level lower than 10^{-3} (Casini 2007) and may only be available in Stokes V measurements (*de la Cruz Rodriguez et al. 2010*) over the whole disk. Measuring prominence magnetic fields embedded within the corona provides a more complete picture of field configurations than could be provided from a coronal or chromospheric magnetometer alone. As a result of this sensitivity requirement the tunable filter must have a

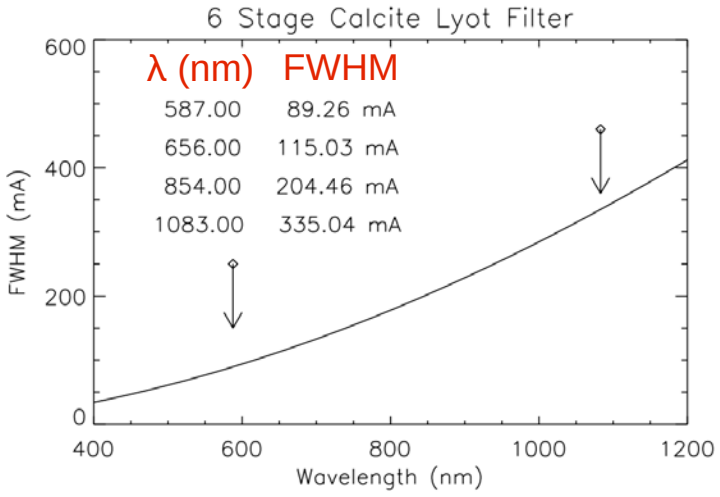


***Figure 2**: Mapping the Full-Width-Half-Max (FWHM) of the ChroMag six-stage Lyot filter. The diamonds and arrows indicate the maximum "tolerance" in spectral resolution resulting from the simulations of Casini (2007). The six-stage ChroMag filter design considerably exceeds this tolerance.*

[1] COSMO, or the COronal Solar Magnetism Observatory, is a conceptual community facility to be operated by HAO as a replacement for the aging Mauna Loa Solar Observatory (MLSO). The components of COSMO are discussed in the white paper submitted by Steve Tomczyk.

[2] With a 2" resolution the EUVI telescope on the STEREO spacecraft can capture these events propagating into the corona (McIntosh et al. 2010). It is highly likely that much of the discrete energy release happens below this scale, but that it happens with enough spatial coherence that the ChroMag resolution will capture the ensemble effect.

FWHM of 240mÅ in the visible region and 460mÅ in the IR and these requirements are met by the preliminary ChroMag design (Elmore 2007; Fig. 2).

The Preliminary Design of ChroMag

ChroMag is a polarimeter designed to measure (on-disk) chromospheric and (off-disk) prominence magnetic fields using the chromospheric spectral lines of He I (587.6 and 1083nm), Hα (656.3nm) and Ca II (854.2nm) with a full FOV of 2.5 solar radii and a spatial resolution of 1.5" (photospheric magnetograms will also be possible using the Si I 1082.7nm line). ChroMag is a 20-cm Lyot coronagraph operating with an occulting disc for prominence measurements and un-occulted for on-disk measurements. The design maintains an instrumental scattered light level small compared to prominence brightness. Even with this constraint, it is possible to use an air-spaced doublet objective lens allowing for observations over the required large spectral range.

In order to make polarimetric measurements spanning the entire solar disc at a high (<<1 minute) cadence, ChroMag will use a pair of tunable Lyot filters, one for each of the dual beams. The Lyot filters have six stages, are wide-fielded, and are tuned using Ferro-electric liquid crystals (FeLCs). If we use the lifetime of type-II spicules (10-100s) as a benchmark for which we should be able to scan a single spectral line. Sampling that line at a dozen wavelengths means that each scan will be collected in around 15 seconds, an easily attainable target (Judge et al. 2010).

The lens for ChroMag was chosen so that the optical path can have low scattered light and be axially symmetric to minimize telescope polarization. Signal to noise is limited by the total number of photons collected and by atmospheric seeing. High frame rate 2096x2096 (or 4096x4096) pixel detectors are used in order to perform the polarization measurement rapidly, compared to atmospheric seeing. In addition ChroMag will employ a simple tip-tilt mirror system in the secondary mirror to remove the remainder of the aberration introduced by the atmospheric motion. A dual beam polarimeter is used to minimize seeing-induced cross-talk. A set of pre-filters are used to isolate spectral regions around the chromospheric lines of interest.

Extracting Physical Units: Inverting the SpectroPolarimetric Data

Our goal with ChroMag is to investigate the four dimensional (space+time) variations of the magnetic and thermodynamic environment of the chromosphere over the whole disk and how the changes both form and drive the base of the heliospheric system. This goal mandates that we measure the full Stokes vector (IQUV) using an instrument as described above and in the associated technical notes. While we are making great strides in the analysis and interpretation of the (often multi-component) intensity ("I") measurements (Rouppe van der Voort et al. 2009; De Pontieu et al. 2010 - "How Is The Solar Corona Filled With Hot Plasma?", Submitted to Science) the magnetic components require considerable theoretical and modeling efforts to "invert".

Progress in inverting Stokes Q, U, and V is much more arduous. Recalling that the local transverse fields are identified through their linearly polarized components ("Q" and "U") in spectral lines, and the line of sight fields are reflected in the circularly polarized ("V") components of the Stokes vector. Casini (2007) uses a Principal Component Scheme to synthesize a multi-dimensional grid of sample line profiles, where each dimension is a physical variable in the model atmosphere generating the synthetic line profile. Those profiles are then matched with the observations using the standard χ^2 metric, reproducing the measured Stokes vector reliably with relatively coarse variable grids (Judge et al. 2010). Such an approach will provide ChroMag reliable "quick-look" magnetograms. Alternative approaches to inverting the Stokes vector can follow the SDO/HMI "VFISV" (Borerro et al. 2010), or "NICOLE" method of

Socas-Navarro et al. (2000). There are alternatives too in the HeLIx, and HAZEL inversion formalisms.

Another important portion of the ChroMag effort will involve the detailed modeling of multi-fluid behavior, complex multi-dimensional radiative transfer, and partial ionization of the chromospheric plasma. As discussed in this white paper on the state of chromospheric physics these advances are coming, but require considerable theoretical effort and the multi-dimensional radiation MHD numerical techniques to reach a level of maturity/scalability that make the full rMHD "forward-modeling" approach to inverting the data practical.

ChroMag Heritage & Readiness

The ChroMag concept is feasible, having been extensively studied by the HAO Instrumentation Group (e.g., Casini 2007; Elmore 2007) and vetted by the COSMO Science Advisor Panel. ChroMag is largely motivated by the design, and success, of the Lyot-based Coronal Multi-channel Polarimeter (CoMP; Tomczyk et al. 2008). CoMP unveiled the ubiquitous Alfvénic motions of the corona (Tomczyk et al. 2007) and was recently deployed to the Mauna Loa Solar Observatory to provide the community with daily synoptic observations of coronal magnetism and dynamics. The scientific requirements of the instrument have been determined, as have the properties of the Lyot filters and optical system. The cost of ChroMag will be in the $4-5M range. If the initial ChroMag investigation is deemed successful we will investigate further community partnerships to construct a network of similar telescopes and increase the duty cycle of these important synoptic observations.

We should stress that ChroMag is, in no way, a replacement for the very highly detailed chromospheric imaging spectropolarimetry of SST/CRISP, DST/IBIS, or the upcoming Visible Tunable Filter (VTF) for the ATST. Those instruments are the vanguard "microscopes" into the complex chromospheric domain and are essential to understand the full complexity of the broader passband ChroMag full-disk measurements.

ChroMag & "The Final Frontier"?

NASA should consider a chromospheric magnetometry mission as a high priority for the coming decade. Solar-C "Plan-B" fits the bill, but if it does not go forward NASA must be prepared to endorse a SMEX or MIDEX mission to fill this critical science need. A narrow FOV version of ChroMag's wavelength-diverse filter-based imaging spectropolarimeter system is a very viable option at the core of such a payload. In the seeing-free environment of space, and with near 24-hour continuous coverage afforded by a satellite mission, the potential to precisely measure the vector field in the chromosphere becomes almost unlimited and the subsequent leap in our understanding of the detailed physics of the innermost boundary of the heliosphere will undoubtedly be huge.

Interstellar Probe

Voyager 1 and 2 observations and recent observations from the Interstellar Boundary Explorer (IBEX) and Cassini orbiter are providing significant new information about the interaction of the solar wind and surrounding interstellar space. Direct information requires *in situ* measurements to understand the global nature of our local galactic environment, which is significantly more complex than previously thought. Given the large distances involved and the limited remaining lifetimes of the Voyager spacecraft, such a mission is eminently timely. The central problem is that of providing a means of propulsion to accelerate a probe from the Solar System. Even with a low-mass spacecraft, achieving the high speeds needed has remained problematic. We consider the various promises and problems for launching such a mission during the coming decade. We conclude that such a mission is feasible with current or on-going technical developments. We discuss, in the order suggested[*], (1) the concept (approaches, spacecraft, payload), (2) implications for solar and space physics, (3) estimated cost, and (4) how this concept meets Heliophysics Decadal Survey criteria.

1 Summary of the Concept

The heliospheric boundary region, recently entered by the Voyagers, and its interaction with the interstellar medium is still one of the last frontiers of uncharted territory in heliophysics. Only an Interstellar Probe with modern instruments and measurement requirements better defined by these recent observations can provide the new information required.

1.1 Implementation Approaches

Voyager 1 is the fastest object escaping the solar system (~3.6 AU/yr) by virtue of its double gravity assists by Jupiter and Saturn. The synodic period of these two bodies is just under 20 years, but with an added constraint of the asymptotic trajectory being confined to a small range of ecliptic latitude and longitude, such mission opportunities become rarer[1]. It has typically been assumed in the scientific community that for such a dedicated effort to be worthwhile, an asymptotic speed at least double that of Voyager 1 – and preferably higher still – is a prerequisite for the mission. Hence, some form of "advanced" propulsion has always been viewed as enabling and essential.

The use of nuclear electric propulsion (NEP) was favored initially, but such systems tended to be large[2-4]. More recent work using the Project Prometheus architecture[5] came to the conclusion that the specific mass (kg/kWe) – as well as the gross mass – of that technical implementation was too large to offer the fast transit times required[6]. Such characteristics have been common for attempts to design fission systems for space[7].

Ballistic, powered, near-Sun, gravity assists[8] were recognized as offering a different potential solution, but with challenges in the high-thrust propulsion capability required[9]. More detailed studies have studied how best to supply such a capability. Specific impulses ~1000 sec at high thrust are required and solar-thermal propulsion offers a solution[10-15]. However, the approach required a very-low probe mass (~150 kg) along with a combined perihelion engine whose function was mission critical, could not be tested in the most stressing environment, and required long-term (3+ year) storage of liquid hydrogen (LH2).

Low-thrust approaches include solar sails[16-19] and RTG-powered electric propulsion[9], the latter now known as radioisotope electric propulsion (REP)[20,21]. Solar electric propulsion (SEP) has also been recently reexamined in conjunction with REP and gravity assists[22]. All of these approaches have limitations driven by materials and mass issues, but appear feasible.

---

[*] *Decadal Study-Request for Information (RFI) from Community* from http://sites.nationalacademies.org/SSB/CurrentProjects/SSB_056864#White_Papers_and_Community_Input

1.2 Solar Sails and Solar Electric Propulsion

Solar sails make use of the radiation pressure exerted by sunlight. At 1 AU, the solar constant momentum flux is p_{sun} = 4.5605 μNm^{-2}. The relevant physical parameter is the "lightness number" β defined as the ratio of solar radiation pressure force on the spacecraft to the gravitational force. The decreasing effectiveness of solar sails with distance, e.g. past the orbit of Jupiter, has led to the mission-design strategy of performing so-called multiple "solar photonic assist" maneuvers that turn the trajectory into a hyperbolic one[23]. The European Space Agency (ESA) has studied this approach for the Interstellar Heliopause Probe/Heliospheric Explorer (IHP/HEX). Mission analyses suggest flyout times to 200 AU in ~20 to 30 years with launch masses of a few hundred kg could be accomplished for sail areal masses of ~5 g/m^2, areas > 10^4 m^2 with approaches to ~0.2 AU of the Sun[24,25]. Such estimates are in accord with previous calculations[18,19].

A similar strategy can be followed with SEP. Common to both approaches is the exploitation of increased solar radiation flux and gravity by first going to the inner solar system and taking up momentum there. The SEP approach has been studied in conjunction with REP (discussed in the next section). In this case the SEP stage provides the energy otherwise provided by a large ballistic booster (see §1.3). Fast escapes still require gravity assists and significant electrical power (~50 kWe at 1 AU) with a higher initial spacecraft mass. However, in addition to a smaller required launch vehicle, a less close approach to the Sun of ~0.7 AU is required compared to ~0.2 AU in the solar sail case[22].

1.3 Radioisotope Electric Propulsion (REP)

A "small" interstellar probe requires power for the onboard electronics and instruments, which will be, at a minimum, ~150 to 200 watts of electricity (We). This requirement, in turn, is most easily fulfilled at large solar distances with a radioisotope power supply (RPS) (rather than with a nuclear fission reactor), the distances being so large as to make solar arrays totally inapplicable. By increasing the power supply one can, in principle, provide power to run ion engines, providing a constant thrust to the spacecraft.

This approach implies that a power conditioning system, ion engines, larger power supply (up to ~1 kWe), and appropriate propellant must all be supplied. Such an approach has been studied for implementation on both a Delta IV Heavy[26-28] and an Ares V launch vehicle[29]. With an Ares V and a Centaur upper stage the flyout time to ~200 AU remains ~28 years for realistic technologies. About 5 years can be trimmed off this amount by also including a gravity assist at Jupiter. For a specific target region on the sky, this also implies limiting the optimal launch windows to about every twelve years. For all of these cases, the probe moves from Earth orbit away from the Sun, so there are no increased thermal requirements as with the approaches of the previous section.

1.4 Spacecraft

The five solar-system-escaping spacecraft to date (Table I) all share the use of a high-energy "kick stage"; a large, spacecraft-fixed high-gain antenna (HGA), an RPS powered by plutonium-238 (^{238}Pu); redundant, fault-tolerant spacecraft electronics; low onboard, propulsive capability; and, single-string instruments. Similar features characterize the Innovative Interstellar Explorer (IIE) concept, with the exception of the lack of large onboard propulsion capability[28,29]. The Pioneer 10 and 11 spacecraft far outlasted their 900-day mission requirement, and the Voyagers their five-year mission requirement; New Horizons is now in its fourth year of flight in a nominal 16-year mission[30]. The IHP/HEX solar-sail spacecraft has subsystems similar to those of Ulysses[16]. These previous efforts demonstrate that an Interstellar Probe can be implemented in the near term to pass through the terminations shock, the heliopause, and into the possibly-shocked interstellar medium. Besides directly detecting the low-energy component of galactic cosmic rays, the interstellar magnetic field (ISMF), and neutral and particle composition and dynamics, the new questions in this region will also be answered (§2).

Spacecraft	Instruments		Spacecraft (dry) (kg)	Payload mass fraction
	Number	Mass (kg)		%
Voyager	10	104.32	721.90	14.45
Pioneer	11	28.98	251.79	11.51
New Horizons	6	28.43	385.00	7.38
IHP	12	25.6	517	4.95
IIE	10	35.2	516.2	6.82

Table I: Deep-space spacecraft, instruments, and their mass fractions. Voyager and New Horizons totals are from the National Space Science Data Center (NSSDC). Pioneer totals are from[31],, IHP totals are from[32] and IIE totals (option 2) are from[28]. The small mass fractions on IIE and IHP are driven by dry mass associated with the propulsion systems.

1.5 Payload

The mass limitations inherent in designing a spacecraft for fast solar-system escape imply the need for judicious selection of instrumentation. The new Voyager, IBEX, and Cassini results provide better insight into the types of instruments to fly and the magnitude of the expected signals. Prior to these results, a sample payload was worked out for IIE[28], based largely upon results for the Pioneers, Voyagers, Ulysses, and other field-and-particles robotic spacecraft. The instrument resources in the IIE payload need to be rethought, given these measurements and ongoing instrumentation developments[16,17]. For example, one would like to include neutral-atom imagers of similar sensitivity to those on IBEX[33] to follow the incoming ISM flow as it becomes increasingly pristine and to provide a tomographic view of the heliospheric boundary layers together with concurrent 1 AU observations, as well as a very sensitive plasma composition capability.

Fields and particles instruments from various flight missions are listed in Table II along with notional IIE and IHP payloads. These need re-examination in light of the recent science results.

Spacecraft ➔ / Instrument ↓	IIE	IHP	Helios	Pioneer	Voyager	New Horizons	Ulysses	IBEX	STEREO
Vector helium magnetometer	8.81	1.5	4.40	2.7	5.6		2.332		
Fluxgate magnetometer			4.75	0.3			2.4		0.27
Plasma wave sensor	10.0	5.8	NA		9.1		7.4		13.23
Plasma	2.00	2	15.696	5.5	9.9	3.3	6.7		2.37
Plasma composition		1.5					5.584		11.4
Energetic particle spectrometer	1.50	3.0	3.50	3.3	7.5	1.5	5.8		1.63
Cosmic-ray spectrometer: anomalous and galactic cosmic rays	3.50	3.5		3.2	7.5		14.6		1.92
Cosmic-ray spectrometer: electrons/positrons, protons, helium	2.30	1.5	7.15	1.7					1.98
Geiger tube telescope				1.6					
Meteoroid detector			8.93	3.2					
Cosmic dust detector	1.75	1.1		1.6		1.6	3.8		
Solar X-rays and gamma-ray bursts							2.0		
Neutral atom detector	2.50							12.1	
Energetic neutral atom detector	2.50	4.5					4.3	7.7	
Lyman-alpha detector / UV measurements	0.30	1.2		0.7	4.5	4.4			
Infrared measurements				2.0	19.5				
Imaging photopolarimeter				4.3	2.6	8.6			
Imaging system			8.93		38.2	10.5			48.1
Common electronics, harness, boom, etc.								5.4	19.1
Totals	35.2	25.6	72.2	30.1	104.4	29.9	54.9	25.2	100.0

Table II. Instrument masses on deep-space, robotic spacecraft.
Notes: Instrument names are from IIE (Table 2 of [28].). Equivalences to payload elements on other spacecraft are notional and sometimes very divergent with respect to capabilities; they provide a rough guide only. Pioneer instrument masses are from [31], Voyager masses from the NSSDC, New Horizons masses from [34], Ulysses masses from *Astron. Astrophys. Suppl. Series*, 92, 207 *et seq.*, 1992; IBEX masses from [33]; STEREO masses from *Space Sci. Rev.* 136 (1-4) 2008; Helios masses from *Raumfahrtforschung*, Band 19, Heft 5, September/Oktober 1975 (not all are available; marked "NA"); and, IHP masses from [32].

2 Advancing Solar and Space Physics

Davis[35] first called attention to the possible modification of the local interstellar medium by solar activity prior to the postulation[36] and confirmation[37] of the near-constant supersonic solar wind. The effects of this action on the nature of near-Sun space has been a subject of scientific speculation ever since[38,39].

All five spacecraft with speeds sufficiently high to escape the solar system are planetary missions but include instrumentation capable of making *in situ* particle and/or field measurements relevant to probing the nature of the Sun's interaction with the very local interstellar medium (VLISM)[§]. The earlier two spacecraft, Pioneer 10 and 11, launched 2 March 1972 and 5 April 1973, fell silent on 23 January 2003 and 30 September 1995, respectively[41], while both were still inside the termination shock. Voyager 1 and 2, launched 5 September and 22 August 1977, continue to return data, now from the heliosheath, having crossed the termination shock of the solar wind at 94.0 AU on 16 December 2004[42] and at 83.7 AU on 30 August 2007[43], respectively. Both should continue to return data until at least ~2020[44]. Voyager 1 is the fastest (~3.6 AU/yeart) and most distant of the five. The New Horizons spacecraft launched to Pluto on 19 January 2006, remains on course for that object and a Kuiper Belt Object (KBO) beyond at roughly the same heliographic longitude as Voyager 2, but near the plane of the ecliptic[45].[¶]

Following the Pioneer 10 and 11 Jupiter flybys, there was an initiative for flying a dedicated "interstellar precursor mission"[2,8,46-48]. Subsequently, the scientific rationale for such a mission has been repeated in a number of NASA and National Academy of Sciences documents (see §4.2 below). However, the ongoing *in situ* measurements by the Voyagers in the heliosheath (the Voyager Interstellar Mission (VIM), coupled with remote measurements of energetic neutral atoms (ENAs) from the interaction region (from the Interstellar Boundary Explorer (IBEX)[49-54] and Cassini missions[55-57]) have shown that the scientific need for new measurements with modern instruments from the remote reaches of the heliosphere are even more compelling than had been thought. Indeed these recent and ongoing observations have revealed that the interaction of the heliosphere with the VLISM is much more complex than heretofore assumed by our present-day concepts. These puzzling discoveries also call for a major re-examination of the strategy for the Interstellar Probe mission.

The *in situ* instruments on Voyager 1 and Voyager 2 continue to reveal significant fluxes of energetic particles in the heliosheath, including a well-defined suprathermal ion "tail" in which the differential intensities fall off ~$E^{-1.5}$ above ~30 keV[58]. At even higher energies (~100 MeV), there is no "unfolding" of the energy spectrum of the anomalous cosmic rays (ACRs), thus pointing to a more remote location for the modulation region and source[59]. The plasma flow in the heliosheath is peculiar; at Voyager 1, some 20 AU into the heliosheath, the radial-flow velocity component (V_r) is trending towards zero. Most strikingly, direct measurements of the shocked, solar-wind flow speed obtained from Voyager 2 revealed that the core flow remains supersonic in the heliosheath beyond the termination shock[60], contrary to previous thought. All of these particle observations at both Voyagers, taken together, unambiguously imply that the bulk of the energy density in the plasma resides in a non-thermal component that extends to very high energies. Strong implications, both quantitative and qualitative, follow from this fact for the overall heliosheath structure. We have never encountered a large-scale plasma regime in which the non-thermal ion pressure dominates the thermal pressure and overwhelms the magnetic field stresses. The closest parallel regime lies in localized regions of planetary magnetospheres during extremely disturbed conditions, but in the heliosheath these conditions always exist everywhere. This means that no simple distribution function (neither maxwellian nor "kappa" distribution) is adequate to describe the essential physics. This is why even sophisticated MHD models failed to predict anything like the striking new features that have just been observed in the last two years.

Then, in 2009, remote sensing of the heliosheath proton population by IBEX and the Ion and Neutral Camera (INCA) on Cassini revealed stunningly unexpected structures on a variety of scales[52,57]. In addition to the general "glow" of the sky in ENAs, IBEX data show a relatively

[§] Generally taken as 0.01 parsecs (pc) or 2,063 astronomical units (AU) so as to be outside the limit of influence of the Sun[40].

[¶] Unfortunately, the available, but old, ^{238}Pu used in the power supply on that mission will likely limit the mission life to sometime before 2040 at a heliocentric distance of less than 100 AU.

narrow "ribbon" of atomic hydrogen emission from ~200 eV to ~6 keV, roughly circular, but asymmetric in intensity, suggesting that it is ordered by the interstellar magnetic field (ISMF). It passes through, rather than being centered on, the direction towards the "nose" where the interstellar plasma flow around the heliosphere stagnates, suggesting that the flow is not the primary driver of the system as has been thought, but rather it is the pressure of the interstellar field that configures the heliosheath. The hydrogen ENAs from both the glow and ribbon are also characterized by non-thermal distribution functions. INCA on Cassini sees at higher energies (10s of keV) a "belt" of emission in ENAs, broader than the ribbon, and tilting significantly away from it, and exhibiting a much steeper energy spectrum than observed in the IBEX energy range[56]. Further, an estimate of total plasma pressure in the heliosheath, combining the Voyager *in situ* and the ENA remote measurements[56] suggests an ISMF strength < 0.6 nT, higher than the previously estimated values of ~ 0. 25 nT.

Attempts to explain consistently all the afore-mentioned fascinating observations exhibit no clear trend towards a consensus. All the diverse *in situ* and remote observations obtained to date only serve to emphasize the need for a new generation of the more comprehensive measurements that will be required to understand the global nature of our Sun's interaction with the local galactic environment. Only an Interstellar Probe with modern instruments and measurement requirements better defined by these recent observations can provide the critical information required. We now know that we are dealing with a strongly non-thermal plasma, and that the complex large and small scale structures in the heliosheath are produced by physical processes yet to be adequately described by theory or simulation.

3 Estimating the Cost

An Interstellar Probe mission pushes on the technical limits of the three enablers for any deep space mission traveling away from the Sun: a highly capable, yet "affordable" launch vehicle, electric power from a low-specific-mass RPS, and reliable, sensitive, deep-space communications at Ka-band. All three elements of this robotic-mission-infrastructure triad are necessary for such a mission to take place; they also can be available in the coming decade.

Launch vehicle prices must be negotiated for any given mission. For the REP approach, a large vehicle with an energetic upper stage is required. Delta IV H costs have likely not decreased from their reported level of $254M in late 2004[61]. While Star 48A motors are available, the type of custom stack required for this application would require design work. The overall stack could likely be produced for less than $400M, including the National Environmental Protection Act (NEPA) costs associated with use of an RPS for spacecraft power. An appropriate launch vehicle for the IHP/HEX-solar-sail approach would likely be in the ~$150M or less category, including NEPA, provided that current mass and initial launch energy (C3 ~ 0) requirements hold.

From past usage of RPSs, we estimate a price of ~$30M per unit (cf. Discovery 2010 Announcement of Opportunity, Amendment I, §5.9.3) with 6 required for the REP approach and 2 for a solar-sail spacecraft. Key to this and other deep-space robotic missions is restarting the production of Pu-238[62]. Estimated power costs are then ~$180M and ~$60M, respectively.

The aperture fee for use of NASA's Deep Space Network (DSN) can be estimated using a variety of NASA resource tools. For three 8-hr tracks/week with the 34-m, high-efficiency (HEF) antennas following a Jupiter flyby some five years after launch (with more frequent tracks earlier in the mission), we estimate 30-year tracking costs (FY2025$) of ~$70M (Waldherr, private comm., 28 Oct 2010)[§§]. Mission-operations-and-data-analysis (MO&DA) costs of ~$5M to $10M per year compare well with recent costs for Ulysses and VIM.

For 12 instruments at an average cost of $15 M each and $500 M for the spacecraft bus, the basic unit for either propulsion approach should be buildable for ~$950M, including a 40%

[§§] S. Waldherr, DSN Commitments Engineer, DSN Mission Commitments Office.

reserve. For a solar-sail implementation, we might guess that a ~$150M precursor mission with an additional ~$50M in development may be needed to prove out the required technology.

Tallying these hypothetical costs for the two approaches suggests the Delta IV H/Star 48/REP mission could be built for ~$1.1 B (spacecraft + instruments + 6 RPS + reserve) and launched for another $500M, a total of ~$1.6B. A solar-sail unit (spacecraft + instruments + 2 RPS + reserve) could be built for $1.0B, qualified for ~$200M and launched for another ~$100M, a total of $1.3B, if the technology works out. Running and tracking either mission for the first 10 years could total ~$25M for the Deep Space Network (DSN) and ~$50M to $100M for ten years of operations in fixed year costs[¶].

4 Meeting Heliophysics Decadal Survey Criteria

4.1 Identification as a high priority or requirement in previous studies or roadmaps

In the U.S. the scientific case for an interstellar precursor mission continues to be made in reports by both NASA (National Aeronautics and Space Administration) and the NRC acting under the National Academy of Sciences (NAS). The latter have included:

1. Physics through the 1990's - Panel on Gravitation, Cosmology, and Cosmic Rays (D. T. Wilkinson, chair), 1986 NRC report
2. Solar and Space Physics Task Group Report (F. Scarf, chair), 1988 NRC study Space Science in the 21st Century - Imperatives for the Decade 1995-2015
3. Astronomy and Astrophysics Task Group Report (B. Burke, chair), 1988 NRC study Space Science in the 21st Century - Imperatives for the Decade 1995-2015
4. The Decade of Discovery in Astronomy and Astrophysics (John N. Bahcall, chair)
5. The Committee on Cosmic Ray Physics of the NRC Board on Physics and Astronomy (T. K. Gaisser, chair), 1995 report Opportunities in Cosmic Ray Physics
6. A Science Strategy for Space Physics, Space Studies Board, NRC, National Academy Press, 1995 (M. Neugebauer, chair)
7. The Sun to the Earth - and Beyond: A Decadal Research Strategy in Solar and Space Physics, 2003
8. Exploration of the Outer Heliosphere and the Local Interstellar Medium, 2004
9. Priorities in Space Science Enabled by Nuclear Power and Propulsion, 2006

Past NASA documents and reports include:

1. Outlook for Space, 1976
2. An Implementation Plan for Solar System Space Physics, S. M. Krimigis, chair, 1985
3. Space Physics Strategy-Implementation Study: The NASA Space Physics Program for 1995-2010
4. Sun-Earth Connection Technology Roadmap, 1997
5. Space Science Strategic Plan, The Space Science Enterprise, 2000
6. Sun-Earth Connection Roadmaps, 1997, 2000, 2003
7. NASA 2003 Strategic Plan
8. The New Science of the Sun - Solar System: Recommended Roadmap for Science and Technology 2005 - 2035, 2006

The most recent NAS/NRC document advocating the mission[63] recommended "NASA should conduct further study of the following mission concepts, which have the most potential to demonstrate the scientific opportunities provided by the Constellation System: …, Interstellar Probe, …."

The explanatory note[*] to the proposal team in the Cosmic Vision 2015 – 2025 competition explaining the rejection of the proposed solar-sail version of the mission (which included a substantial proposed NASA collaboration) noted that this is "…an innovative mission that addressed our place in the universe and which should be done at some stage…." This mission was endorsed in NASA's Heliophysics Roadmap[65] as a potential mission for an international

[¶] N.B. the launch plus power plus DSN costs are dependent upon NASA infrastructure.

[*] "The [reviewers] considered the concept of a mission to the outer heliosphere to be extremely interesting…. The main issues are with the timeliness of the main science return from the mission, the technical feasibility of some of the elements and the need to preserve technical information across several generations of scientists/engineers." Technical feasibility concerns primarily centered on the 60,000 m^2 solar sail and what was seen as a required demonstrator mission to prove the concept and implementation. A secondary concern was the need for "very efficient" RPSs, which would be a NASA contribution. Efficient RPSs remain a recognized issue, but that problem is being addressed[62] and the mission flyout time to a given distance is contemplated as being faster than that of the Voyagers, but it will still be long. On the other hand, other missions including those of the Voyagers, Ulysses, and the Solar and Heliospheric Observatory (SOHO) have demonstrated that the appropriate maintenance of corporate knowledge across multiple decades is a manageable problem[64].

partnership[†].

4.2 Makes a significant contribution to more than one of the Panel themes

Interstellar probe is primarily a solar and heliophysics mission/investigation. To the extent that it can uncover more information about the conditions in the very local interstellar medium (VLISM) and how these are modified in the interaction region, it *may* provide more information on the spectra of galactic cosmic rays and how they evolve in reaching human-occupied space and planetary magnetospheres. Any contribution will be indirect, but could be significant.

4.3 Contributions to important scientific questions facing solar and space physics today

The exact formulation of the science questions has varied with particular studies. The formulation used with the NASA Vision Mission study (IIE) is:

1. What is the nature of the nearby interstellar medium?
2. How do the Sun and galaxy affect the dynamics of the heliosphere?
3. What is the structure of the heliosphere?
4. How did matter in the solar system and interstellar medium originate and evolve?

This set of questions (from the 3rd Interstellar Probe Science and Technology Definition Team Meeting, 17-19 May 1999) feeds into objectives and questions articulated in NASA's IPSTDT Report and was used to establish a Traceability Matrix for the mission[28]. In the recent NASA Roadmap[65], the science context flows from Research Focus Areas (RFAs) under each of the three broad science objectives in that report, all of which couple to the priority investigations of determining:

1. What is the composition of matter fundamental to the formation of habitable planets and life?
2. How do the heliosphere and the interstellar medium interact?
3. What is the magnetic structure of the Sun-heliosphere system?

as well as to Decadal Survey Challenge 2: "Understanding heliospheric structure, the distribution of magnetic fields and matter throughout the solar system, and the interaction of the solar atmosphere with the local interstellar medium."

Along with the Voyager Interstellar Mission (VIM), comprised of Voyagers 1 and 2 launched in 1977, the Advanced Composition Explorer (ACE) mission launched in 1997, and the Interstellar Boundary Explorer (IBEX), launched in 2008, and now providing paradigm-shifting results [49,50,52-54,57], the Interstellar Probe is called out as fundamental to addressing these questions.

Similarly, the IHP/HEX mission has had the announced science goals of [16,17,64]

1. How do solar wind and interstellar medium interact to form the heliosphere and how does this relate to the universal phenomenon of the formation of astrospheres?
2. What are the properties of the very local interstellar medium and how do they relate to the typical ISM?
3. How do plasma, neutral gas, dust, waves, particles, fields, and radiation interact in extremely rarefied, turbulent, and incompletely ionized plasmas?
4. What is the cause of the Pioneer Anomaly?[66] (a potential "bonus" science goal).

While emphasizing different features and levels of detail, all of these formulations point to the science of our local neighborhood in the galaxy with an emphasis on what it, and the Sun's interaction with it can tell us about ourselves, our solar system, and the what lies beyond.

4.4 Contributions to applications and/or policy making

There are no direct contributions to societal applications or imminent policy decisions.

[†]"The nature of composition and dynamics of the interstellar medium are among the highest ranked science questions in heliophysics. No international partnership opportunity to explore the interstellar boundary is known at this time. Were it to materialize, a spacecraft directly sampling the environment outside the heliosphere could address these questions.

"The next logical step in exploration would be to directly sample the medium that lies beyond the extended solar atmosphere. The solar wind and magnetic field keep the unique plasma of the interstellar medium outside the heliosphere. A partnership mission to interstellar space would allow us to sample its unique dynamics and composition and to access the regime of low-energy cosmic rays that helps us understand cosmic particle acceleration processes for the first time." Indeed, "the [Roadmap] team identified one high-priority science target for a potential international partnership, interstellar mission…." (p.64 of [65]).

4.5 Complementing other observational systems or programs available

An interstellar probe will complement the entire set of heliospheric spacecraft by providing the "outer boundary condition" for the heliosphere. Launch during the operational lifetime of the Voyagers or a New Horizons extended mission, i.e., within the end of the decade under consideration, would, of course, provide an even tighter complementarily with fields and particles observations from those platforms.

4.6 Cost-benefit

The spacecraft and instrument hardware are not the primary drivers, except indirectly in terms of mission lifetime requirements. Development through launch will not cost less than New Horizons but probably not more than Cassini (§3). The question of what the mission would really cost requires further study.

4.7 Degree of readiness (technical, resources, people)

An Interstellar probe has consistently been rated as technically ready using near-term instrumentation. The issue is that of propulsion. Recent work has indicated that neither ballistic nor NEP approaches are credible. Further full-up study is required to determine credible flyout times for REP alone, REP in conjunction with SEP, and solar sail approaches; initial studies suggest they are all feasible (see §1 and included references).

4.8 Fit with other national and international plans and activities

A European-led international team proposed the similar IHP/HEX mission under the Cosmic Vision 2015 – 2025 competition as noted in §4.1.

5 The International Interstellar Probe Team

The International Interstellar Probe Team is an open group of 92 scientists and engineers from 16 countries on 5 continents dedicated to reaching the beginnings of interstellar space[††].

[††] Signatories to this white paper follow:

1	Joe Mazur	USA	Joseph.E.Mazur@aero.org	47	Kelvin Long	UK	kelvin.long@tesco.net
2	Jonathan Slavin	USA	jslavin@cfa.harvard.edu	48	Timothy Horbury	UK	t.horbury@imperial.ac.uk
3	Matthew E. Hill	USA	Matthew.Hill@jhuapl.edu	49	Robert J. Forsyth	UK	r.forsyth@imperial.ac.uk
4	Alexander Lazarian	USA	alazarian@facstaff.wisc.edu	50	Malcolm Macdonald	UK	malcolm.macdonald.102@strath.ac.uk
5	Eberhard Moebius	USA	Eberhard.Moebius@unh.edu	51	Richard Marsden	The Netherlands	Richard.Marsden@esa.int
6	Stuart D. Bale	USA	bale@ssl.berkeley.edu	52	Wing-Huen Ip	Taiwan	wingip@astro.ncu.edu.tw
7	Merav Opher	USA	mopher@gmu.edu	53	Peter Bochsler	Switzerland	bochsler@soho.unibe.ch
8	Nathan Schwadron	USA	n.schwadron@unh.edu	54	Stas Barabash	Sweden	stas@irf.se
9	Dave McComas	USA	DMcComas@swri.edu	55	Adri Burger	South Africa	Adri.Burger@nwu.ac.za
10	Priscilla Frisch	USA	frisch@oddjob.uchicago.edu	56	Marius Potgieter	South Africa	Marius.Potgieter@nwu.ac.za
11	Frederic Allegrini	USA	fallegrini@swri.edu	57	Vladislav Izmodenov	Russia	vlad.izmodenov@gmail.com
12	Karlheinz Trattner	USA	trattner@spasci.com	58	Yury Malama	Russia	malama@ipmnet.ru
13	Pauette Liewer	USA	paulett.c.liewer@jpl.nasa.gov	59	Dmitry Aleksashov	Russia	alexash@ipmnet.ru
14	Ralph L. McNutt, Jr.	USA	ralph.mcnutt@jhuapl.edu	60	Sergey V. Chalov	Russia	chalov@ipment.ru
15	Edmond C. Roelof	USA	Edmond.Roelof@jhuapl.edu	61	Vladimir B. Baranov	Russia	baranov@ipmnet.ru
16	Michael Gruntman	USA	mikeg@usc.edu	62	Olga Katushkina	Russia	okat@iki.rssi.ru
17	Stamatios M. Krimigis	USA	Tom.Krimigis@jhuapl.edu	63	Elena Provornikova	Russia	provea@mail.ru
18	Randall Smith	USA	rsmith@cfa.harvard.edu	64	Mikhail I. Panasyuk	Russia	panasyuk@sinp.msu.ru
19	Mark Wiedenbeck	USA	mark.e.wiedenbeck@jpl.nasa.gov	65	Yuri Yermolaev	Russia	yermol@hotbox.ru
20	Pontus C. Brandt	USA	Pontus.Brandt@jhuapl.edu	66	Stan Grzedzielski	Poland	stangrze@cbk.waw.pl
21	David Humm	USA	David.Humm@jhuapl.edu	67	Roma Ratkiewicz	Poland	roma@cbk.waw.pl
22	Nicola J. Fox	USA	Nicola.Fox@jhuapl.edu	68	Luca Sorriso-Valvo	Italy	lucasorriso@gmail.com
23	Gregory Matloff	USA	GMatloff@CityTech.Cuny.Edu	69	Gaetano Zimbardo	Italy	gaetano.zimbardo@fis.unical.it
24	John D. Richardson	USA	jdr@space.mit.edu	70	Marco Casolino	Italy	Marco.Casolino@roma2.infn.it
25	Bryan A. Palaszewski	USA	bryan.a.palaszewski@nasa.gov	71	Karl-Heinz Schartner	Germany	Karl-Heinz.Schartner@exp1.physik.uni-giessen.de
26	David J. Lawrence	USA	David.J.Lawrence@jhuapl.edu	72	Horst W. Loeb	Germany	H.W.Loeb@exp1.physik.uni-giessen.de
27	Brian Wood	USA	brian.wood@nrl.navy.mil	73	Bernd Dachwald	Germany	dachwald@fh-aachen.de
28	Alan Cummings	USA	ace@srl.caltech.edu	74	Andreas Ohndorf	Germany	andreas.ohndorf@dlr.de
29	William Kurth	USA	william-kurth@uiowa.edu	75	Wolfgang Seboldt	Germany	wolfgang.seboldt@dlr.de
30	Louis J. Lanzerotti	USA	ljl@njit.edu	76	Dieter Breitschwerdt	Germany	breitschwerdt@astro.physik.tu-berlin.de
31	Peter Wurz	USA	peter.wurz@space.unibe.ch	77	Klaus Scherrer	Germany	dat-hex@t-online.de
32	George C. Ho	USA	George.Ho@jhuapl.edu	78	Horst Fichtner	Germany	hf@tp4.rub.de
33	Gary P. Zank	USA	garyp.zank@gmail.com	79	Eckart Marsch	Germany	marsch@linmpi.mpg.de
34	Robert Henshaw	USA	Bob.Henshaw@jhuapl.edu	80	Karl Mannheim	Germany	asti070@mail.uni-wuerzburg.de
35	Nikolaos Paschalidis	USA	Nick.Paschalidis@jhuapl.edu	81	Ralf Srama	Germany	ralf.srama@mpi-hd.mpg.de
36	Adam Szabo	USA	adam.szabo-1@nasa.gov	82	Bernd Heber	Germany	heber@physik.uni-kiel.de
37	Carey M. Lisse	USA	Carey.Lisse@jhuapl.edu	83	Wolfgang Droege	Germany	Wolfgang.Droege@astro.uni-wuerzburg.de
38	Eric R. Christian	USA	eric.r.christian@nasa.gov	84	Robert F. Wimmer-Schweingruber	Germany	wimmer@physik.uni-kiel.de
39	Mihir I. Desai	USA	mdesai@swri.edu	85	Milan Maksimovic	France	milan.maksimovic@obspm.fr
40	Christopher Paranicas	USA	Chris.Paranicas@jhuapl.edu	86	Jean-Louis Bougeret	France	jean-louis.bougeret@obspm.fr
41	Robert E. Gold	USA	robert.gold@jhuapl.edu	87	Rosine Lallement	France	rosine.lallement@latmos.ipsl.fr
42	Edgar Rhodes	USA	Ed.Rhodes@jhuapl.edu	88	Pekka Janhunen	Finland	pekka.janhunen@fmi.fi
43	Robert Decker	USA	Rob.Decker@jhuapl.edu	89	Timo Laitinen	Finland	timo.laitinen@utu.fi
44	Steven D. Howe	USA	Steven.Howe@inl.gov	90	Ingrid Mann	Belgium	imann@uni-muenster.de
45	Louis Friedman	USA	tps.ldf@planetary.org	91	Helmut O. Rucker	Austria	helmut.rucker@oeaw.ac.at
46	Jakobus le Roux	USA	jar0013@uah.edu	92	Iver Cairns	Australia	cairns@physics.usyd.edu.au

6 References

1 Fiehler, D. I. & McNutt, R. L., Jr. Mission design for the Innovative Interstellar Explorer vision mission. *J. Spacecraft* **43**, 1239-1247, doi:DOI: 10.2514/1.20995 (2006).
2 Jaffe, L. D. *et al.* An interstellar precursor mission. *J. Brit. Int. Soc.* **33**, 3-26 (1980).
3 Pawlik, E. V. & Phillips, W. M. A nuclear electric propulsion vehicle for planetary exploration. *J. Spacecraft* **14**, 518-525 (1977).
4 Nock, K. T. in *19th AIAA/DGLR/JSASS International Electric Propulsion Conference.* 7.
5 NRC. Priorities in Space Science Enabled by Nuclear Power and Propulsion. (The National Academies Press, Washington, D.C., 2006).
6 Zurbuchen, T. H. *et al.* in *NASA Space Science Vision Missions* Vol. 224 *Progress in Astronautics and Aeronautics* (ed Marc S. Allen) 434 (2008).
7 Greene, S. R. in *American Nuclear Society* 10 (San Diego, California, 2005).
8 Holzer, T. E., Mewaldt, R. A. & Neugebauer, M. The Interstellar Probe: Scientific objectives for a Frontier mission to the heliospheric boundary and interstellar space. (NASA, Ballston, VA, 1990).
9 Mewaldt, R. A., Kangas, J., Kerridge, S. J. & Neugebauer, M. A small interstellar probe to the heliospheric boundary and interstellar space. *Acta Astron.* **35, Suppl.** (1995).
10 McNutt, R. L., Jr. *et al.* A realistic interstellar explorer. *Adv. Space Res.* **34**, 192-197 (2004).
11 McNutt, R. L., Jr. *et al.* in *Solar Wind X*. eds M. Velli & et al.) 830-833 (American Institute of Physics, NY).
12 McNutt, R. L., Jr. *et al.* in *Space Technology and Applications International Forum (STAIF)-2000.*
13 McNutt, R. L., Jr. *et al.* in *COSPAR Colloquium on The Outer Heliosphere: The Next Frontiers COSPAR Colloquia Series.* eds K. Scherer, H. Fichtner, H.-J. Fahr, & E. Marsch) 431-434 (Pergamon Press, New York).
14 McNutt, R. L., Jr. *et al.* Low-cost interstellar probe. *Acta Astron.* **52**, 267-279 (2003).
15 Lyman, R. W., Ewing, M. E., Krishnan, R. S., Lester, D. M. & McNutt, R. L., Jr. in *37th AIAA/ASME/SAE/ASE Joint Propulsion Conference.*
16 Wimmer-Schweingruber, R. F. *et al.* Interstellar heliospheric probe/heliospheric boundary explorer mission—a mission to the outermost boundaries of the solar system. *Exp. Astron.* **24**, 9-46, doi:DOI 10.1007/s10686-008-9134-5 (2009).
17 Wimmer-Schweingruber, R. F., McNutt, R. L., Jr. & IHP/HEXTeam. The Interstellar Heliopause Probe: Heliospheric Boundary Explorer Mission to the interstellar medium. *Earth Moon Planet* **104**, 17-24, doi:DOI 10.1007/s11038-008-9249-8 (2009).
18 Liewer, P. C. *et al.* in *COSPAR Colloquium on The Outer Heliosphere: The Next Frontiers COSPAR Colloquia Series* Vol. 11 *COSPAR Colloquia Series* eds K. Scherer, H. Fichtner, H.-J. Fahr, & E. Marsch) 411-420 (Pergamon Press, 2001).
19 Liewer, P. C., Mewaldt, R. A., Ayon, J. A. & Wallace, R. A. in *Space Technology and Applications International Forum (STAIF)-2000.*
20 Noble, R. J. Radioisotope electric propulsion of sciencecraft to the outer solar system and near-interstellar space. *Acta Astron.* **44**, 193-199 (1999).
21 Noble, R. J. in *Missions to the outer solar system and beyond; 2nd IAA Symposium on Realistic Near-Term Advanced Scientific Space Missions.* 121-127 (International Academy of Astronautics).
22 Loeb, H. W., Schartner, K.-H., Dachwald, B., Ohndorf, A. & Seboldt, W. in *Space Propulsion 2010* (San Sebastian, Spain, 2010).
23 Sauer, C. G. in *AIAA/AAS Astrodynamics Conference*, Vol. 76-792 (San Diego, CA, 1976).
24 Dachwald, B. Optimal solar-sail trajectories for missions to the outer solar system. *Journal of Guidance, Control, and Dynamics* **28**, 1187 - 1193 (2005).
25 Dachwald, B., Seboldt, W. & Lämmerzahl, C. in *Lasers, Clocks, and Drag-Free Control - Exploration of Relativistic Gravity in Space Astrophysics and Space Science Library* eds Dittus, Lämmerzahl, & S. G. Turyshev) 365-381 (Springer, 2008).
26 McNutt, R. L., Jr. *et al.* in *IGPP 5th Annual International Astrophysics Conference.*
27 McNutt, R. L., Jr. *et al.* in *Solar Wind 11 - SOHO 16.* 693-696 (European Space Agency).
28 McNutt, R. L., Jr. *et al.* in *41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit.* (AIAA).
29 McNutt Jr, R. L. & Wimmer-Schweingruber, R. F. Enabling interstellar probe. *Acta Astronautica* **In Press, Corrected Proof** (2010).
30 Fountain, G. H. *et al.* The New Horizons Spacecraft. *Space Science Reviews* **140**, 23-47 (2008).
31 Christiansen, R. A., Jackson, R. W., Sperans, J., Tendeland, T. & Wolfe, J. H. Pioneer H Jupiter Swingby Out-of-the-Ecliptic Mission Study - Final Report. Report No. NASA-TM-108108, 128 (Ames Research Center, Moffett Field, CA 94035, 1971).
32 Wimmer-Schweingruber, R. F. & Team, R. L. M. J. I. H. P. H. E. X. The Interstellar Heliopause Probe/Heliospheric Explorer: IHP/HEX. *AIP Conference Proceedings* **1216**, 655-658 (2010).
33 McComas, D. J. *et al.* IBEX - Interstellar Boundary Explorer. *Space Sci. Rev.* **146**, 11-33, doi:10.1007/s11214-009-9499-4 (2009).
34 Russell, C. T. 402 (Springer, 2009).
35 Davis, L. Interplanetary Magnetic Fields and Cosmic Rays. *Physical Review* **100**, 1440 (1955).
36 Parker, E. N. Dynamics of the interplanetary gas and magnetic fields. *Astrophys. J.* **128**, 664-676 (1958).
37 Neugebauer, M. & Snyder, C. W. Solar Plasma Experiment. *Science* **138**, 1095-a-1097 (1962).
38 Parker, E. N. The stellar-wind regions. *Astrophys. J.* **134**, 20-27 (1961).
39 Zank, G. P. Interaction of the solar wind with the local interstellar medium: A theoretical perspective. *Space Sci. Rev.* **89**, 1-275 (1999).
40 Suess, S. T. The heliopause. *Reviews of Geophysics* **28**, 97-115 (1990).
41 NASA-ARC. *The Pioneer Missions*, <http://www.nasa.gov/centers/ames/missions/archive/pioneer.html> (2007).
42 Stone, E. C. *et al.* Voyager 1 explores the termination shock region and the heliosheath beyond. *Science* **309**, 2017-2020 (2005).
43 Stone, E. C. *et al.* An asymmetric solar wind termination shock. *Nature* **454**, 71-74, doi:doi:10.1038/nature07022 (2008).

44 NASA-JPL. *Voyager, The Interstellar Mission*, <http://voyager.jpl.nasa.gov/index.html> (2010).

45 McNutt, R. L. *et al.* The Pluto Energetic Particle Spectrometer Science Investigation (PEPSSI) on the New Horizons Mission. *Space Science Reviews* **140**, 315-385 (2008).

46 Allen, J. P. *et al.* Outlook for Space: Report to the NASA Administrator by the Outlook for Space Study Group. Report No. NASA SP-386, (Technical Information Office, NASA, Washington, D. C., 1976).

47 Jaffe, L. D. & Ivie, C. V. Science aspects of a mission beyond the planets. *Icarus* **39**, 486-494 (1979).

48 Jaffe, L. D. & Norton, H. N. A prelude to interstellar flight. *Astro. Aero.* **18**, 38-44 (1980).

49 Funsten, H. O. *et al.* Structures and Spectral Variations of the Outer Heliosphere in IBEX Energetic Neutral Atom Maps. *Science* **326**, 964-966 (2009).

50 Fuselier, S. A. *et al.* Width and Variation of the ENA Flux Ribbon Observed by the Interstellar Boundary Explorer. *Science* **326**, 962-964 (2009).

51 McComas, D. J. & al., e. in *Physics of the Outer Heliosphere.* eds V. Florinski, N. V. Pogorelov, & G. P. Zank) 162-181 (AIP, Melville, NY).

52 McComas, D. J. *et al.* Global Observations of the Interstellar Interaction from the Interstellar Boundary Explorer (IBEX). *Science* **326**, 959-962 (2009).

53 Möbius, E. *et al.* Direct Observations of Interstellar H, He, and O by the Interstellar Boundary Explorer. *Science* **326**, 969-971 (2009).

54 Schwadron, N. A. *et al.* Comparison of Interstellar Boundary Explorer Observations with 3D Global Heliospheric Models. *Science* **326**, 966-968 (2009).

55 Krimigis, S. M. *et al.* Magnetospheric imaging instrument (MIMI) on the Cassini mission to Saturn/Titan. *Space Sci. Rev.* **114**, 233-329 (2004).

56 Krimigis, S. M., Mitchell, D. G., Roelof, E. C. & Decker, R. B. in *Pickup Ions: Throughout the Heliosphere and Beyond.* in press (Aip Conf.).

57 Krimigis, S. M., Mitchell, D. G., Roelof, E. C., Hsieh, K. C. & McComas, D. J. Imaging the Interaction of the Heliosphere with the Interstellar Medium from Saturn with Cassini. *Science* **326**, 971-973 (2009).

58 Decker, R. B., Krimigis, S. M., Roelof, E. C. & Hill, M. E. in *Pickup Ions: Throughout the Heliosphere and Beyond.* in press (AIP Conf.).

59 Richardson, J. D. & Stone, E. C. The solar wind in the outer heliosphere. *Space Sci. Rev.* **143**, 7-20 (2009).

60 Richardson, J. D., Kasper, J. C., Wang, C., Belcher, J. W. & Lazarus, A. J. Cool heliosheath plasma and deceleration of the upstream solar wind at the termination shock. *Nature* **454**, 63-66 (2008).

61 Astronautica, E. *Delta IV*, <http://www.astronautix.com/lvs/deltaiv.htm> (2009).

62 NRC. *Radioisotope Power Systems: An Imperative for Maintaining U.S. Leadership in Space Exploration*. (The National Academies Press, 2009).

63 NRC. *Launching Science: Science Opportunities Provided by NASA's Constellation System*. (The National Academies Press, 2008).

64 Wimmer-Schweingruber, R. F., McNutt, R. L., Jr. & IHP/HEX Team. in *Solar Wind 12, AIP Conf. Proc.* **1216**, 655-658 (2010).

65 Christensen, A. B. *et al.* Heliophysics, The Solar and Space Physics of a New Era: Recommended Roadmap for Science and Technology 2009–2030. Report No. NP-2009-08-76-MSFC, x+151 (National Aeronautics and Space Administration, George C. Marshall Space Flight Center, Huntsville, AL., 2009).

66 Anderson, J. D. *et al.* Study of the anomalous acceleration of Pioneer 10 and 11. *Phys. Rev. D* **65**, 082004 (2002).

The Importance of Polar Observations in Understanding the Solar Dynamo

M. Miesch[1], M. Rempel[1], A. Kosovichev[2], T. Sekii[3], H. Hara[3], T. Yokoyama[4], S. Brun[5], T. Tarbell[6], T. Appourchaux[7], B.P. Brown[8], J. Toomre[9]

[1]High Altitude Observatory, NCAR, [2]Stanford University, [3]National Astronomical Observatory of Japan, [4]University of Tokyo, [5]CEA Saclay, France, [6]Lockheed Martin Solar and Astrophysics Laboratory, [7] Institut d'Astrophysique Spatiale, France [8] University of Wisconsin,[9] JILA/University of Colorado

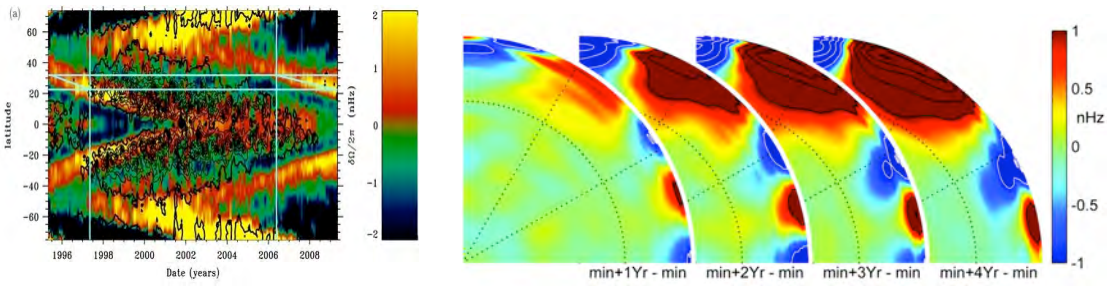


Fig 1: Torsional oscillations in the solar convection zone; Dynamical signatures of the solar dynamo (courtesy of Howe etal. 2009, ApJ 701, 87 and Vorontsov etal. 2002, Science 296, 101)

Abstract

The dynamics and magnetism of the polar regions of the Sun play a critical role in the operation of the global solar dynamo, which determines the variable strength of the solar activity cycle. This problem has become central for our understanding of the basic mechanisms underlying solar activity and for predicting the Sun's influences on the heliosphere, the geospace environment, and the Earth's climate system. However, investigation of the Sun's polar regions is a formidable challenge, and can be done only by an out-of-ecliptic solar mission. The key science problems required for solving the dynamo problem include observations of the photospheric magnetic flux distribution and evolution, and also helioseismic measurements of the differential rotation, meridional circulation, large-scale convection flows at the surface and in the interior. The observations must resolve three fundamental scales of the solar dynamics and activity: supergranulation, large magnetic flux structures, and global circulation of the convection zone. They should be done for a significant part of a solar cycle from an out-of-ecliptic orbit with at least 30 degrees inclination. These goals can be achieved with modest instrumental and telemetry requirements, and with the orbit options proposed for the Solar-C Plan A and Solar Polar Imager (SPI) missions.

Introduction

Understanding the origins of solar magnetic activity has been at the forefront of solar physics since the discovery of the 11-year sunspot cycle nearly two centuries ago. Unraveling this mystery has broad implications not only for promoting a deeper knowledge of the Sun itself but also for understanding the Sun's influences on the heliosphere, the geospace environment, and the Earth's climate system. Such influences regulate space weather, with increasing economic impacts on our technological society as our reliance on telecommunications systems, power grids, and airline travel continues to grow. As a readily observable example of an astrophysical magneto-hydrodynamic dynamo, the Sun also provides unique insights into the generation of magnetic fields by turbulent plasma flows throughout the universe, from planetary and stellar interiors to stellar and galactic accretion disks to interstellar clouds.

The global magnetic polarity of the Sun reverses during each 11-year sunspot cycle so the overall period of the solar magnetic activity cycle is 22 years. Understanding how such regularity arises from the highly turbulent conditions of the solar convection zone and how magnetic flux emerges from the solar interior to energize the solar atmosphere and power solar variability is a formidable challenge. Large-scale flows (differential rotation and meridional circulations) established by turbulent convection, plasma instabilities, and nonlinear feedbacks all play an important role, spanning many orders of magnitude in spatial and temporal scales.

Prediction of the solar sunspot cycles is not only of great practical importance, it also represents an ultimate test for our understanding of solar magnetism. So far, the predictions based on dynamo models have not been conclusive. *Solar cycle predictions based on measurement of the polar field strength appear promising but their reliability and the underlying physical mechanisms remain unclear*. The polar magnetic field reaches its maximum strength during the sunspot minima, and it correlates very well with the maximum sunspot number of the following solar cycle. In addition, there is evidence that the strength of geomagnetic indices measured during the solar minima may be correlated with the following sunspot maxima. This correlation is probably caused by the high-speed solar wind from the polar regions. However, how the polar magnetic field strength is translated into the future sunspot number is unknown. Thus, for developing physics-based forecasts of solar activity it is extremely important to investigate the structure and evolution of the polar magnetic fields, mechanisms of magnetic flux transport, and interaction between the old and new magnetic flux.

Modern solar observations coupled with sophisticated theoretical and numerical models have yielded profound insights into many aspects of solar magnetism but the basic physical mechanisms responsible for generating these fields are still not understood. ***An out-of-ecliptic heliophysics mission (such as Solar-C Plan A or SPI) that focuses on observations of the magnetic structure, dynamics and solar-cycle evolution of the polar regions of the Sun will provide data of critical importance for our understanding of the solar dynamo mechanism and the cyclic nature of solar activity.*** This task will be accomplished by launching instruments that can characterize processes occurring inside the solar convection zone and draw the connection to their manifestation on the solar surface and in the heliosphere. By pursuing a complementary and integrated approach between measurements addressing dynamo signatures in the solar convection zone and magnetism at the solar surface, a polar-viewing mission, such as Solar-

C/Plan A or SPI, can make great scientific progress in our understanding of the Sun and the fundamental origins of cosmic magnetism.

Science Challenges

The solar magnetic cycle is one of the greatest puzzles in astrophysics. Similar magnetic cycles have been observed in other stars; and the turbulent dynamo that establishes these cycles is a key to understanding the generation of magnetic fields throughout astrophysics. It has been known for centuries that the number and distribution of sunspots changes quasi-periodically with 11-year cycle, as manifested in the well-known solar "butterfly" diagram. Many recent solar dynamo models are based on the so-called "flux-transport" paradigm whereby the advection of magnetic flux by the axisymmetric flow in the meridional plane regulates the period and amplitude of the activity cycle. Most flux-transport models attribute the generation of mean poloidal field to the emergence and subsequent dispersal of photospheric active regions by the combination of poleward circulation and turbulent diffusion in the upper convection zone, known as the "Babcock-Leighton" mechanism. The reversal of the global poloidal field occurs as opposite-signed field from lower latitudes converges on the polar regions. However, how this processes couples to the dynamics and magnetic topology of the deeper convection zone is currently unknown. Recent Hinode observations reveal the polar field is highly structured, with magnetic field concentrations reaching 1kG, which is not captured by mean-field models. How the polar "magnetic landscape" changes with the solar cycle is one of the most critical questions for understanding the nature of the solar magnetism. The magnetic field evolution is intimately coupled to the flow field, including convection, differential rotation, and the mean meridional flow. In particular, the latitudinal bands of photospheric magnetic activity reflected in the solar butterfly diagram are closely linked to systematic variations in the differential rotation known as "torsional oscillations" (Fig. 1). A surprising result from helioseismology is that the depth and strength of these flows increase in the near-polar regions. Furthermore, models of surface flux transport indicate that the polar field strength is sensitive to variations in the amplitude and structure of the high-latitude meridional flow. However, reliable measurements of the magnetic and flow fields in the near-polar regions are currently lacking. ***There is a substantial gap in our knowledge of polar magnetism and dynamics, challenging our understanding of the basic mechanisms underlying solar magnetic activity, and inhibiting the development of physics-based forecasts of future solar activity.***

The outstanding issues confronting our current understanding of the solar dynamo may be summarized through several key scientific questions:

Q1) What physical processes dominate the generation of the global dipole magnetic field of the Sun, and where do they operate?

Q2) What is the dynamical origin of photospheric active regions? How does the generation of toroidal flux from poloidal flux and its subsequent emergence help regulate cyclic magnetic activity?

Q3) What are the structure and the evolution of the solar differential rotation, and the mean, large-scale meridional flow in the polar regions and deep convection zone, how are they maintained and how do they vary with the solar cycle?

Q4) Can we predict sunspot cycles and periods of high solar activity? Why might the activity cycle amplitude correlate with the polar field strength?

Progress on these scientific questions requires detailed observations of the solar polar regions, where data is currently scarce and where much of the subtle interplay between plasma flows and magnetic fields that gives rise to cyclic polarity reversals is thought to occur. ***Out-of-ecliptic observations of the Sun, for the first time, will provide opportunity for detailed investigations of the magnetic structure and dynamics of the polar regions.*** High-latitude photospheric observations will also provide an unprecedented vantage point for helioseismic imaging that can be used to probe flows and fields in the deep convection zone and tachocline where solar activity is ultimately thought to originate. With this in mind, we propose the following prime measurement targets for a future heliophysics mission, such as Solar-C Plan A:

T1) Differential rotation and meridional flow in the polar regions and the deep convection zone

T2) Photospheric magnetic flux distribution and evolution in the polar regions

T3) Dynamical coupling between magnetic fields and flows

T4) Structure and evolution of solar convection

T1. Differential Rotation and Meridional Flow in the Polar Regions and Convection Zone

In the past three decades, determination of the solar internal rotation profile and near-surface meridional circulation by means of global and local helioseismology has revolutionized solar dynamo theory. Rotational shear has long been an essential ingredient in all solar dynamo models, as the principle generation mechanism for the global-scale toroidal magnetic flux that ultimately emerges from the solar interior to form active regions. More recent flux-transport dynamo models have further identified flux advection by the mean meridional circulation as a key factor in the establishment of cyclic magnetic activity. In particular, the poleward advection of emergent toroidal flux from lower latitudes may account for the polarity reversals of the polar field and thus the cyclic reversals of the global dipole moment. Possible correlations between the high-latitude meridional flow speed and activity patterns such as cycle amplitude and duration are necessary to distinguish among various dynamo paradigms. Regardless of the dynamo mechanism, surface flux transport models indicate that the strength and distribution of the polar magnetic field (crucial for coupling to the heliosphere) is sensitive to the amplitude and structure of the high-latitude meridional circulation. ***Determination of the differential rotation and the meridional circulation in the polar regions will thus provide unprecedented insights into the dynamics of the convection zone, the operation of the solar dynamo, and the solar-terrestrial interaction, bringing the helioseismology revolution to its ultimate fruition.***

The unprecedented vantage point will also allow us to improve measurements of the large-scale and meridional flows in the deep convection zone by means of the ground-breaking techniques of stereoscopic helioseismology. Coordinated observations with another instrument such as HMI or GONG would provide the very long baselines needed to measure the relatively long-wavelength modes that sample the deep interior with local helioseismology. Such measurements are challenging but extremely important, since deep-seated flow signatures such as zonal jets may offer the most reliable probe of sub-surface magnetism.

T2: Magnetic Flux Distribution and Evolution in the Polar Regions

Synoptic observations of solar magnetic fields have shown evidence of transport of magnetic flux of the following polarity of bipolar active regions towards the poles. When a sufficient amount of the flux reaches the polar regions it changes the mean polarity of these regions, which

in turns reverses the global polarity of the mean poloidal field. ***However, how the process of the polarity reversal actually occurs in detail and how it couples to the deeper convection zone is unknown.*** What processes determine the distribution and subsequent evolution of surface and subsurface polar flux after the reversal? These most basic and intriguing questions of the solar magnetism problem will be addressed by a magnetograph viewing the pole from an inclination much greater than the 7 degrees provided by the ecliptic.

T3: Dynamical Coupling Between Fields and Flows

The most well-established signature of dynamical coupling between the zonal and meridional flows, magnetic fields, and thermal gradients is that of the solar torsional oscillations (see Fig. 1). These are alternating bands of east-west zonal flow that evolve in close correspondence with the solar activity cycle. Two components are evident from the data, a low-latitude branch that propagates equatorward in conjunction with bands of magnetic activity and a high-latitude branch that propagates poleward on a comparable time scale. The high-latitude branch is stronger and deeper and may arise from distinct dynamical influences. Correlated meridional flows and thermal signatures are known for the low-latitude branch but are currently undetectable for the high-latitude branch. An out-of-the ecliptic vantage point will provide an unprecedented opportunity to observe such correlations and in particular their phase relationships, providing important constraints to theoretical and numerical models of the solar convection zone and dynamo.

Variations of the meridional flow with the solar cycle induced by magnetic or thermal forcing also have important consequences for determining the strength and distribution of the polar field and the timing of polarity reversals. ***Coordinated observations of the meridional flow and the magnetic field in the polar regions will provide quantitative estimates of flux transport and will thus elucidate the physical mechanisms that underlie cyclic solar activity.***

T4: Structure and Evolution of Solar Convection

Magnetic flux transport by turbulent convection may contribute to the cyclic reversal of the large-scale poloidal field by working in concert with the meridional circulation. Not only does convection influence magnetism, but the converse is also true; magnetism can influence the structure of convection and investigating the nature of this nonlinear feedback provides valuable insight into both phenomena. In particular, the structure and evolution of supergranulation in polar coronal holes, where a net unipolar flux permeates the photosphere, may be significantly different than that at lower latitudes where the flux distribution exhibits mixed polarity. The Lorentz force tends to decrease convective length scales but magnetically-induced enhancements in radiative cooling may counteract this effect. Careful observations over extended time intervals (at least hundreds of days for reliable statistics) at high latitudes are needed to

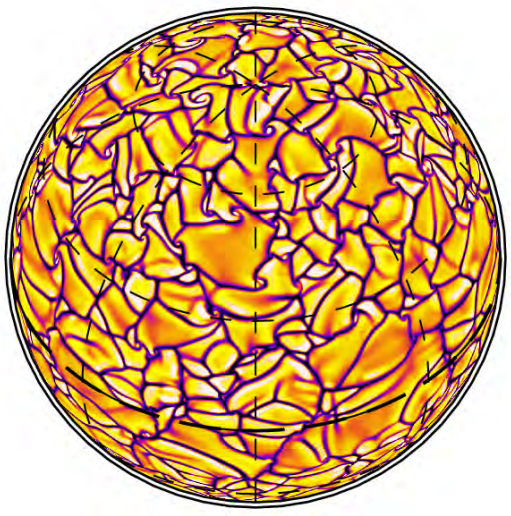


Fig 2: Polar view of solar convection simulation. Cyclonic vortices characterize the polar convection, while the equator has broader structures (Miesch et al, ApJ, 2008).

elucidate the subtle nonlinear feedbacks between solar convection, magnetism, and radiation.

Observational signatures of giant cells are notoriously difficult to glean from photospheric observations but a high-latitude vantage point will provide new opportunities. The maintenance of mean flows by global convective motions is expected to produce thermal gradients between the equator and pole that may be detectable by helioseismic inversions or photospheric irradiance measurements. Furthermore, theoretical and numerical models predict a change in morphology between global convective motions at high and low latitudes in rotating spherical shells (Fig 2). The transition between polar and equatorial convective modes occurs near the so-called tangent cylinder, a cylindrical surface aligned with the rotation axis and tangent to the base of the convection zone. This tangent cylinder intersects with the solar surface at about 45 degrees latitude. ***Possible changes in the subsurface flow fields inferred from local helioseismic inversions inside and outside the tangent cylinder may provide a valuable and previously unexploited observational signature of the elusive but extremely important giant cells. Hinode measurements have already revealed a systematic high-latitude alignment of supergranules that may reflect the underlying influence of giant cells.***

Helioseismic Observing Requirements

There are two main objectives of heliseismology observations from an inclined orbit:

1) ***High-latitude local helioseismology***: measure the local wavefield in the high-latitude region better than from the ecliptic place, for local helioseismic analyses of various flows in the near-surface (and if possible, deeper) layers, such as meridional flow.
2) ***Stereoscopic helioseismology***: in combination with observing facilities in the ecliptic plane, observe waves with large skip angles better, for helioseismic analyses of the deep interior, such as mid-latitude tachocline regions.

Helioseismology requires long continuous observing periods at high time cadence to resolve the various spatial and temporal scales associated with processes of generation and dissipation of solar magnetic fields. These scales range from supergranulation, which defines the magnetic network, is a primary source for magnetic field diffusion and has a typical spatial scale of 30 Mm and lifetime of about 1 day, to active regions occupying 50-100 Mm and evolving on the scale of a week, and to the global meridional and zonal flows occupying the whole convection zone 200 Mm deep and varying with the solar cycle on the time scale of a year.

The fundamental scales of solar dynamics and activity define three basic observing patterns *that are needed to achieve the primary objective above: relatively short (8 hour long) but high-resolution (512x512 pixels) time series to investigate the structure and dynamics on the supergranular and magnetic network scales; longer (7-14 days) and high-resolution (512x512 pixels) series to follow the evolution of large-scale magnetic flux regions, and long time period (36-72 day) low resolution (256x256 pixels) observations to investigate the global dynamics and circulation of the convection zone down to the tachocline by both local and global helioseismology techniques.*

All these observing programs require a 1-minute cadence for photospheric Dopplergrams. However, a strict temporal continuity requirement (>95% coverage) is required for 8-hour

chunks of these programs. Each 8-hour helioseismology observing run will provide a single snapshot of subsurface flows in the depth range of approximately 0-20 Mm. Sufficiently long sequences of the 8-hour uninterrupted runs are needed to study the dynamics of the flows. The longer observing programs may have gaps. However, the total coverage should be at least 70-80%, in particular for the global helioseismology 36-72-day runs, to prevent the aliasing problems in measurements of mode frequencies and frequency splitting. The precision of the frequency measurements is inversely proportional to the square root of the observing time. Currently, 72-day runs are used in the SOHO/MDI Structure program that is similar to the low-resolution (256x256 pixels) program described here. These runs provide inversion results for solar structure, asphericity and rotation through the whole convection zone, tachocline and outer radiative zone.

The current global and local helioseismology results for differential rotation, meridional flows and others subsurface flows are obtained only up to about 60 degrees latitude, a limit set by many factors but most crucially by projection and foreshortening effects. Successful measurements of the polar region dynamics therefore require observations from at least 30 degrees latitude. To ensure that the spacecraft stays above this latitude long enough for global helioseismology, the inclination must be about 40 degrees or more for a 1-AU Keplerian orbit.

Table 1. Helioseismology Observing Requirements

Measurement Objective(s)	Technique	Length of observation	Image size (pixels)	Image cadence
Differential rotation and torsional oscillations at high latitudes, sub-surface polar jets	Global Local	36-72 days 8 hrs	256x256 512x512	1 min 1 min
Meridional circulation, latitudinal & longitudinal structures, secondary cells, relationship to active longitudes, magnetic flux transport	Local	8 hrs	512x512	1 min
Supergranulation and large-scale convection patterns in polar regions (super-rotation, wave-like behavior, network, flux transport, relationship to coronal holes)	Local	8 hrs	512x512	1 min
Structure and dynamics of the high-latitude tachocline (oblateness, flows)	Local	36 days	256x256	1 min
Tomography of the deep interior:stereoscopic observations (with SDO or GONG)	Local	36 days	256x256	1 min

Conclusion

Polar observations with an out-of-ecliptic mission is one the most compelling challenges for the next decade in our quest for understanding the solar dynamo, the Sun's magnetic activity, and the origin of cosmic magnetism in general. The observations can be done with currently available technology and methods. The investigation of the dynamics and magnetism of the polar regions is a critical part of an integrated research strategy for understanding the mechanisms of solar activity and variability, and the influences on the Earth and the Solar System.

NASA's Balloon Program:
Providing World-Class Science, Technology Development, and Vital Training of the Next Generation of Space Physicists

R. M. Millan[1], D. M. Smith[2], M. P. McCarthy[3], R. H. Holzworth[3], J. G. Sample[4], E. A. Bering[5], and R. P. Lin[4]

Summary

High-altitude balloon experiments have a rich history in both Solar and Heliophysics. Balloons continue to offer a unique science platform and essential training for the next generation of scientists and engineers. This paper summarizes examples of balloon-based science accomplishments and future opportunities. A strong balloon program is vital to the future of NASA and to Space Physics research. We are advocating continued support of the Ultra Long Duration Balloon (ULDB) program, increased support and flight opportunities for small and mid-sized payloads and balloon flotillas, and appropriately scaled funding opportunities for the development of new, sophisticated payloads that produce world-class science.

1. Balloon-Based Science

Balloons continue to contribute compelling science across the field of Space Physics - from gamma ray solar flares to particle precipitation to measurements of large-scale magnetospheric electric fields.

In Solar Physics, balloons offer a low-cost method for carrying heavy payloads high above the atmosphere. Such missions have led to scientific discovery and are an ideal platform for developing and testing new spacecraft instrumentation. For example, in the 1980's, balloon-based X-ray instruments discovered hard X-ray microflares [Lin et al., 1984] and superhot flare plasmas [Lin et al., 1981]. Balloon missions (e.g. HEIDI, HIREGS) were also essential for the development of the RHESSI (Ramaty High Energy Solar Spectroscopic Imager) Small Explorer that has operated for nearly a decade. Optical telescopes have used balloon platforms to improve solar imaging and spectroscopy from Project Stratoscope in the 1950's to the SUNRISE payload today [Solanki et al., 2010].

In 1957, shortly before the discovery of the radiation belts, X-rays were detected by a balloon-borne particle detector launched near Minneapolis; the X-rays were believed to be bremsstrahlung produced by energetic electrons as they enter Earth's atmosphere and collide with neutrals [Winckler et al., 1958]. Since that time, balloons have made important contributions to our understanding of both auroral and radiation belt particle precipitation [Parks et al., 1993]. Electron microbursts (Figure 1) were discovered with balloons [Anderson and Milton, 1964] and later studied extensively at relativistic energies by the SAMPEX satellite [Blake et al., 1996]. We now know that microbursts are an important loss mechanism for the radiation belts [Lorentzen et al., 2001, O'Brien et al., 2003] and may be an ideal test case for studying non-linear wave-particle interactions [Cattell et al., 2008].

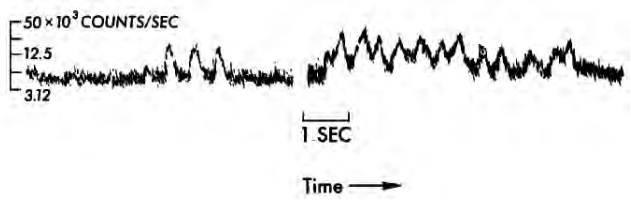


Figure 1: Electron microbursts were first observed by balloons [Anderson and Milton, 1964]

Balloons offer a unique platform for precipitation studies that is complementary to spacecraft measurements such as those made by SAMPEX. In particular, a balloon is

nearly stationary over the time of a typical electron precipitation event, thus resolving temporal and spatial ambiguity. These observations have revealed complex spatial and temporal structure (Figure 2) that is neither understood nor incorporated into current models. Balloons may be the only effective way to study these temporal variations, which may be caused by modulation of plasma wave growth, in turn modulating the precipitation rate.

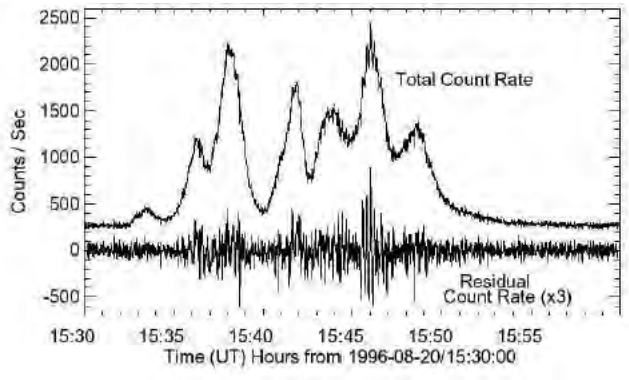


Figure 2: First observation of relativistic electron precipitation from a balloon showing modulation of precipitation at ULF timescales [Foat et al., 1998].

The BARREL (Balloon Array for RBSP Relativistic Electron Precipitation) project will fly 40 small (∼20 kg) balloon payloads during two Antarctic campaigns in 2013 and 2014 to provide a global view of electron precipitation and the ability to distinguish complex temporal and spatial variations.

Balloon-borne observations of substorm electric fields have contributed to our understanding of the substorm process for more than 30 years [Kelley et al., 1971, Mozer, 1971, Mozer and Lucht, 1974, Kokorowski et al., 2008]. The idea that processes in the high-latitude ionosphere are images of activity in the magnetosphere has been used to study the magnetosphere for many years [Baker et al., 2004]. One of the assumptions is that the conductivity of the magnetospheric plasma is so high that magnetic field lines are electrostatic equipotentials. This assumption has allowed use of ionospheric plasma flow and electric field observations to study large scale magnetospheric dynamics [Mozer and Manka, 1971, Mozer, 1973, Mozer and Gonzalez, 1973, Madsen et al., 1976, Foster and Doupnik, 1984, Reiff and Burch, 1985, Bering and Benbrook, 1987, Hairston and Heelis, 1990].

2. Training Opportunities

Suborbital investigations historically support the development of young scientists and provide the experience required for PIs of Explorer class missions. A majority of Principal Investigators on NASA Space Physics missions got their formative experiences this way. In particular, balloons and rockets provide hands-on experience with hardware, and an ideal mission duration that exposes students to the entire lifecycle of a mission, from the proposal process through the presentation of scientific results.



Figure 3: BARREL student involvement.

For example, the BARREL project has just advanced into Phase C and has conducted two test campaigns as part of the Phase B development. Students have been involved in every aspect including hardware design and fabrication, I&T, environmental testing, GSE software, thermal modeling, field operations, data analysis, and re-

views (Figure 3). So far, participation has included 1 high school student, 14 undergraduates, and 8 graduate students. Six students made presentations at the Preliminary Design Review. Two masters thesis have been completed and multiple Ph. D. dissertations are expected.

The balloon flight environment poses many of the same challenges as spaceflight, including a complex thermal environment, telemetry constraints, and solar power systems required for long duration missions. Thus, balloons provide an ideal training ground in addition to being an ideal test platform for new instrumentation. This training is necessary for future experimental scientists but also provides a unique opportunity to train personnel in systems engineering and project management. In their recently released report, the NRC Committee on NASA's Suborbital Research Capabilities made the recommendation, "To increase the number of space scientists, engineers, and system engineers with hands-on training, NASA should use the suborbital program elements as an integral part of on-the-job training and career development for engineers, experimental scientists, systems engineers, and project managers."

3. Future Directions

Heavy Lifting and Long Duration



Figure 4: SUNRISE balloon launch [Report of the NRC Committee on NASA's Suborbital Research Capabilities, 2010]

NASA currently supports large payloads launched primarily during northern and southern summers from a select few locations. They also continue to develop the capability for extended duration (∼100 day) flights through the Ultra Long Duration Balloon (ULDB) program. This capability remains critical for solar research; X-ray, gamma-ray, and neutron instruments are generally very heavy due to the stopping power required to stop high energy photons and the long observing window required to catch rare large gamma-ray flares. The next generation GRIPS (Gamma-Ray Imaging Polarimeter for Solar flares) instrument has been funded for development and is planned for LDB flights, but the ULDB capability would significantly increase the science return. Balloons can also carry large and powerful optical telescopes such as SUNRISE, which weighs more than 6 tons (Figure 4), high above the atmosphere where they make ultra high-resolution observations at the ∼50 km spatial scale (Figure 5), three times better than the state-of-the-art Hinode space mission [NRC Committee on NASA's Suborbital Research, 2010]. The SUNRISE measurements additionally extend to UV wavelengths unreachable from the ground.

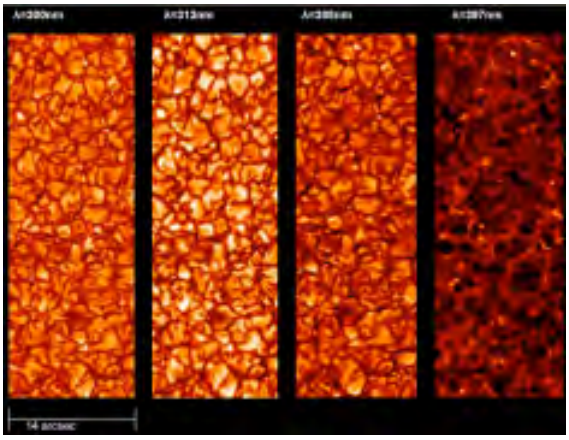


Figure 5 - SUNRISE High resolution image of magnetic solar region [Image: MPI for Solar System Research]

In 2009, NASA conducted a highly successful 54-day flight of a mid-sized superpressure balloon as part of the ULDB test plan (Figure 6). The BARREL team flew a prototype instrument system on this flight to conduct initial testing. Because of the very long observing times, we obtained interesting measurements of electron precipitation, even during a very magnetically quiet period. Balloon payloads can also be recovered, refurbished, and reflown at a very low cost.

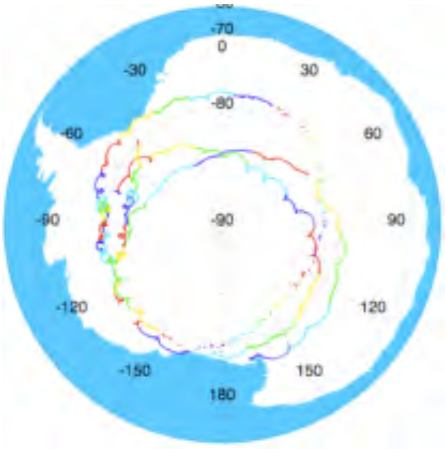


Figure 6: Flight trajectory over Antarctica of 54-day ULDB test flight in 2009 which carried BARREL prototype instruments.

Balloons remain the cheapest and most effective method for carrying huge (>1000 kg) payloads to the edge of space. NASA should continue development of the ULDB program, which promises observing times comparable to satellite missions.

Support for Small and Mid-sized Payloads

Magnetospheric science is a system science often requiring multi-point measurements and coordinated measurements from different platforms. The BARREL campaigns will take place during the RBSP (Radiation Belt Storm Probes) mission, providing conjunctions which will quantitatively reveal the physics of wave-particle interactions that cause precipitation. Since equatorial spacecraft cannot resolve the small atmospheric loss cone, low altitude measurements are necessary for studying precipitation.

Unlike the magnetospheric magnetic field, which penetrates to the ground for open field lines, the dynamics of the large scale electric field in the ionosphere and magnetosphere cannot be deduced from the ground. A BARREL-like flotilla capable of measuring vector electric fields would allow for an instantaneous determination of the large scale magnetospheric electric field. There is no other cheap way to get direct, real-time, mapped-in-situ, large scale magnetospheric electric fields with 1 to 10 second resolution, and global, instantaneous coverage. Charged particles cannot be energized without the particle seeing an electric field in its frame of reference, with a component parallel to its velocity. Yet the large-scale, remote electric field cannot be easily determined from satellite measurements (without averaging for minutes to hours). Balloons offer a very low cost method for obtaining this critical information on large scale electric fields.

MINIS was the predecessor to BARREL and consisted of a total of six small balloon payloads launched from the northern and southern hemispheres in 2005. Each of the southern payloads carried an X-ray detector, DC and searchcoil magnetometers, and electric field instruments [Kokorowski et al., 2006, Kokorowski et al., 2008]. MINIS was funded by the National Science Foundation and the balloons were hand-launched by our group because at that time, NASA did not support small balloon payloads at all. Since that time, the NASA balloon program has developed some support systems for small balloon payloads. This is an important step in supporting magnetospheric physics.

NASA should continue to support and develop control systems for small and mid-sized payloads, and should develop the capability for supporting balloon flotillas. Because small payloads do not carry the same safety or cost risk as larger payloads, NASA should appropriately relax its constraints on launch locations for small payloads.

Funding Opportunities

Balloon-based instruments are making world-class scientific observations, and have thus, not surprisingly, increased in complexity. Missions such as BARREL require multiple simultaneous flights. Although the individual payloads are relatively smaller and cheaper than the heavy solar experiments, the cost and complexity is found in the number of payloads. In both cases, balloon experiments are developing beyond the current funding envelope. LCAS is the main program supporting suborbital science missions at the ∼$350k/year level (not including launch costs). According to the 2010 NASA ROSES Announcement, $1.2M was expected to be available in FY2011 to support new investigations. The maximum grant duration is currently capped at 3 years which must include instrument development, launch, and data analysis. This places tight constraints on the development of new balloon-based experiments. The NRC Committee on NASA's Suborbital Research Capabilities found that current funding levels for science payloads is inadequate and should be increased to take advantage of the new, near space observatory-class capabilities.

BARREL takes advantage of a unique funding opportunity since it was selected as a LWS Geospace Mission of Opportunity. This opportunity allowed for development of a mission well beyond LCAS or NSF funding levels, but still very modest (<$10M). However, such opportunities are rare for suborbital missions. For BARREL, it was first necessary obtain an amendment of the AO for the proposal to even be allowable. Moreover, at the time of the proposal review, BARREL was subject to the same management and review requirements as a spacecraft mission. Creation of opportunities designed specifically for suborbital projects would avoid these unnecessary difficulties. BARREL, as it is currently being managed, could be used as a precedent for similar suborbital missions. Funding opportunities, however, still remain sparse.

We suggest these ways of enhancing the opportunities for low-cost suborbital missions:

1. Re-instate NASA's University Explorer mission class (UNEX) which could provide the funding necessary for both balloon and rocket experiments to achieve world-class science results. Funding of UNEX missions (upward of $10M) is at the level appropriate for the development and support of advanced balloon-based instrumentation and concepts.

2. Create opportunities for targeted suborbital and ground-based campaigns that directly support spacecraft missions. Prior to BARREL, there are several examples of successful suborbital and ground-based investigations that significantly contributed to spacecraft mission science. In 1992, sounding rockets were launched to study chemical releases from the CRRES (Combined Release and Radiation Effects Satellite) spacecraft. Ground-based measurements are commonly combined with spacecraft measurements to understand magnetospheric processes. In some cases, it has proven beneficial to install additional ground-based instruments in support of a specific spacecraft mission. This is exemplified most recently by

the THEMIS mission which included a 20-camera all-sky imaging array, and 21 magnetometers installed in the northern U.S. and Canada. Such opportunities are currently rare and NASA should look for ways to increase them because it is a very cost-effective method for enhancing the relatively small number of missions and for directly involving a larger part of the community, particularly students. An increase in these opportunities can be achieved for PI-led Explorer missions by encouraging and allowing for targeted supporting instrumentation at the proposal stage. For larger scale missions such as RBSP or MMS, a category of Missions of Opportunity specific to suborbital and ground-based instruments should be created.

Implementation of these or similar ideas would enhance balloon-based science, as well as rocket-based and ground-based science.

4. Conclusions

The NRC Committee on NASA's Suborbital Research Capabilities found that NASA's suborbital programs "play vital and necessary strategic roles in NASA's research, innovation, education, employee development, and spaceflight mission success, thus providing the foundation for achievement of agency goals." In a constrained funding environment, we must do everything possible to stretch available resources to advance our field. Balloons continue to produce important science at a nearly negligible cost compared with large spacecraft missions. Moreover, balloons, along with rocket and ground-based experiments, can enhance the large missions where most resources are spent. Balloons provide an ideal environmental test platform to bring new instruments to technology readiness level appropriate for space flight. Our strategy for Space Physics must use these low-cost programs to support NASA missions. A strong balloon program is vital to the future of NASA and to Space Physics research.

[1]Department of Physics and Astronomy, Dartmouth College
[2]Department of Physics and SCIPP, U. C. Santa Cruz
[3]Department of Earth and Space Sciences, University of Washington
[4]Space Sciences Laboratory and Physics Department, U. C. Berkeley
[5]Department of Physics, University of Houston

References

[Anderson and Milton, 1964] Anderson, K. A. and Milton, D. W. (1964). Balloon observations of X rays in the auroral zone. *J. Geophys. Res.*, 69:4457–4479.

[Baker et al., 2004] Baker, D. N., Kanekal, S. G., Li, X., Monk, S. P., Goldstein, J., and Burch, J. L. (2004). An extreme distortion of the Van Allen belt arising from the 'Hallowe'en' solar storm in 2003. *Nature*, 432:878–881.

[Bering and Benbrook, 1987] Bering, III, E. A. and Benbrook, J. R. (1987). Conjugate ionospheric electric field measurements. *Annales Geophysicae*, 5:485–502.

[Blake et al., 1996] Blake, J. B., Looper, M. D., Baker, D. N., Nakamura, R., Klecker, B., and Hovestadt, D. (1996). New high temporal and spatial resolution measurements by SAMPEX of the precipitation of relativistic electrons. *Adv. Space Res.*, 18(8):171–186.

[Cattell et al., 2008] Cattell, C., Wygant, J. R., Goetz, K., Kersten, K., Kellogg, P. J., von Rosenvinge, T., Bale, S. D., Roth, I., Temerin, M., Hudson, M. K., Mewaldt, R. A., Wiedenbeck, M., Maksimovic, M., Ergun, R., Acuna, M., , and Russell, C. T. (2008). Discovery of very large amplitude whistler-mode waves in earths radiation belts. *Geophys. Res. Lett.*, 35(L01105).

[committee on NASA's suborbital research capabilities, 2010] committee on NASA's suborbital research capabilities, N. R. C. (2010). *Revitalizing NASA's Suborbital Program: Advancing Science, Driving Innovation, and Developing a Workforce*. The National Academies Press, Washington, D. C.

[Foat et al., 1998] Foat, J. E., Lin, R. P., Smith, D. M., Fenrich, F., Millan, R., Roth, I., Lorentzen, K. R., McCarthy, M. P., Parks, G. K., and Treilhou, J. P. (1998). First detection of a terrestrial MeV X-ray burst. *Geophys. Res. Lett.*, 25(22):4109–4112.

[Foster and Doupnik, 1984] Foster, J. C. and Doupnik, J. R. (1984). Plasma convection in the vicinity of the dayside cleft. *J. Geophys. Res.*, 89:9107–9113.

[Hairston and Heelis, 1990] Hairston, M. R. and Heelis, R. A. (1990). Model of the high-latitude ionospheric convection pattern during southward interplanetary magnetic field using DE 2 data. *J. Geophys. Res.*, 95:2333–2343.

[Kelley et al., 1971] Kelley, M. C., Starr, J. A., and Mozer, F. S. (1971). Relationship between magnetospheric electric fields and the motion of auroral forms. *J. Geophys. Res.*, 76(22):5269–5277.

[Kokorowski et al., 2008] Kokorowski, M., III, E. B., Ruohoniemi, M., Sample, J., Holzworth, R., Bale, S., Blake, J., Collier, A., Hughes, A., Lay, E., Lin, R., McCarthy, M., Millan, R., Moraal, H., OBrien, T., Parks, G., Pulupa, M., Reddell, B., Smith, D., Stoker, P., and Woodger, L. (2008). Magnetospheric electric field variations caused by storm-time shock fronts. *Adv. Space. Res.*, 42:181–191.

[Kokorowski et al., 2006] Kokorowski, M., Sample, J. G., Holzworth, R. H., Bering, E. A., Bale, S. D., Blake, J. B., Collier, A. B., Hughes, A. R. W., Lay, E., Lin, R. P., McCarthy, M. P., Millan, R. M., Moraal, H., O'Brien, T. P., Parks, G. K., Pulupa, M., Reddell, B. D., Smith, D. M., Stoker, P. H., and Woodger, L. (2006). Rapid fluctuations of stratospheric electric field following a solar energetic particle event. *Geophys. Res. Lett.*, 33:20105–+.

[Lin et al., 1984] Lin, R. P., Schwartz, R. A., Kane, S. R., Pelling, R. M., and Hurley, K. C. (1984). Solar hard X-ray microflares. *Astrophysical Journal*, 283:421–425.

[Lin et al., 1981] Lin, R. P., Schwartz, R. A., Pelling, R. M., and Hurley, K. C. (1981). A new component of hard X-rays in solar flares. *Astrophys. J. Lett.*, 251:L109–L114.

[Lorentzen et al., 2001] Lorentzen, K. R., Looper, M. D., and Blake, J. B. (2001). Relativistic electron microbursts during the GEM storms. *Geophys. Res. Lett.*, 28:2573–2576.

[Madsen et al., 1976] Madsen, M. M., Iversen, I. B., and Dangelo, N. (1976). Measurements of high-latitude ionospheric electric fields by means of balloon-borne sensors. *J. Geophys. Res.*, 81(22):3821–3824.

[Mozer, 1971] Mozer, F. (1971). Origin and effects of electric fields during isolated magnetospheric substorms. *J. Geophys. Res.*, 76(31):7595–7608.

[Mozer, 1973] Mozer, F. (1973). Electric fields and plasma convection in the plasmasphere. *Reviews of Geophysics*, 11(3):755–765.

[Mozer and Gonzalez, 1973] Mozer, F. and Gonzalez, W. D. (1973). Response of polar cap convection to the interplanetary magnetic field. *J. Geophys. Res.*, 78(28):6784–6786.

[Mozer and Lucht, 1974] Mozer, F. and Lucht, P. (1974). The average auroral zone electric field. *J. Geophys. Res.*, 79:1001–1006.

[Mozer and Manka, 1971] Mozer, F. and Manka, R. H. (1971). Magnetospheric electric field properties deduced from simultaneous balloon flights. *J. Geophys. Res.*, 76(7):1697–1712.

[O'Brien et al., 2003] O'Brien, T. P., Lorentzen, J. R., Mann, I. R., Meredith, N. P., Blake, J. B., Fennel, J. F., Looper, M. D., Milling, D. K., and Anderson, R. R. (2003). Energization of relativistic electrons in the presence of ULF power and MeV microbursts: Evidence for dual ULF and VLF acceleration. *J. Geophys. Res.*, 108:doi:10.1029/2002JA009784.

[Parks et al., 1993] Parks, G. K., Freeman, T. J., McCarthy, M. P., and Werden, S. (1993). The discovery of auroral X-rays by balloon-borne detectors and their contributions to magnetospheric research. In *Auroral Plasma Dynamics*, Geophysical Monograph 80, pages 17–23, Washington, D. C. American Geophysical Union.

[Reiff and Burch, 1985] Reiff, P. H. and Burch, J. L. (1985). IMF B(y)-dependent plasma flow and Birkeland currents in the dayside magnetosphere. II - A global model for northward and southward IMF. *J. Geophys. Res.*, 90:1595–1609.

[Solanki et al., 2010] Solanki, S. K., Barthol, P., Danilovic, S., Feller, A., Gandorfer, A., Hirzberger, J., Riethmüller, T. L., Schüssler, M., Bonet, J. A., Martínez Pillet, V., del Toro Iniesta, J. C., Domingo, V., Palacios, J., Knölker, M., Bello González, N., Berkefeld, T., Franz, M., Schmidt, W., and Title, A. M. (2010). SUNRISE: Instrument, Mission, Data, and First Results. *Astrophys. J. Lett.*, 723:L127–L133.

[Winckler et al., 1958] Winckler, J. R., Person, L., Arnoldy, R., and Hoffman, R. (1958). X-rays from visible aurora at minneapolis. *Phys. Rev.*, 110:1221.

NASA's Explorer Program as a Vital Element to Further Heliophysics Research

Eberhard Möbius

Space Science Center and Department of Physics, University of New Hampshire, Durham, NH

Eric R. Christian, Goddard Space Flight Center, Greenbelt, MD

Christina M. S., Cohen, Richard A. Leske, California Institute of Technology, Pasadena, CA

Stephen A. Fuselier, Lockheed Martin Advanced Technology Laboratory, Palo Alto, CA

Eberhard Gruen, Max-Planck-Institut für Kernphysik, Heidelberg, Germany, and Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO

Michael Gruntman, University of Southern California, Los Angeles, CA

Lynn M., Kistler, Harald Kucharek, Nathan A. Schwadron, Roy B. Torbert, Space Science Center and Department of Physics, University of New Hampshire, Durham, NH

Janet G. Luhman, Space Sciences Laboratory, University of California, Berkeley, CA

David J. McComas, Mihir Desai, Southwest Research Institute, San Antonio, TX

Christopher T. Russell, Robert Strangeway, University of California, Los Angeles, CA

Mark A. Wiedenbeck, Jet Propulsion Laboratory, Pasadena, CA

Introduction

This White Paper will make the case for an enhanced Explorer program as a vital program element for Heliophysics. The Explorer program has been one of the greatest success stories of NASA since its very beginning. Explorers, which include stand-alone missions of various sizes as well as the Explorer-run Missions of Opportunity (MoOs), are PI-led missions that have returned outstanding discoveries. Missions, such as COBE, AMPTE, SAMPEX, FAST, ACE, TRACE, WMAP, RHESSI, THEMIS, and most recently IBEX, have produced breakthrough science results. This program is vital to an extraordinary range of science in both the Heliophysics and Astrophysics science programs. It is the only competed mission line that has consistently offered relatively frequent opportunities to carry out small missions with well-focused science goals. These missions are developed and launched in a short (approximately three to four-year) timeframe. Several of the missions have led to breakthrough results in Heliophysics, such as the Interstellar Boundary Explorer (IBEX) with its first all-sky maps of the heliospheric boundary and the discovery of the "Ribbon" as a marker of the surrounding interstellar magnetic field orientation.

For the health of the field, it is important to maintain a steady stream of small missions between the increasingly sparse large mission opportunities. Also, these PI-led missions have historically stayed within budget and on schedule. It is also worth noting that in the most recent Senior Review the TWINS mission, one of the Explorer Missions of

Opportunity, was rated #1 in science per dollar among all 13 heliospheric missions under consideration. Therefore, we call for an increase in the Explorer Budget so that on average one launch per year is achieved again, with the typical accessibility of 50% of the missions being Heliophysics missions.

This white paper is applicable to all three Decadal Survey study panels.

Discussion

Explorer Missions have successfully spearheaded science topics that range from the Earth's atmosphere, auroral regions, magnetospheric substorms via the Sun and the boundary of our solar system with the interstellar medium to most distant luminous objects in the universe at an overall cost that is well below those of larger scale missions. ISEE has provided deep insight into particle acceleration and transport through the magnetosphere and has enabled the first ever encounter with a comet; AMPTE has pioneered active experiments in the solar wind and advanced composition observations throughout the magnetosphere; FAST has re-written the understanding of auroral phenomena; ACE has taken solar wind, interstellar, solar energetic particle, and cosmic ray composition to the next level and in the process changed several paradigms about solar and galactic cosmic rays; TRACE has provided detailed insight into chromospheric and coronal structures on the Sun; RHESSI has provided unprecedented solar flare diagnostics; and IBEX is changing the paradigms of the heliosphere-interstellar medium interaction. As is evidenced by the frequency of these missions, the Explorer program was intended to provide frequent and rapid access to space in order to be able to react to new promising and important scientific ideas. However, over the past decade the rate at which Explorers are initiated has substantially diminished.

Simultaneously over the last few decades, increased launch and mission costs have reduced the launch rate for all SMD missions, in particular, the larger mission lines. Already now, there are only a few places in the country left that are capable of supporting a flight instrument development and implementation. This is particularly true for space science groups at universities, where other large-scale developments are not present as, for example, in large research centers. The remaining hardware centers at universities are in constant danger of losing valuable engineers and technicians from their just barely critical skill mix, since they don't have any cushion to balance a "bursty" funding profile, which can become particularly uneven through very few large hardware programs. Having access to a reasonable rate of Explorer opportunities would be an important element to smooth the potential funding flow.

The Explorer Program has not only provided the opportunity to take up important focused investigations on a reasonably paced schedule, it has also been a determining factor in many scientific careers in Heliophysics, and has provided graduate and undergraduate students at participating universities with an invaluable training ground. The reason is that participation in Explorer programs has been much more in line with the mission and capabilities of university laboratories than the ever-increasing requirements for large missions. Therefore, increased access to Explorer opportunities will also provide a valuable training ground for NASA's technical base, with its ability to attract and train young talent.

Implementation and Cost

Currently, the Explorer line contains a mix of Small and Mid-Size Explorers and a separate SALMON (Stand-Alone Mission of Opportunity) component that provides the means to join forces with other agencies and/or other nations' space programs. Therefore, this mix of sizes and opportunities should be maintained. However, the overall program should be restored to the effort level of the late 1990's through the early 2000's. In 2001, the Explorer budget was approximately $140M. The FY11 budget request for Explorers is only $120M, which is still a large improvement in the funding over the last four or five years. However, to restore the Explorer program to its earlier level would require a budget of more than $170M, after adjusting for inflation.

Conclusion

Heliophysics has derived a number of its successes and breakthroughs in the past from successful Explorer missions. To remain innovative and maintain a healthy program with a steady stream of observations, punctuated by a high frequency of new opportunities, the field needs a mix of large missions and small PI-led frequent mission opportunities. This mix is also needed to maintain a diverse infrastructure in the field that includes viable university centers and to train the future generation of engineers and scientists. Such a balanced approach can be achieved by a targeted increase in the Explorer funding so that on average a launch rate of one mission per year is reached again. The required funding increase is of the order of $50M per year over last year's funding level.

Heliospheric Science at Low Radio Frequencies

Divya Oberoi[1], Leonid Benkevitch[1], Mario M. Bisi[2], Andrew Breen[2], Roger J. Cappallo[1], John Clover[3], Bernard V. Jackson[3], Elizabeth Jensen[4], Colin J. Lonsdale[1], Lynn D. Matthews[1], Christopher Russell[5], Alan R. Whitney[1], Angelos Vourlidas[6] and the MWA Project

ABSTRACT

Passage through the magnetized solar wind plasma leaves an imprint on electromagnetic radiation traversing it. These propagation effects, most easily studied at low radio frequencies, provide sensitive probes of the heliosphere. While known for many decades, the use of some of these tools to explore and understand the heliosphere had remained somewhat limited, primarily due to technological limitations. With the enormous progress in digital signal processing and computational capacity, we can now expect the new generation low radio frequency instruments, like the Murchison Widefield Array (MWA), to make effective use of these tools. This white paper reviews the science applications of a few different propagation effects - interplanetary scintillations, Faraday rotation and dispersion measure, identifies their key contributions, and the potential for detailed characterization of the heliosphere. We also identify crucial areas for future funding support to maximize the heliospheric science returns from the investments in MWA and its successors.

The heliosphere has traditionally been poorly observed – A vast array of instrumentation observes the solar disc across an impressively wide band of the electromagnetic spectrum. A large and capable fleet of satellites provide *in-situ* measurements of the solar wind plasma, most of them in the Earth's orbit and the ecliptic plane. In the vast inner heliosphere, however, there are far fewer means available to investigate the tenuous solar wind and this region has traditionally been poorly investigated. Yet it is the evolution and dynamics in this region which plays a dominant role in space-weather and needs to be understood in order to relate the observations close to the Sun to the *in-situ* measurements close to 1 AU.

The solar wind is more complicated that previously thought – Comparatively recently, imaging of Thomson scattered white light, a technique long used for coronal imaging, has

[1]MIT Haystack Observatory, Westford, MA USA

[2]Aberystwyth University, United Kingdom

[3]University of California, San Diego

[4]Planetary Science Institute / ACS Consulting

[5]Institute of Geophysics and Planetary Physics, University of California, Los Angeles

[6]Naval Research Laboratory

been extended to cover the inner heliosphere and to provide information about large scale solar wind density and velocity. Pioneered by the Solar Mass Ejection Imager (SMEI) on-board the *Coriolis* satellite (Eyles et al. 2003) and vastly improved by the next generation Heliospheric Imagers (HIs) on board the STEREO satellites (Harrison et al. 2009), these instruments now provide us detailed views of the solar wind in the inner heliosphere. One unmistakable conclusion from these observations is that the solar wind is significantly more structured and dynamic than previously thought.

Radio observations are a powerful tools for studying the heliosphere – Due to their wavelength dependence, the propagation effects imposed by the passage of electromagnetic radiation through the magnetized solar wind plasma are large enough at low radio frequencies ($\lesssim$300 MHz, in the present context) to serve a diagnostics for the solar wind plasma. These propagation effects provide a set of powerful and complementary tools to investigate the heliosphere. The propagation effects of particular interest in the current context are interplanetary scintillations (IPS), Faraday rotation (FR) and dispersion measure (DM) observations.

Murchison Widefield Array: A *Radio* Heliospheric Imager – The Murchison Widefield Array (MWA) is a new generation radio inteferometer currently under construction in the radio quiet Western Australian outback (Lonsdale et al. 2009). The MWA design exploits the advances in digital signal processing and computational capacity to rise to the challenge of low radio frequency imaging. As emphasized in the name, the wide fields-of-view ($\sim 10^3 \mathrm{deg}^2$ at 200 MHz) of MWA provide simultaneous access to a large part of the heliosphere. MWA design is the most aggressive implementation of a *large-N* design yet, where N refers to the number of elements in the intereferometer array. With its 512 elements, it will provide images with unprecedented imaging fidelity and dynamic range. With a point source sensitivity of 20 mJy (1 s integration with 32 MHz bandwidth at 200 MHz), the MWA will be a sensitive probe of propagation effects. With no moving parts and electronic pointing agility, the MWA will quickly be able to scan the entire sky. The MWA frequency range of 80–300 MHz spectral range is a good choice based on criteria ranging from magnitude of the propagation effects to the availability of suitable background sources. All these features make MWA exceptionally well suited for heliospheric science, and solar, heliospheric and ionospheric science are amongst its key science objectives. Other advantages of MWA are highlighted at appropriate places in the following text.

Interplanetary Scintillation: Information content – IPS is the radio analogue of optical twinkling of stars. It was amongst the earliest techniques used to investigate the solar wind (e.g. Hewish et al. 1964) and has lead to many interesting revelations about the solar wind. The discovery of high velocity polar solar wind well before the Ulysses polar pass (Coles et al. 1980) is a good example. In addition to the *macro* properties like velocity, the primary IPS observable, the power spectrum of intensity fluctuations, is also sensitive to *micro* properties like turbulence – the level and spectral slope of electron density fluctuations. Turbulence is likely to be an important mechanism for dissipation of energy as Coronal Mass Ejections (CMEs) travel through the heliosphere and hence plays an important role in CME dynamics, but is inaccessible to HI measurements.

A common property of all propagation effects is that the observable is an integral along the entire ray path through the medium. Depending upon the nature of available data and the sophistication of the analysis, the weighted averaging along the line-of-sight (LoS), can be its greatest strength or its primary weakness. In principal, a large number of carefully arranged IPS observations providing a sufficiently dense sampling of the heliosphere can be inverted to provide a global 3D/4D distribution of the solar wind parameters, a feat impossible to match for any local sampling technique.

Interplanetary Scintillation: Current status and the challenge – In practise, the tomographic reconstruction just mentioned only seems to work well either when the solar wind structure is comparatively simple or when it is dominated by a few strong features. This, however, results not from an intrinsic limitation of the technique but due to the inability of the current instruments to sample the heliospheric adequately to be able to reconstruct it. The challenge therefore is to be able to achieve a sampling of the heliosphere dense enough to allow a reliable and unique tomography-like inversion.

Much of the IPS tomographic work has been done using the Solar Terrestrial Laboratory IPS Array in Japan, which is currently undergoing an upgrade to increase its sensitivty so that the number of sources it can observe can increase from ∼40 to few 100 per day. With a physical collecting area of 15,900 m^2, the Ooty Radio Telescope (ORT) in India is currently the most sensitive IPS instrument. Impressive progress has been made in using IPS data along with that from other available resources to piece together a more comprehensive understanding of individual large events (e.g. Bisi et al. 2010b; Manoharan 2010) as well as trying to build global models of the heliosphere (e.g. Bisi et al. 2010a). There is an effort underway in the worldwide IPS community to work towards a common data format and analysis methods.

The advantages of next generation of radio instruments – In the present context, the ability to easily form multiple beams allows modern digital instrumentation to meet the challenge of achieving a much denser sampling of the heliosphere by having multiple simultaneous beams. The MWA will have 32 simulataneous beams, each using the full collecting area of the instrument, which can independently be pointed anywhere in its wide field-of-view. The ORT is currently being upgraded to digital signal path and backend which will provide ∼100 simultaneous beams and a $\sim 40^\circ \times 2^\circ$ field-of-view.

These instruments will increase the heliospheric sampling density by ∼2 orders of magnitude. Given that we are able to do reasonable reconstructions during favourable, less complicated instances, lends credence to the argument that the new generation datasets will lead to a significant improvement in quality and robustness of heliospheric reconstruction.

Faraday Rotation: Information content – Due to the dependence of geo-effectiveness on CMEs on their magnetic field orientations, the determination of the magnetic fields of CMEs has long been regarded as the "missing link" of space weather prediction. Radio observations of Farady rotation (FR) provide the *only* known remote sensing technique to estimate CME magnetic field

strength and orientation. The measurement requires monitoring the plane of linear polarization of radiation from a background polarized source due to passage through heliospheric magnetized plasma. If such measurements are available along a large number of lines of sight distributed over a large part of the sky, such a dataset will provide enough information to constraint the CME 3D magnetic field topology sufficiently and estimate its geo-effectiveness (Liu et al. 2007; Jensen et al. 2010).

Faraday Rotation: Current status and challenge – Heliospheric FR measurements were first done ∼40 years ago using satellite beacons on Pioneer I and II and have since been done with many satellite missions, most recently with the Cassini spacecraft (Jensen et al. 2005). Such measurements have never yet been done using polarized astronomical sources. A few independent reasons conspire to make this a difficult measurement – the fractional polarization at low radio frequencies is intrinsically low; the polarization performance of low frequency interferometers is difficult to accurately characterize; and the ionospheric FR is expected to be about an order of magnitude larger than the heliospheric signature.

The expectations from MWA – Observations from Westerbork telescope suggest that the sky density of polarized sources with sufficient flux density (20 mJy, typical) to be useful for these observations can be expected to be of order 1 per few deg squared (Haverkorn et al. 2003). In addition the diffuse galactic background emission is also known to have complicated polarization structure, which can suitable instances serve as polarized background sources (Haverkorn et al. 2004). The primary science driver for MWA, the Epoch of Reinonization (EoR) studies require measurements of a very weak (few mK) signal in the presence of enormous foreground emissions (hundreds of K). The MWA design has hence been optimized for low systematics and accurate calibration. We believe that we stand a good chance of being able to routinely make these measurements over the wide MWA field-of-view. If successful these measurements will routinely allow us to estimate CME magnetic field orientations and predict their geo-effectifveness.

Dispersion Measure – While used routinely in pulsar studies to characterize the interstellar medium, this technique is not yet used for studying the solar wind. The main reason behind this has been that DM expected due to the heliospheric plasma ($\sim 10^{-2}$–10^{-4} pc cm^{-3} in the region 1°–10° from the Sun) has been smaller than the typical DM measurement uncertainties. A technique recently pioneeded at the Giant Meterwave Radio Telescope (GMRT) has brought these uncertainties down to 10^{-4} pc cm^{-3} (Ahuja et al. 2005). This development and the fact that there are a few pulsars which pass through close to the Sun, make this a potentially interesting tool to get "clean" measurements of heliospheric electron density. Pulsar often have a significant linearly polarized flux, the angle of which varies over the phase of the pulse. Pulsar studies are hence a useful tool for polarization calibration as well as FR and DM studies.

Enabling Heliospheric Science – As mentioned earlier, the design, technology development and construction of the MWA is already being resourced. A companion white paper presents the case for low radio frequency solar imaging science (Oberoi et al. 2010a). Much of the support

specific to heliospheric science however still needs to be resourced. This support will leverage the significant US and international investment in MWA and maximize the heliospheric science returns.

We list below the major tasks which will need to be resourced to enable heliopsheric science. The tasks common for both solar and heliospheric science, e.g. operations costs, data capture and management support and firmware augmentations, are listed only in the companion white paper.

- **Haystack Interplanetary Scintillation Software System** – We have developed the software framework for archiving and analysing the MWA IPS data. This framwork is general enough to accommodate data from any IPS instrument and provides a common interface for all IPS data(Oberoi & Benkevitch 2010). The key aspects of the data archiving part of the framework have been tested in a sufficiently mature implementation. The data analysis functionality currently available is however limited and needs to be developed. This is a significant task and much of the complexity for a 3D/4D tomographic reconstruction will reside here.

- **IPS Survey** – MWA (Lat -26°) will be the first IPS instrument in the Southern hemisphere. The sky visible to MWA will have a significant overlap with that visible to ORT (Lat +11°), and we expect to have access to the catalogue of IPS sources generated at ORT over the last few decades. We will however need to survey a large fraction of the Southern sky to find suitable IPS sources. This will be a significant effort spanning, at the minimum, about an year, as one needs to observe the sources over a large range of elongations (angular distance from the Sun) to determine their suitability for IPS.

- **FR Survey** – The Southern sky has never been surveyed for polarized sources at low radio frequencies. In order to be efficient at heliospheric FR monitoring, we need to build a database of source of polarized emission in the Southern sky. Some of the aspects of this survey task are common with the need to build a global sky model to meet the calibration needs.

- **Ionospheric calibration** – Separating the heliospheric FR signature from the ionospheric, requires knowledge of *absolute* ionospheric total electron content (TEC) along the LoS to the source. While interferometric observations can determine *differential* TEC to few mTEC Units (1 TEC Unit = 10^{16} el m^{-3}), they are not sensitive to absolute TEC values. The best estimates for absolute TEC values come from dual frequency GPS receivers. However GPS receiver bias calibration needs to improved by a factor of few before the error on ionospheric FR can be reduced to desirable levels. This project also shares its needs with the standard MWA polarization calibration.

- **Data analysis effort** – Sensitive measurements requiring developing a detailed understanding of the instrumental characteristics and calibration, which comes only from a sustained effort working with the data. We will need to continue this effort and ramp it up as we go forward.

- **Building a more complete picture** – The MWA will provide heliospheric observations sensitive to different aspects of the heliosphere – IPS to velocity and turbulence charateristics; FR to electron density and magnetic field; and DM to electron density. In addition HI observations are sensitive to electron density and velocity. Though it will offer its own challenges, the simultaneous use of all these dataset will enable a far more detailed investigation of the heliosphere than has been possible before. Building a collaboration to work towards these goals will be necessary to allow extraction of maximum benefit/information from these combined datasets.

Technical Risk – The status of the MWA hardware development and the technical challenges associated with building and operating an unprecendentedly large interferometer have been discussed in the companion white paper (Oberoi et al. 2010a)

The IPS observations are amongst the simplest observations to be done by the MWA. Once the hardware array beamformer, a part of the MWA correlator system, is in place there is little technical risk associated with them. The FR observations are more challenging and will likely require a period of learning, where we optimize our calibration schemes and build a better understanding of the low radio frequency and its polarized component.

Early results from 32T – A 32 element engineering protytpe (32T) currently exists on site. In addition to serving its primary role of providing test bed for end-to-end engineering validation of hardware and software, it also provides an early science opportunity. The 32T does not have the sensitivity to do any of the observations proposed here or test the efficacy of the proposed calibration scheme, it does, however, provide a sensitive test of hardware performance. The data from 32T have already lead to some remarkable solar results (Oberoi et al. 2010b). The snapshot images from 32T have a dynamic range an order of magnitude better than those from the current facilities and show non-thermal emission features much weaker than ever seen before.

Cost and leveraging of resources – The support of this effort will leverage an international investment in the 20–25M range (Oberoi et al. 2010a). We anticipate submitting proposals to the Atmospheric and Geospace Sciences division of NSF seeking support for the aspects mentioned here at appropriate times in the future.

Astro-2010 endorsement of HERA – As discussed in a little more detail in the companion white paper (Oberoi et al. 2010a), the low radio frequency interferometry received a strong endorsement from the recently concluded Astronomy and Astrophysics Decadal Review (Astro-2010).

REFERENCES

Ahuja, A. L., Gupta, Y., Mitra, D., & Kembhavi, A. K. 2005, MNRAS, 357, 1013

Bisi, M. M., Jackson, B. V., Breen, A. R., Dorrian, G. D., Fallows, R. A., Clover, J. M., & Hick, P. P. 2010a, Sol. Phys., 265, 233

Bisi, M. M., et al. 2010b, Sol. Phys., 265, 49

Coles, W. A., et al. 1980, Nature, 286, 239

Eyles, C. J., et al. 2003, Sol. Phys., 217, 319

Harrison, R. A., et al. 2009, Sol. Phys., 256, 219

Haverkorn, M., Katgert, P., & de Bruyn, A. G. 2003, A&A, 403, 1031

—. 2004, A&A, 427, 549

Hewish, A., Scott, P. F., & Wills, D. 1964, Nature, 203, 1214

Jensen, E. A., Bird, M. K., Asmar, S. W., Iess, L., Anderson, J. D., & Russell, C. T. 2005, Advances in Space Research, 36, 1587

Jensen, E. A., Hick, P. P., Bisi, M. M., Jackson, B. V., Clover, J., & Mulligan, T. 2010, Sol. Phys., 265, 31

Liu, Y., Manchester, IV, W. B., Kasper, J. C., Richardson, J. D., & Belcher, J. W. 2007, ApJ, 665, 1439

Lonsdale, C. J., et al. 2009, IEEE Proceedings, 97, 1497

Manoharan, P. K. 2010, Sol. Phys., 265, 137

Oberoi, D., & Benkevitch, L. 2010, Sol. Phys., 265, 293

Oberoi, D., Benkevitch, L., Cappallo, R. J., Lonsdale, R. J., Matthews, L. D., Whitney, A. R., & the MWA Collaboration. 2010a, White paper submitted to Decadal Strategy for Solar and Space Physics

Oberoi, D., et al. 2010b, To be submitted to Astrophysical Journal Letters, currently under MWA Collaboration review. In the interim available at http://www.mwatelescope.org/tmp

This preprint was prepared with the AAS LaTeX macros v5.2.

High-Fidelity Coronal Imaging at Low Radio Frequencies

Divya Oberoi[1], Leonid Benkevitch[1], Roger J. Cappallo[1], Colin J. Lonsdale[1], Lynn D. Matthews[1], Alan R. Whitney[1], Angelos Vourlidas[2] and the MWA Project

ABSTRACT

Emission from the Sun at low radio frequencies (ν <300 MHz) provides a uniquely powerful diagnostic of the dynamics and physical conditions of the solar corona at heights of $R \gtrsim 0.1\ R_\odot$. Much of this region is largely inaccessible to observations at X-ray and EUV. Spectroscopic imaging of the Sun at low radio frequencies is poised to make key contributions to solar physics in the coming decade, as a new generation of low frequency radio interferometers comes on-line. The Murchison Widefield Array (MWA), currently under construction in the Australian outback, is one such instrument. The MWA will enable high time resolution, high spectral resolution solar imaging with unprecedented fidelity and dynamic range over an observing band of 80–300 MHz. This white paper reviews key science applications of such data, including studies of coronal magnetic field topology, magnetic reconnection, and possible contributions to coronal heating. Crucial areas for future funding support and technology development needed in order to maximize the potential of the MWA as a solar instrument are also addressed.

Origin of low-radio frequencies – Emission at low radio frequencies ($\lesssim$ 300 MHz) arises in the solar coronal regions at heights $\gtrsim$0.1 $R_\odot$ and has long been known to be a sensitive probe of the solar corona (e.g. Pick & Vilmer 2008; McLean, D. J. & Labrum, N. R. 1985, and references therein). The background solar emission in this frequency range is dominated by thermal bremsstrahlung from the hot, optically thin coronal plasma. The non-thermal emission, varying on short spectral and temporal scales, is dominated by plasma emission, a resonant emission mechanism in which electrostatic Langmuir waves are converted to electromagnetic radiation at the local plasma frequency and its harmonic. The middle and upper corona is especially interesting, as this is where the solar wind originates and space-weather phenomena are rooted. Most of this region is also difficult to probe by other means; X-rays and EUV tend to probe denser, hotter plasma, which is usually closer to the solar surface or confined to small volumes higher in the corona.

Information content of solar low-radio frequency observations – Large gradients in electron density and temperature in the coronal region lead to an even larger gradient in the absorption coefficient, κ. The changes in κ translate into the optical depth, τ, being a strong function of the ray path, which in turn is determined by refraction due to large scale density gradients. Brightness temperature, T_B, the primary radio observable, is an integral along the

[1]MIT Haystack Observatory, Westford, MA USA

[2]Naval Research Laboratory

entire ray-path through the corona, and is sensitive to practically all coronal parameters of interest - electron density and temperature distributions and magnetic field (line-of-sight component).

Role of "Quiet" Sun low-radio frequency observations – Due to refractive bending, the rays at different frequencies pass through different regions of the corona, making it possible to put together a consistent 3D picture of the corona using wideband low radio frequency observations in a manner similar to peeling the layers of an onion. Low radio frequency observations are orders of magnitude more sensitive to non-thermal electron heating events in the corona than any other available means (Ramesh et al. 2010a). Our recent observations with the MWA Prototype push this sensitivity advantage another order of magnitude (Oberoi et al. 2010b). A proper accounting of these contributions is essential to making progress on balancing the coronal heating budget. Radio observations of circular polarization provide the only means to quantify the line-of-sight magnetic field component at these coronal heights. Though they do not yet make the ranks of routine coronal diagnostics, such observations are begining to appear (e.g. Ramesh et al. 2010b). Understanding the background corona is a prerequsite to understanding the active corona.

Role of low-radio frequency observations for the "Active" Sun – A large number of Coronal Mass Ejections (CMEs) are seen in coronagraphs, but their radio counterparts have remained more elusive. The advantages of radio over optical and soft X-ray observations have been apprecited for some time now, and the metric and decimetric bands are widely regarded as the most promising part of the spectrum for CME studies (Bastian & Gary 1997). With only a handful of successful detections reported in the literature (e.g. Bastian et al. 2001; Magdalenić et al. 2008), routine direct detection of CME radio emission has remained largely beyond the reach of the current instrumentation. In some of these cases where the radio burst emission has actually been imaged, it is difficult to interpret the observations in the framework of the accepted models for CMEs and type II burst emission.

Imaging studies of type II and III bursts will help us verify the validity of theoretical models and emission mechanisms, and help improve them. Having broadband imaging information on type II and III solar bursts which drift through the spectrum, and the possibility of imaging the emission at both the fundamental and the harmonic simultaneously will help us improve our understanding of these pheomenon. Imaging studies of type III solar bursts will provide information about the magnetic field topology, by tracing the motion of the burst location (Pick et al. 2009). Due to propagation effect like scattering and ducting, these observations can also provide information about small scale properties of the coronal medium. Being sensitive to low level mangnetic reconnection activity (Oberoi et al. 2010b), these observations will help to explore the possibility of using activity seen in radio for identifying precursors of significant events (Jackson & Sheridan 1979). Low radio frequency radio observations are hence destined to play an important role in developing our understanding of solar and coronal physics.

The challenge and the current status – In spite of a large number of compelling reasons in its favour, much of the promise of low radio frequency solar studies is yet to be realised.

This is because high-fidelity coronal radio imaging at low frequencies is very challenging. Coronal emission features are complex and dynamic, evolving rapidly in space, time and frequency. The limited instantaneous spatial and spectral coverage provided by present day interferometers has been inadequate to simultaneously resolve transient solar phenomena in all of these domains. At low radio frequencies, the only operational radioheliographs (dedicated solar radio interferometers) are the Gauribidanur Radioheliograph (40-150 MHz) in India and the Nançay Radioheliograph (5 frequencies in the range 150-450 MHz) in France. They both provide non-contiguous spectral coverage at a few spot frequencies with a spectral resolution of order MHz, and their best dynamic range performance is of order few hundred.

Murchison Widefield Array: a capable new generation solar radio imager – The demands from a capable solar imaging radio telescope have remained well beyond the reach of available technological and financial resources, until recently. Leveraging the massive advances in digital signal processing and computational capacity, a new generation of low frequency intereferometers is expected to become available in the next few years. One of these, the Murchison Widefield Array (MWA) (Lonsdale et al. 2009), has solar, heliospheric and ionopheric science amongst its key science objectives and is exceptionally well suited for the needs of low radio frequency solar imaging. MWA specifications are tabulated in Table 1 below. The key relevant features of the MWA design for solar imaging are:

1. **Unprecedented imaging fidelity and dynamic range** – The large number of elements (512) spread over a compact footprint (3 km). The resulting 131,072 physical baselines will lead to an instantaneous monochromatic *uv* sampling approaching Nyquist (for solar imaging), providing full polarimetric radio imaging capability of unprecedented fidelity and flexibility.

2. **Spectral coverage and flexibility** – Its 80-300 MHz observing band, corresponds to coronal heights of $\sim$ 0.1–1 $R_\odot$. The 30.72 MHz bandwidth can be distributed across the entire band. Further more, an exceptionally radio-quiet site in the Western Australian outback provides seamless access to practically the entire observing band, enabling broadband studies.

Every significant advance in our solar observing capability to date has revealed that the Sun and the corona are more dynamic and richer in phenomenon than previously thought. SDO and HINODE represent the most recent steps in that progression in space and we can expect MWA to do the same from the ground.

Table 1: MWA Specifications

Frequency range	80-300 MHz
No. of receptors	8,192 dual polarization dipoles
No. of antenna tiles	512
No. of baselines	130,816
Collecting area	$\sim$8000m^2
Field of view	$\sim$15°–50° (1000 deg^2 at 200 MHz)
Array Configuration	Core array $\sim$1.5 km diameter (97% of area) + Extended array $\sim$ 3 km diameter
No. of spectral channels	768 (40 kHz spectral resolution)
Temporal resolution	0.5 s uncalibrated 8 s calibrated
Polarization	Full Stokes
Continuum point source sensitivty	20mJy (1 s, 200 MHz, full bandwidth) 0.34mJy (1 hr, 200 MHz, full bandwidth)
Array voltage beams	32, single polarization

Early results from the MWA Prototype – A 32 element engineering protytpe (32T) currently exists on site. In addition to serving its primary role of providing a test bed for end-to-end engineering validation of hardware and software, it also provides an early science opportunity. Our expectations from the MWA are borne out by our recent results from the 32T (Oberoi et al. 2010b). Our snapshot images have dynamic ranges of $\sim$2500, and we see short-lived, narrow-band, non-thermal features significantly weaker than seen before. It is evident that in many respects we have stepped well beyond the capabilities of other present day instruments. As it is built out to completion from its current stage, the capabilities of the MWA solar observing will continue to grow



Fig. 1.— One of the 32 MWA tiles currently deployed in the Western Australian outback at Boolardy.

steadily. The full MWA will offer an order of magnitude improvement in angular resolution, 16 times as much collecting area and 16^2 times as many baselines, leading to significant improvements in imaging sensitivity dynamic range.

Technical Risk – The development phase for much of MWA technology is essentially over, and it is expected to be deployed in the field in the next few years. The prototype hardware has been deployed in the field and is performing well. The project successfully cleared a NSF mandated detailed testing and verification exercise early this year. The number of elements of an interferometer add complexity to its design and operation, and in this regard, the MWA is in a league of its own. There is a technical challenge associated with constructing, commissioning and operating such an interfeometer. The MWA Project is well aware of it and considerable effort is being expended to anticipate and retire the risks on these fronts.

Cost and leveraging of resources – For the MWA project, the total cost of physical completion is estimated to in the 20–25 MUSD range, met by comparable contributions from US and Australia and a smaller in-kind contribution from India. This price-tag is dominated by the NRE cost for development of a radically new technology. A proposal seeking the US share of construction funding for the MWA has recently been submitted to the Astronomy division of the NSF. We anticipate submitting proposals to the Atmospheric and Geospace Sciences division of NSF seeking support for the aspects mentioned here at appropriate times in the future.

Astro-2010 endorsement of HERA – The Hydrogen Epoch of Reinoization Array (HERA) Phase II was endorsed by the recently concluded Astronomy and Astrophysics Decadal Review (Astro-2010) as a compelling project suitable for funding through the proposed mid-scale instrumentation program at $85 M level. A prerequisite for this is the successful completion of HERA Phase I, a project comprising MWA and Precision Array for Probing Epoch of Reionization (PAPER). HERA I received a strong recommendation from the *Panel on Radio Millimeter and Submillimeter Astronomy from Ground* (RMS Panel) for robust funding support over the next five years.

Timeliness – The construction of the MWA is expected to be funded by other sources leading to commissioning of the full MWA in mid-2012. The support to do the necessary preparatory and collaborative work needed to be able to extract the maximum science benefit from the MWA would be very timely. The availability of supporting complementary, detailed and sensitive, observations from platforms like Solar Dynamics Observatory (SDO), HINODE and STEREO which will provide information about other independent pieces of the puzzle, will enable us to construct a more complete picture than possible otherwise.

Enabling Solar Imaging Science – As mentioned earlier, the design, technology development and construction of the MWA is being funded from independent resources. The recent endorsements from the Astro-2010 decadal survey bodes well for the support of low-frequency radio astronomy effort over the next decade. We however need to put in the effort to fund the work needed to leverage the US investment in MWA to maximize the solar science returns. We list below

the major tasks which will need to be resourced by the MWA solar and heliospheric community, ranging from operations of the array to building the tools and collaborations for interpretation of the data. These tasks are either common to both solar and heliospheric science or specific to solar science. A companinon white paper (Oberoi et al. 2010a) presents the case for heliospheric low radio frequency science.

- **MWA operations** – The only science applications for the MWA when the Sun is above the horizon will be solar, heliospheric and ionospheric science. We will need to pick up our share of the tab for operational cost of the MWA.

- **Solar data storage and management** – Solar science requires data at higher time resolution than the MWA defaults (8 s), so a subset of the MWA data will be directly written to disc at the native resolution of the correlator (0.5 s). The fraction of data which can be written out will be determined in part by the available capacity for high bandwidth data capture, storage and management.

- **Augmenting MWA firmware for solar applications** – Much of the MWA functionality resides in the firmware implemented on the digital boards and the baseline design provides only one mode of operation. It is possible to design observing modes better suited for solar and heliospheric science by suitably modifying the firmware. These modes can range from a comparatively simple trade off of some the reduandant baselines for higher time resolution, to including another layer of beamforming at the receivers to reduce the number of elements by a factor of 8 and gain the same factor in bandwidth, giving us simultaneous imaging capability across the entire 80-300 MHz.

- **Data analysis effort** – We have been working the data from the 32T and it has already been very rewarding (Oberoi et al. 2010b). In order to better understand both the coronal emission and instrumental characteristics and calibration, we need to sustain and build up this effort.

- **Radiative transfer through the corona** – The same effects which make low-radio frequency coronal imaging sensitive to coronal conditions and rich in its information content, also make it more complicated to model. In order to be able to interpret coronal radio images and relate them to models of coronal electron density, temperature and magnetic field distributions, one need to build a pretty sophisticated software suite capable of ray tracing and following the evolution of Stokes parameters along the ray path in magnetized plasma. We have recently developed a numerically efficient algorithm for ray tracing in the solar corona (Benkevitch et al. 2010) and have just started working tracing the evolution of Stokes parameters.

- **Developing an interferometric data analysis pipeline** – It will be impossible to analyze the very large amounts of data generated by the MWA following the traditional human

intensive inteferometric data analysis approach. A sophisticated and efficient data analysis pipeline will necessarily be required. This effort is expected to leverage NSF investments in development of *Common Astronomy Software Applications*[1] (CASA).

- **Extraction of information from 4D data-sets and data visualization** – With its 769 spectral channels and 0.1 s time resolution, the 32T can provide about half a million images a minute (460,800). Though the temporal resolution of the MWA will be a factor of 5 coarser, the fact remains that it is humanly impossible to extract information from such a large number of images. Some clever techniques for data visualization and analysis will be needed to distill the useful information into a more digestible form and volume.

- **Ensuring compatibility of HERA with solar science needs** – As plans progress for HERA Phase II, the solar community will need to engage with the HERA astronomy community to ensure that the instrument design and capabilities meet the needs of the solar science and resource this effort.

- **Building appropriate collaborations** – To derive the maximum science benefit from the MWA observations, one will use complementary information available in other wavebands - optical, UV/EUV and X-rays. Significant effort will need to be spent in setting up and supporting productive collaborations to bring together the relevant pan-spectrum expertize.

REFERENCES

Bastian, T. S., & Gary, D. E. 1997, J. Geophys. Res., 102, 14031

Bastian, T. S., Pick, M., Kerdraon, A., Maia, D., & Vourlidas, A. 2001, ApJ, 558, L65

Benkevitch, L., Sokolov, I., Oberoi, D., & Zurbuchen, T. 2010, ArXiv e-prints

Jackson, B. V., & Sheridan, K. V. 1979, Proceedings of the Astronomical Society of Australia, 3, 383

Lonsdale, C. J., et al. 2009, IEEE Proceedings, 97, 1497

Magdalenić, J., Vršnak, B., Pohjolainen, S., Temmer, M., Aurass, H., & Lehtinen, N. J. 2008, Sol. Phys., 253, 305

McLean, D. J. & Labrum, N. R., ed. 1985, Solar radiophysics: Studies of emission from the sun at metre wavelengths (Cambridge University Press)

[1]CASA is a radio astronomical analysis software being developed in support of the next generation of radio telescoped by an international consortium of institutions including National Radio Astronomical Observatories in the US. More information available at `http://casa.nrao.edu`.

Oberoi, D., Benkevitch, L., Cappallo, R. J., Lonsdale, R. J., Matthews, L. D., Whitney, A. R., & the MWA Collaboration. 2010a, White paper submitted to Decadal Strategy for Solar and Space Physics

Oberoi, D., et al. 2010b, To be submitted to Astrophysical Journal Letters, currently under MWA Collaboration review. In the interim available at http://www.mwatelescope.org/tmp

Pick, M., Kerdraon, A., Auchère, F., Stenborg, G., Bouteille, A., & Soubrié, E. 2009, Sol. Phys., 256, 101

Pick, M., & Vilmer, N. 2008, A&A Rev., 16, 1

Ramesh, R., Kathiravan, C., Barve, I. V., Beeharry, G. K., & Rajasekara, G. N. 2010a, ApJ, 719, L41

Ramesh, R., Kathiravan, C., & Sastry, C. V. 2010b, ApJ, 711, 1029

This preprint was prepared with the AAS LaTeX macros v5.2.

Deepening the Solar/Stellar Connection for a Better Understanding of Solar and Stellar Variability

A White Paper Submitted to the Heliophysics Decadal Survey

Rachel A. Osten, Space Telescope Science Institute





The Sun has had enormous scientific attention paid to it due to its proximity to Earth and the important influences it has on space weather for our increasingly technical society. Yet the Sun is only the closest of billions of stars in our galaxy. And while it has been the subject of daily ground- and space-based observations for decades now, the Sun provides a sample size of one for understanding stellar astrophysics. The primary focus of this white paper is on solar and stellar variability. The purpose is to motivate a deepening of the solar/stellar connection, by recognizing it as a mutual relationship from which both solar physicists and stellar astronomers can benefit. With the impressively detailed nature of solar observations, one may think of the solar/stellar connection as more akin to a one-way street than a multi-lane throughway. However, by exploring the range of behaviors that stars exhibit we can access regions of parameter space which it is not possible to probe with solar observations alone. This white paper addresses four key questions which can be probed with a deeper connection between solar and stellar variability.

Question 1: How typical is the Sun as a solar-like star?

Solar flares produce emissions across the electromagnetic spectrum, from km-wavelength radio emission to gamma-rays. The interrelationship between various emissions can be used to infer the dynamics and energetics of the flare process, and thus gain insight into the physics involved in flares. Solar white-light or visible flares are difficult to see on the Sun, due to the generally small amount of contrast between the disk-integrated light and the small physical area involved in producing white-light flare emissions. The importance of the white light flare arises from their dominance in the flare energy budget of the radiative emissions from the lower chromosphere and photosphere via UV and white-light continuum emissions (Woods et al. 2006), even if the flare contrast is not as high as in other wavelength ranges (5-30% continuum enhancement in the models of Allred et al. (2005). The frequency of solar flares is distributed in energy according to a power-law in the differential flare frequency distribution $dN/dE \propto E^{-\alpha}$ with $\alpha=1.7\pm0.1$ (Kucera et al. 1997), where dN/dE is the flare frequency per unit flare energy. The largest solar flares release about 10^{32} erg in a matter of minutes (Woods et al. 2004) while a more typical energy release is 10^{29} erg. The flare frequency is such that the largest events are observed typically only a few times per solar cycle of 11 years. Detailed studies of the energy budget in solar flares and coronal mass ejections (Emslie et al. 2005) led to an estimate that the total solar irradiance as diagnosed in the UV-optical bands indicated total radiated energies roughly 100 times that determined from X-ray studies alone.

Studies of stellar variability hold the key to understanding the potential extremes of solar behavior. Infrequent events can be detected much more easily on a large sample of stars than by intense scrutiny of one Sun. For example, determination of flares as large as the Carrington flare on the Sun (Cliver & Svalgaard 2004) are currently suggested to occur every 500 years, but that is an uncertain number. Schaefer et al. (2000) found evidence for superflares on solar-like stars with radiated energies up to 10^6 times larger than the largest solar flares, with estimated frequencies of once every several hundred years. As human society becomes more technologically advanced, we expose

ourselves to increasing vulnerability to solar behavior. Having a better sense of the frequency of such events lets us know what precautions might be needed. The Kepler mission currently operating has a primary mission of finding eclipsing extrasolar planets, but with a monitoring campaign of >100,000 stars over 3+ years, it is collecting a vast repository of stellar data, which can be used to investigate sunspot cycles and flare variability. This is an example of where stellar input data can be used by solar physicists to gain a better perspective of the Sun as a star.

Question 2: How do flares on solar-like stars and the Sun compare to flares on extremely magnetically active stars?

While observations of stellar flares are not as numerous as those of flares on our well-studied Sun, the general consensus appears to be that stellar flares comprise the same basic physical processes occurring in solar flares. This is despite the fact that the types of stars studied so far can be significantly different from the Sun: F dwarfs (Mullan & Mathioudakis 2000), G and K giants (Ayres et al. 1999, Ayres et al. 2001, Testa et al. 2007), tidally locked RS CVn binary systems (Osten et al. 2004), dMe flare stars (Osten et al. 2005), flares on hyperactive young Suns containing star-disk interactions (Favata et al. 2005), and very low mass stars near the substellar limit (Stelzer et al. 2006). With the advent of sensitive X-ray and radio telescopes, multi-wavelength campaigns on magnetically active stars have revealed many similarities in behavior between solar and stellar flares, which implies that common physical processes unite the phenomena. Yet, because of the observational biases in the types of stars targeted for in-depth study, our understanding of flares and associated effects may be incomplete.

Despite the generally good agreement between solar flare models and stellar flare observations, there are still areas of apparent disagreement. M dwarfs have been the subject of flare monitoring due to their propensity for frequent and extreme flaring. The white light and continuum response in a stellar flare on a dMe flare star observationally may take the form of a blackbody with temperatures near 10^4 K and area coverages $\sim$0.01% of the stellar disk area (Hawley et al. 2003), with the flare enhancement peaking to the blue end. Radiative hydrodynamic simulations of the response of a stellar flaring atmosphere to the input of a beam of electrons can reproduce other features observed in stellar flares, such as line emission and observed velocity shifts and Stark broadening, but still cannot reproduce the magnitude of the observed continuum enhancements (Allred et al. 2006). This is in contrast to the same kinds of models applied to solar flares, which do show general agreement between observed white-light flares and models (Allred et al. 2005). The solar flare beam heating in an M dwarf stellar atmosphere does not produce the observed continuum increases. The electron beam cannot penetrate deeply into the photosphere to provide direct heating which would result in the observed continuum emission, and backwarming from the X-ray and EUV emission cannot provide the required energy input.

The distribution of stellar flares with energy on extremely magnetically active flaring stars appears to be similar to the distribution of large solar flare events (Osten & Brown 1999, Wolk et al. 2005), with evidence for M dwarf flare stars to have a distribution

favoring smaller amplitude events (α in the flare frequency-energy distribution ≥2; Güdel et al. 2003). The typical individual large stellar flare represents energetic events at least as powerful as the largest solar flares (and often larger), by virtue of sensitivity thresholds. Power stellar flares many orders of magnitude larger than the largest solar flares have been observed: radiated energies at least 10^{38} erg, and possibly larger, have been observed on the most hyperactive flare stars (Favata et al. 2005, Osten et al. 2007). Flares whose luminosity at X-ray energies approach or even exceed the star's normal radiative output across all wavelengths (Favata et al. 2000, Osten et al. 2007, Osten et al. 2010) have been detected. These observations call out for a description of the maximum flare energy and energy storage in active regions from a theoretical perspective, which the solar community could provide.

Interpretation of stellar flares requires the spatial resolution of solar flare data to provide plausible scenarios for what might be happening on flaring stars. The interaction between stellar astronomers and solar physicists in this regard can be a fruitful one, as stellar astronomers can gain from the large solar datasets available, and solar physicists can learn how stellar flares differ from solar flares.

Question 3: What is the dependence of flare rate on stellar parameters?

Due to their dependence on the presence and dynamics of interacting magnetic fields, stellar flares are a transient diagnostic of magnetic activity, in addition to more commonly used persistent indicators such as chromospheric/coronal emissions, starspots, and rotation. Magnetic activity is a function of convection zone depth and rotation rate, as quantified by the Rossby number (Noyes et al. 1984). This relates ultimately to the dynamo generation of magnetic fields producing the large-scale magnetic fields whose stresses and plasma interactions cause the mechanical heating and reconnection observed as different magnetic activity phenomena.

In solar-like stars, magnetic activity is a function of age, with young stars exhibiting magnetic hyperactivity, and a consequent fall-off with increasing age accompanying spin-down. Likewise, the stellar flare rate appears to decline as a function of age. In the young suns of Orion (0.9-1.2 M_{sun}, F7-G5 dwarfs, age ~10^6 years), Wolk et al. (2005) estimated X-ray flares more energetic than 10^{34} erg occurred roughly once per week. This minimum flare energy is more than 100 times more energetic than the largest observed solar flares. Stelzer et al. (2000) did a systematic study of the X-ray flare rates of T Tauri stars at ages 1-10 MY, and studies of stars in the Pleiades (age ~100 MY) and the Hyades (age ~600 MY). They noted that the G stars in their group had the smallest flare rates by spectral type. Corrected for sensitivity biases, their flare rate for the Pleiades was 0.67±0.13%, for the Hyades 0.32±0.17% expressed as a fraction of time in the flaring state compared to total observation time.

That magnetic activity can survive to old age is evidenced by the flare observed serendipitously on the old solar neighborhood M dwarf Barnard's star, at an estimated age of 11-12 GY (Paulson et al. 2006). The decline of flaring with age is generally assumed to follow the trends of other activity indicators, but this is an open question.

Evidence that magnetic activity may not decline monotonically at GY ages comes from a few sources: Silvestri et al. (2005) concluded that chromospheric activity in M dwarfs did not decline in the 1-10 GY range as fast as predicted based on extrapolating from objects with ages < 1 GY. Studies of chromospheric activity and its dependence on age and rotation led Pace & Pasquini (2004) to conclude that they found no evidence of decay in chromospheric activity for stars older than a few GY; instead, the major decline in activity was in objects at ages of the Hyades and earlier (0.6 GY), for clusters of 1.7 GY and older (up to 4.5 GY) the same activity level was seen. Flares as magnetic activity indicators are more sensitive to the transient build-up and release of magnetic stresses, which may happen less frequently but survive for longer than persistent levels of magnetic activity. Thus determining the existence of flares in a large sample of cool stars is a method complementary to measuring the amount of magnetic activity using other means. A large sample size can additionally constrain the dependence of flare rate on other stellar parameters, such as spectral type and rotation rate.

The primary benefit would be inputs to dynamo models, as flares are fundamentally an expression of the dynamo generation of magnetic fields.

Question 4: What are the astrobiological implications of stellar flares?

The effects of solar flares are experienced dramatically here on Earth -- disruptions to power grids and spectacular aurorae are among the many manifestations. Solar storms heat the upper atmosphere causing increased drag which accelerates the orbital decay of satellites. The Halloween solar flares of 2003 produced significant effects at Earth -- the Earth's Van Allen belt suffered a major distortion by the radiation and particles in the flares, being compressed by a factor of two (Baker et al. 2004). Even larger flares than observed from the Sun have potential implications for solar system evolution and habitability. Support for the existence of large stellar flaring events at young ages comes from meteorite studies, which indicates indirectly that energetic events were common in the young Sun. The mystery of melting chondrules in meteorite studies has been suggested as due to magnetic reconnection flares in the early Sun (Feigelson & Montmerle 1999). Likewise, one hypothesis to explain isotopic abundance anomalies in meteorites uses MeV particle spallation of meteoritic material, which could arise from particle production in stellar flares. The results of Favata et al. (2005) suggest that large flares on young star can involve loops connecting the star with its accretion disk, and flares may stimulate turbulence in the disk, which affects planetary formation. While still extremely speculative, such a scenario does cast stellar flares in a favorable role.

The renewed interest in planets around M dwarfs (Tarter et al. 2007) has enlivened the debate about the effects of stellar flares on exoplanets in such systems, due to the known high flare rate of dMe flare stars. Paradoxically, while the discovery of planets around M dwarfs (starting with Marcy et al. 1998) has led to the revitalization of the idea of the impact of flares on planets, the currently poorly constrained superflare frequency on solar-like stars has gained little attention. Smith et al. (2004) showed that a significant fraction (up to 4%) of the ionizing radiation from a transient flare event can be converted into UV emissions which reach the planetary surface, and that this is orders

of magnitude higher than the fraction of the primary ionizing radiation which reaches the surface. Khodachenko et al. (2007) and Lammers et al. (2007) considered the effect of a nearly continual rate of coronal mass ejections and flaring on an exoplanet in the habitable zone around an active M dwarf. They concluded that planets with weak magnetic moments would have expanded atmospheres, which coupled with compression of the magnetosphere might result in atmospheric erosion. Recently, Segura et al. (2010) calculated the effects of radiation and protons on the atmospheric chemistry of an Earth-like planet in the habitable zone around an M dwarf, and concluded that flares may not present a direct hazard for life.

While life did evolve around one flaring G star, the effect that flares and associated events (mass ejections, energetic particles) are predicted to have on other planets still varies widely. This is largely an result of incomplete knowledge about how to extrapolate solar flare data to stellar scenarios, as well as partial inclusion of possible effects in such calculations. With the discovery of exoplanets around a magnetically active late-type star (Lanza et al. 2010) and the coming flood of M dwarf Earth-like exoplanet discoveries, such investigations will quickly need to provide real constraints on habitability for specific systems. Solar physicists can aid in this with studies correlating key parameters which are currently unconstrainable from the stellar side.

Summary

The primary plea is for solar physicists to talk to stellar astronomers and vice versa. There is a wealth of solar data, and a relative paucity of stellar data. Despite that seeming imbalance, there are areas where interaction can be mutually beneficial. Stellar astronomers can mine solar data in ways that solar physicists might not have anticipated. This requires a funding mechanism that crosses the boundary lines between pure solar research and pure stellar research.

References

Allred, J. C., et al. 2005 Astrophysical Journal 630, 573
Allred, J. C., et al. 2006 Astrophysical Journal 644, 484
Ayres, T. R., et al. 1999 Astrophysical Journal 526, 445
Ayres, T. R., et al. 2001 Astrophysical Journal 562, L83
Baker, D. N., et al. 2004 Nature 432, 878
Cliver, E. W. & Svalgaard, L. 2004 Solar Physics, 224, 407
Emslie, A. G., et al. 2005 Journal of Geophysical Research, 110, A11103
Favata, F., et al. 2000 Astronomy & Astrophysics, 353, 987
Favata, F., et al. 2005 Astrophysical Journal Supplement Series 160, 469
Feigelson, E., D, & Montmerle, T. 1999 Annual Reviews of Astronomy & Astrophysics 37, 363
G¨ udel, M.,et al. 2003 Astrophysical Journal 582, 423
Hawley, S. L., et al. 2003 Astrophysical Journal 597, 535
Khodachenko, M. L. et al. 2007 Astrobiology 7, 167
Kucera, T. A., et al. 1997, Astrophysical Journal, 475, 338

Lammers, H. et al. 2007 Astrobiology 7, 185
Lanza, A. F. et al. 2010 Astronomy & Astrophysics 520, 53
Marcy, G. W., et al. 1998 Astrophysical Journal 505, L147
Mullan, D. J., & Mathioudakis, M. 2000 Astrophysical Journal 544, 475
Noyes, R. W. et al. 1984 Astrophysical Journal, 287, 769
Osten, R. A., & Brown, A. 1999 Astrophysical Journal 515, 746
Osten, R. A., et al. 2004 Astrophysical Journal Supplement Series 153, 317
Osten, R. A., et al. 2005 Astrophysical Journal 621, 398
Osten, R. A., et al. 2007 Astrophysical Journal 654, 1052
Osten, R. A. et al. 2010 Astrophysical Journal 721, 785
Pace, G.; Pasquini, L. 2004, Astronomy & Astrophysics 426, 1021
Paulson, D. B., et al. 2006 Publications of the Astronomical Society of the Pacific 118, 227
Schaefer, B. E., et al. 2000 Astrophysical Journal, 529, 1026
Segura, A. et al. 2010 Astrobiology 10, 751
Silvestri, N. M., et al. 2005 The Astronomical Journal 129, 2428
Smith, D. S., et al. 2004 Origins of Life and Evolution of the Biosphere 34, 513
Stelzer, B.,et al. 2000 Astronomy & Astrophysics, 356, 949
Stelzer, B.; et al. 2006 Astronomy & Astrophysics 460, L35
Tarter, J. C. et al. 2007 Astrobiology, 7, 30
Testa, P., et al. 2007 Astrophysical Journal, 663, 1232
Wolk, S. J., et al. 2005 Astrophysical Journal Supplement 160, 423
Woods, T. N.,et al. 2004 Geophysical Research Letters, 31, 10802
Woods, T. N., et al. 2006 Journal of Geophysical Research, 111, A10S14

The Total and Spectral Solar Irradiance Sensor: Response to the National Academy of Science Decadal Strategy for Solar and Space Physics

Peter Pilewskie, Greg Kopp, and Erik Richard
Laboratory for Atmospheric and Space Physics, University of Colorado at Boulder

Robert Cahalan
NASA Goddard Space Flight Center

William F. Denig
National Geophysical Data Center, NOAA/NESDIS

1. Introduction

Radiative energy from the Sun establishes the basic climate of the Earth's surface and atmosphere and defines the terrestrial environment that is humanity's habitat. Solar radiation powers the complex and tightly coupled atmospheric dynamics and chemistry, driving interactions among the atmosphere, oceans, ice, and land. External solar variability on a wide range of temporal and spatial scales ubiquitously affects the Earth system, and combines with forcings internal to the climate system, including anthropogenic changes in greenhouse gases and aerosols, natural modes such as ENSO, and volcanic forcing, to define past, present, and future climates. Understanding these effects requires: 1) Measurement continuity via a 2014 launch of the JPSS/TSIS; 2) climate model refinements using recently available total and spectral solar irradiances; and 3) interagency support.

Although relative variations in the Sun's radiative inputs since the Industrial Revolution are small, the climate sensitivity uncertainties from solar forcing are comparable to those associated with greenhouse gases and aerosols [*IPCC Fourth Assessment*, 2007]. Seemingly small variations in solar irradiance represent large impacts on, or uncertainties in, the total energy input to the Earth. Total solar irradiance (TSI) measurements provide the only *quantitative* record constraining proxy and physical models used for estimating past solar irradiances that are essential for a definitive understanding of the historical record of climate change. Establishing the input solar energy baseline provides the foundation for evaluating all other climate forcings, including those caused by human activities.

The current uninterrupted 32-year TSI climate data record is the result of several overlapping instruments flown on different missions. This record clearly exhibits variability over the 11 year solar cycle and on shorter time scales. Measurement continuity has allowed successive instruments to be linked to the existing data record despite offsets between instruments, but does make this important climate data record susceptible to loss in the event of a gap in measurements. Future improvements in instrument accuracy will reduce this risk, but *a timely 2014 launch of the JPSS/TSIS, intended to provide continuity after the Glory/TIM, is critical.*

A newer record of solar spectral irradiance (SSI) commenced with the launch of the NASA SORCE mission in 2003. *The measurement of SSI is vital for understanding how solar variability impacts climate and for validating climate model sensitivity to spectrally varying*

solar forcing. These data also provide a basis for improving synthetic solar spectra, which can be used to model paleoclimate and validate hypotheses on past and future climate change.

The continuous high-precision TSI record acquired from research missions over the last three decades constrains the total energy available to drive the climate system. The newer SSI measurements provide the details to elucidate the underlying mechanisms and terrestrial interactions responsible for Sun-induced climate changes. Both measurements are essential for constraining the energy input to the climate system and for interpreting the response of climate to external forcing.

Prior decadal surveys have been influential in supporting these measurements with recommendations to move them from research to operational realms to guarantee their long-term continuity, but the ongoing transition to operations relies on further interdisciplinary and interagency support. Many solar irradiance missions are supported by NASA Earth Sciences programs with intent to transition to NOAA for operations. Meanwhile, the physical understanding of why the solar irradiance varies has been advanced primarily through support from Solar and Heliophysics programs at NASA, while climate modeling aspects are often supported by the NSF. Funding to continue and maintain this long-term solar data record is currently not well covered well by any one agency: NASA does not "monitor" long-term irradiances, but NOAA has yet to institute a solar monitoring capability, and the research aspects providing physical understanding are diversely spread and sporadic. *It is critical that agencies work together to ensure that the measurements transition to operational status, and that research funding is provided for the long time periods needed for climate understanding.*

2. Background

Solar radiation is the Earth's primary source of energy, exceeding by four orders of magnitude the next largest source, radioactive decay from the Earth's interior [*Sellers*, 1965]. At the top of the atmosphere (TOA) the radiative balance between the incoming solar radiation and outgoing scattered solar and emitted infrared radiation defines the planet's radiative effective temperature of 255 K, which is significantly lower than the average surface temperature (288 K) because of the atmosphere's greenhouse effect.

Energy from the Sun establishes the basic structure of the Earth's surface and atmosphere and defines its external environment. Solar radiation powers the complex and tightly coupled atmospheric circulation and chemistry, and the interactions among the atmosphere, oceans, ice, and land that maintain the terrestrial environment as humanity's habitat. Natural variability on a wide range of temporal and spatial scales is ubiquitous in the Earth system, and this constant change combines with anthropogenic influences to define the net system state in the past, present, and future.

For this reason a reliable, continuous record of solar irradiance (total and spectral) is essential for climate change understanding and attribution, a fact that is recognized by a number of funding agencies and international organizations. For example, NASA identified solar irradiance as one of 23 crucial measurements for the NASA Earth Observing System (EOS) program over ten years ago [*EOS Science Plan*, 1999]. More recently, the Global Climate Observing System (GCOS, a joint undertaking of the World Meteorological Organization, the Intergovernmental

Oceanographic Commission of the United Nations Educational Scientific and Cultural Organization, the United Nations Environment Programme, and the International Council for Science) has designated solar irradiance (total and spectral) as one of 27 Global Essential Climate Variables and recommends its continuity in the *Implementation Plan for the Global Observing System for Climate* [2010]. Solar irradiance a designated Environmental Data Record (EDR) by the now-reorganized National Polar-orbiting Operational Environmental Satellite System (NPOESS) and remains a required Climate Data Record (CDR) under the NOAA-NASA Joint Polar Satellite System (JPSS).

A continuous 32-year record of total solar irradiance exists from space-based observations, as shown in Figure 1. Evident in this combined record is an 11-year cycle with peak-to-peak amplitude of approximately 0.1% and variations a factor of two to three times greater associated with the short-term transits of sunspots over the disk of the Sun. Variability in TSI occurs over a broad range of time scales, from day-to-day variations through the 11-year solar cycle and longer. Because the Sun's energy input to the Earth is so large, even the small relative fluctuations that occur during the 11-year solar activity cycle can cause detectable climate responses [e.g., *Haigh*, 2003; *van Loon et al.*, 2007]. Similar variability levels are likely to occur on longer time scales, and may have been the chief contributor to warming in the first half of the twentieth century [e.g., *Tett* et al., 2002; *Stott* et al., 2003]. However, the amplitude of long-term change must be deduced indirectly from proxy records tied to the existing TSI data record, which is too short to fully identify the long-term physical mechanisms of solar variability. The *Fourth Assessment Report* of the *Intergovernmental Panel on Climate Change* [IPCC AR4, 2007] estimates the direct radiative forcing due to changes in the solar output since 1750 to be 0.12 W m^{-2} (from its baseline value of 1361 W m^{-2}) with a factor of two uncertainty and a low level of scientific understanding. Shortly prior to 1750, the Maunder Minimum, which corresponded with the Little Ice Age in Europe, may have caused even greater changes in solar forcing.

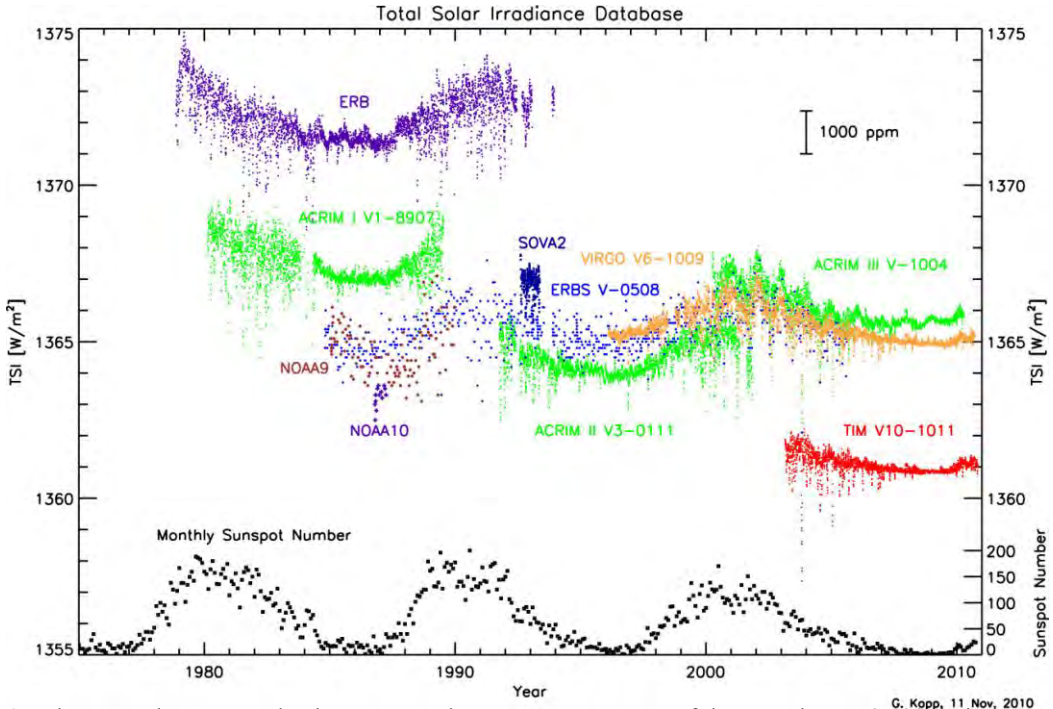


Figure 1. The space-borne TSI database. Space-borne measurements of the TSI show ~0.1% variations with solar activity on 11-year and shorter time scales. Offsets between instruments are the results of calibration differences.

The measurements made by individual radiometers providing the data in Figure 1 [prepared by *G. Kopp*, 2010] exhibit a spread of nearly 1% that is of instrumental rather than solar origin and far exceeds the 11-year or rotational solar variability. The individual TSI datasets in Figure 1 from 1978 to the present time include observations made by ERB on Nimbus-7; ACRIM-I on SMM, ACRIM-II in UARS, and ACRIM III on ACRIMSAT; ERBS on the ERBE satellite; SOVA on EURICA; VIRGO on SOHO; and TIM on SORCE [*Kyle et al*., 1993, *Willson*, 1994, *Froelich*, 1994, *Lee et al*., 1995, *Froelich*, 1996, *Willson*, 2001, and *Kopp et al*., 2005]. While instrument offsets are large, each instrument has high precision and is able to detect small changes in the TSI caused by variability in solar activity. Increases of 0.1% in TSI during times of high solar activity over the 11-year solar cycle are unambiguous. Short term changes of 0.3% are directly attributable to variations in solar magnetic activity [*Fligge, et al.,* 2000]. These data were all recorded with ambient temperature sensors, each of which has its own stated instrumental uncertainty, typically on the order of 0.1% (1000 ppm), with the exception of the Total Irradiance Monitor on SORCE, which has a 350 ppm uncertainty [*Kopp et al*., 2005]. Most of these instruments have internal degradation tracking methods, giving them the best stability of any on-orbit solar sensor so that long-term (secular) changes in solar variability can be monitored given measurement continuity.

The differences between current on-orbit TSI measurements are approximately 0.35%, which significantly exceeds the true solar changes. A 2005 workshop conducted at the National Institute of Standards and Technology (NIST) in Gaithersburg [*Butler et al.*, 2008] has led to investigations into the effects of diffraction and of aperture area measurements on the differences between instruments. Additional recommendations resulting from the workshop included power and irradiance calibrations, and comparisons with a standard cryogenic electrical substitution radiometer. In response to these issues, a new TSI Radiometer Facility (TRF) has been established to provide such calibrations [*Kopp et al*., 2007]. The TRF is the only facility in the world to allow direct irradiance comparisons between a TSI instrument and a reference cryogenic radiometer at full solar power levels and under flight-like vacuum conditions. New understanding of radiometer performance and the quantification of all sources of uncertainties will allow improved constructions of composite irradiance records for CDRs.

In addition to total solar irradiance, reliable knowledge of solar spectrum changes is crucial for the irradiance Climate Data Record. Because of selective absorption and scattering processes in the Earth's atmosphere, different spectral regions contribute to the vertical distribution of deposited energy in distinct ways (Figure 2) [*Pilewskie et al*., 2005]. Various possible mechanisms of climate response to solar variability have been identified and are dependent on wavelength [*Lean et al*., 2005]. The solar radiation in the

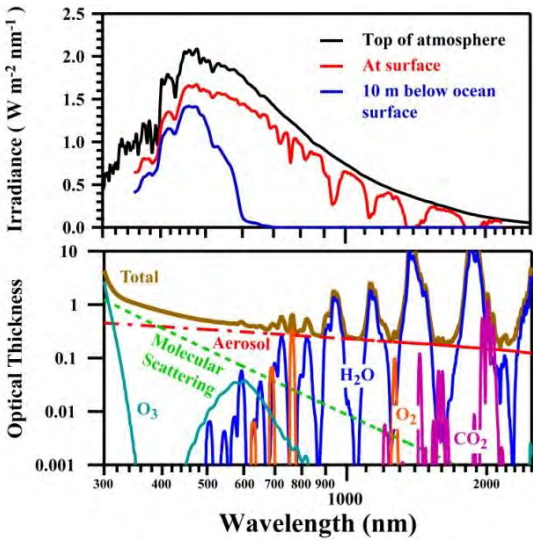


Figure 2. Top panel: SORCE Solar Irradiance Monitor [SIM; *Harder et. al*., 2005; *Harder et al., 2009*] top-of-atmosphere solar irradiance, downwelling spectral irradiance near surface, and at 10 m below ocean surface. Lower panel: Spectral optical thickness for model atmosphere showing major contributors through near-ultraviolet to near-infrared.

ultraviolet, at wavelengths shorter than 300 nm, plays a major role in the photochemistry, dynamics, temperature, composition, and structure of the middle atmosphere. The direct heating of the Earth's surface is a consequence of absorption of solar irradiance in the near ultraviolet, visible, and near infrared spectral regions. When averaged over the globe, roughly half of the incoming solar radiation is either absorbed in the atmosphere or scattered back into space, leaving the remaining half to be absorbed at the surface. Because scattering and absorption are wavelength-dependent, accurate knowledge of the spectral dependence of solar variability is essential to resolving many issues related to solar influences on climate and to determining the mechanisms of climate response to solar forcing.

Solar spectral irradiance inputs are essential for the next generation of Global Climate Models (GCMs), which include parameterizations of atmospheric processes from the surface to the mesosphere on increasingly finer altitude resolution grids and with interactive ozone photochemistry. In particular, while the GCMs in IPCC AR4 primarily used the total solar irradiance to specify solar forcing of climate change, updated state-of-the art models planned for use in IPCC AR5 will require solar spectral irradiance binned onto their respective wavelength grids. The requirements for functional solar irradiance Climate Data Records therefore include accurate time series of total *and* spectral irradiance in the ultraviolet, visible and infrared spectrum on time scales of days to centuries.

3. History of the Total and Spectral Solar Irradiance Sensor

The Total and Spectral Solar Irradiance Sensor (TSIS) is a dual-instrument package that will acquire solar irradiance in the present decade. The TSIS is comprised of the Total Irradiance Monitor, or TIM [*Kopp et al*., 2005], which measures the total solar irradiance that is incident at the top of the atmosphere, and the Spectral Irradiance Monitor, or SIM TIM [*Harder et al*., 2005], which measures solar spectral irradiance over 96% (in integrated irradiance) of the solar spectrum. The TSIS TIM and SIM are heritage instruments to those currently flying on the NASA Solar Radiation and Climate Experiment (SORCE). Both were selected as part of the TSIS because of their unprecedented measurement accuracy and stability, and because both measurements are essential to constraining the energy input to the Earth's climate system and for interpreting the response of climate to external forcing.

The measurements of TSI and SSI from the TSIS were originally mandated under the National Polar-orbiting Operational Environmental Satellite System (NPOESS) in the 1990's. TSIS was de-manifested from NPOESS following a 2006 restructuring, and re-manifested (as Government Furnished Equipment) in 2008 following a 2007 National Academy of Science workshop, *Options to Ensure the Climate Record from the NPOESS and GOES-R Spacecraft.* In early 2010 NPOESS underwent yet another major restructuring, with the National Oceanic and Atmospheric Administration taking responsibility for one of the two NPOESS polar orbits. NASA and NOAA have since established the Joint Polar Satellite System (JPSS), which will provide stewardship for TSIS and all other original NPOESS climate sensors.

The TSIS instrument package underwent Preliminary Design Review in May, 2009, and Critical Design Review in December, 2009. It is on schedule for delivery for spacecraft integration in 2012. However, the spacecraft, launch vehicle, and launch date have yet to be determined. The JPSS afternoon satellite will utilize an identical spacecraft bus to that of the NASA NPOESS

Preparatory Project, scheduled for launch in September, 2011. In its current configuration, the JPSS bus cannot accommodate TSIS. Several options for flying TSIS have been studied, including an optimal Sun-pointing free-flyer option, but a final decision has yet to be made. *Continuity with existing TSI missions, most notably the NASA Glory mission, demands that a TSIS flight decision is acted upon soon in order to meet a planned 2014 launch.*

As a free-flyer, TSIS is a small mission; estimated mission cost is $150M including launch, the TSIS TIM and SIM instruments, spacecraft, mission operations, and science data processing. The free-flyer concept utilizes a SORCE-like spacecraft, which lowers the overall cost, cost uncertainty, and mission risk.

4. TSIS Requirements

The Solar Irradiance Environmental Data Record (EDR) requirements were specified in the original *NPOESS Integrated Operational Requirements Document* (IORD) and updated in version II [2002] and in the *NPOESS Technical Requirements Document* [2002]. Solar irradiance is now recognized as a Climate Data Record (CDR) rather than an EDR to distinguish it from higher latency, lower accuracy requirements of operational *weather* data. The requirement for TSI accuracy is 100 ppm, the same as it is for the NASA Glory mission. This is approaching a level that will help mitigate the risk of a data gap by reducing measurement offsets as shown in Figure 1 and improve our ability to link the Glory and TSIS TSI data records to future measurements, providing that the future measurements are made with similar or better accuracy. The current record relies on measurement overlap and good instrument stability to link independent measurements. The TSIS Level I requirements (see Table 1) appropriately reflect the TSI accuracy and instrument stability measurement requirements to determine long-term solar variability needed for climate studies. Similarly, the requirements for the newer SSI record address current needs for solar spectral irradiance to improve our understanding of the climate system response to solar variability over shorter-duration time scales. The NOAA Climate Data Record Program is updating and refining the requirements for all Global Essential Climate Variables based on improved understanding of the climate system and its response to observable trends, and to improvements in metrology. In addition to these ongoing activities, the *Achieving Satellite Instrument Calibration for Climate Change* (ASIC³) multi-agency sponsored workshop report [2007] provided recommendations for climate observing system requirements (including those for solar irradiance) for the next decade.

Specification	Total Irradiance	Spectral Irradiance
Measurement Range	1310 to 1410 W/m^2	0 - 10 W/m^2/nm
Long Term Stability	0.001%/yr	$\lambda \leq 400$ nm: 0.05%/yr $\lambda > 400$ nm: 0.01%/yr
Measurement Precision	0.002%	0.01%
Measurement Accuracy	0.01% (0.15 W/m^2)	0.25%
Spectral Range	Entire spectrum	200 – 2400 nm
Spectral Resolution	Integrated	$\lambda < 280$ nm: 2 nm 280 nm $\leq \lambda \leq$ 400 nm: 5 nm $\lambda > 400$ nm: 45 nm
Reporting Frequency	4 6-hourly averages/day	2 spectra/day

Table 1. TSIS Level-1 requirements.

5. The Link to Heliophysics

TSIS and the NASA solar irradiance missions before it (ERBE, ACRIMSAT, SORCE, and Glory) have been funded through the NASA Earth Science Division, and rightly so considering their emphasis on climate. The National Academy of Sciences Earth Science decadal survey [*Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond*, 2007] recommended the Glory and TSIS missions for continuity of the solar irradiance record.

Recognizing the importance of these solar measurements, interdisciplinary connectivity, in particular in fields associated with the solar-terrestrial environment, is improving. The continuous measurement of solar irradiance, historically within NASA the domain of Earth science, has important ties to solar physics and space weather. Likewise, terrestrial climate change is now a component of the NASA Heliophysics division within *the Living With a Star* (LWS) Program.

One important cross disciplinary activity funded by the LWS Program is in improving solar models and our understanding of spectral solar variability. For example, the Solar Radiation Physical Model (SRPM) has been developed over several years [*Fontenla et al*., 2009] and provides the necessary physical basis for calculating high resolution spectra and their variations directly from solar theory and analysis of solar images. Using ground-based images of the solar disk and a set of models of active- and quiet-Sun features such as those of *Fontenla and Harder* [2005], the sources of the UV, visible, and IR spectral irradiance variation observed by SIM are being explored. Further studies are underway to investigate visible and infrared variations observed by SIM, and to improve modeling of SOLSTICE data in the MUV and FUV spectral regions, where large non-LTE effects play an important role in solar emissions. The extension and improvement of these studies over the full rise and decay of a solar is needed for understanding the physics of the solar atmosphere in originating the observed SSI variations. This physical insight can guide much more reliable analyses of historical records and future expectations.

Finally, it should be noted that while there is broad international support for the continuity of solar irradiance measurements from space, and that NASA Earth Science has supported and managed most of the missions, it appears that Research and Analysis (R&A) support for solar irradiance studies occasionally ~~–f~~alls through the cracks," with no proper place to call home. NASA Earth and Heliophysics divisions both provide some opportunities for solar irradiance-based studies, but generally within broader programs and at a relatively low level of emphasis. We recommend that cross-division collaborative support within NASA is needed to better address the many aspects of the solar-terrestrial environment and to provide the necessary resources to answer some of the outstanding questions in Sun/Earth/global change climate studies.

References

Barron, E. J., D. L. Hartmann, M. D. King, D. S. Schimel, and M. R. Schoeberl, *EOS Science Plan - The State of Science in the EOS Program*, Ed. EOS Senior Project Scientist, Michael D. King, NASA/Goddard Space Flight Center, Greenbelt, Maryland, http://eospso.gsfc.nasa.gov/science_plan/SciencePlan.pdf, January 1999.

Butler, J. J., B. C. Johnson, J. P. Rice, E. L. Shirley, and R. A. Barnes, Sources of differences in on-orbital Total Solar Irradiance measurements and description of a proposed laboratory intercomparison, *J. Res. Natl. Inst. Stand. Technol.,* **113**, 187-203, 2008.

Fligge, M., S. K. Solanki, and Y. C. Unruh, Modelling irradiance variations from the surface distribution of the solar magnetic field, *Astron. & Astrophys.,* **353**, 380-388, 2000.

Fontenla, J. and J. Harder*, MmSAI*, 76, 826, 2005.

Fontenla, J. M., W. Curdt, M. Haberreiter, J. Harder, and H. Tian, *ApJ*, 707:482–502, doi:10.1088/0004-637X/707/1/482, 2009.

Forster, P. et al. in *Climate Change 2007: The Physical Science Basis*, Fourth Assessment Report of the Intergovernmental Panel on Climate Change, http://ipcc-wg1.ucar.edu/wg1/Report/AR4WG1_Print_Ch02.pdf , 2007

Fröhlich, C., and the VIRGO team, First Results from the VIRGO Experiment, *Transactions AGU*, American Geophysical Union Spring Meeting, Baltimore, Maryland, 20-24 May 1996.

Fröhlich, C., *The Sun as a Variable Star*, Eds.: J. Pap, C. Fröhlich, H. Hudson, and S. Solanki, Cambridge Univ. Press, New York, pp, 28-36, 1994.

Haigh, J. D., *Royal Society of London Phil. Trans.,* Series A, 361, 95, 2003.

Harder J., G. Lawrence, J. Fontenla, G. Rottman, and T. Woods, The Spectral Irradiance Monitor: Scientific Requirements, Instrument Design, and Operation Modes, *Solar Phys.,* **230**, 141-167, 2005.

Harder, J. W. , J. M. Fontenla, P. Pilewskie, E. C. Richard, and T. N. Woods, Trends in solar spectral irradiance variability in the visible and infrared, *Geophys. Res. Lett*., 36, L07801, doi:10.1029/2008GL036797, 2009.

Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC, 2010 Update. Available at http://www.wmo.int/pages/prog/gcos/index.php.

Kopp, G., G. Lawrence, and G. Rottman, The Total Irradiance Monitor (TIM): Science Results, *Solar Phys.,* **230,** 129–139, 2005.

Kopp, Greg, Karl Heuerman, Dave Harber, and Ginger Drake, The TSI Radiometer Facility - absolute calibrations for total solar irradiance instruments. *Proc. of SPIE*, **6677**, 667709, 2007.

Kyle, H. L., D. V. Hoyt, J. R. Hickey, R. H. Maschoff, and G. J. Vallette, "Nimbus-7 Earth Radiation Budget Calibration History - Part 1: The solar channels", NASA Reference Publication 1316, 1993.

Lean, J., G. Rottman, J. Harder, and G. Kopp, SORCE contributions to new understanding of global change and solar variability, *Sol. Phys.*, **230**, 27-53, 2005.

Lee, R. B., M. A. Gibson, R. S. Wilson, and S. Thimas, Long-term total solar irradiance variability during sunspot cycle 22, *J. Geophys. Res.,* **100**, 1667-1675, 1995.

National Research Council, *Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond*, 2007.

NPOESS (National Polar-Orbiting Operational Environmental Satellite System) *Technical Requirements Document*, Prepared for: NPOESS – Integrated Program Office, Silver Springs, MD, Version 7, 24 January 2002.

Ohring, G. (ed.), Achieving Satellite Instrument Calibration for Climate Change (ASIC3), 2007.

Pilewskie, P., G. Rottman and E. Richard, An overview of the deposition of solar radiation in the lower atmosphere: connections to the SORCE mission and climate change, *Sol. Phys*., 230, 55-69, 2005.

Sellers, W., *Physical Climatology*, The University of Chicago Press, Chicago, IL, 1965.

Stott, P. A., G. S. Jones, and J. F. Mitchell, *J. Climate,* **16**, 4079, 2003.

Tett, S. F. B, and 12 co-authors, *J Geophys. Res.,* **107**, D16, 10, 2002.

van Loon, H., G. A. Meehl, and D. J. Shea, *J Geophys. Res.,* **112**, D02108, 2007.

Willson, R. C., "Irradiance observations of SMM, Spacelab-1, UARS, and ATLAS Experiments", *The Sun as a Variable Star*, eds.: J. Pap, C. Fröhlich, H. Hudson, and S. Solanki, Cambridge Univ. Press, New York, pp, 54-62, 1994.

Willson, R. C., The ACRIMSAT/ACRIM3 Experiment—Extending the Precision, Long-Term Total Solar Irradiance Climate Database, *Earth Observ.*, **13**, 14– 17, 2001.

2010 Solar and Space Physics Decadal Survey White Paper: Next Steps in Solar Spectral Irradiance Studies

Mark P. Rast[1,2,3], Jerald W. Harder[2], Roberto Casini[3], Serena Criscuoli[4], Ilaria Ermolli[4], Juan M. Fontenla[2], Hanli Liu[3], Scott W. McIntosh[3], Carolus J. Schrijver[5], Han Uitenbroek[6]

Understanding the physical causes underlying solar spectral irradiance variations and the impact of these variations on terrestrial climate remain critical compelling challenges. On the solar side, the most fundamental unknown is the role of the "quiet-sun[7]," a question brought into focus by spectral irradiance observations during the recent deep minimum, which show opposing trends in the infrared and ultraviolet portions of the spectrum (Harder et al. 2009). On the terrestrial side, while it is clear that the Sun played only a small role in late twentieth century warming, the signature of solar cycle variations in climate are now quite well established (Gray et al. 2010). How this coupling occurs is not fully understood, suggesting something incomplete in our understanding of the climate system.

We suggest that in the coming decade, through the focused efforts outlined below, we can successfully address both of these scientific challenges and emerge with a new more complete understanding of the Sun and its influence on Earth. The essential ingredients in this effort are poised at the edge of our abilities, challenging but not risky, and take full advantage of current technological capabilities. These ingredients are:

- Radiometric imaging of the Sun to assess the absolute center-to-limb brightness of magnetic structures on full disk images at key wavelengths. This can be achieved via the coupling of an imager, capable of high relative precision, and a radiometer, to calibrate the absolute image intensity, into a single space-based instrument, a **Radiometric Solar Imager (RSI)**. Both of the necessary technologies have been demonstrated operationally, and can be readily combined into a single instrument with high return and low risk.

- High-resolution observations of the magnetic features identified in full disk images from the RSI. These are necessary to assess the variance underlying the full-disk pixels as a function of structure type, disk position, and solar cycle phase. Such observations can be achieved in the next decade by the judicious use of data from the Advance Technology Solar Telescope (ATST). The synergy between full disk radiometric imaging, achievable only from space, and ultra high resolution dynamical observations, achievable only with ground based apertures, will allow irradiance modeling to be extended beyond one-dimensional static atmospheres to include the statistical properties of underlying variations due to solar dynamics.

[1] *Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder*
[2] *Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder*
[3] *High Altitude Observatory, National Center for Atmospheric Research, Boulder, Colorado*
[4] *INAF-Osservatorio Astronomico di Roma, Monte Porzio Catone, Italy*
[5] *Lockheed Martin Advanced Technology Center, Palo Alto, California*
[6] *National Solar Observatory/Sacramento Peak, Sunspot, New Mexico*

[7] "Quiet-sun" in this white paper refers to internetwork regions, as is typical in irradiance studies. Internetwork regions are identified by low Ca II K emission, and are often assumed to be regions of low magnetic flux density. "Quiet-sun" as used here should not be confused with the broader non-coronal-hole-region usage found in other discussions of chromospheric, transition region, or coronal emission.

- High-resolution radiative magnetohydrodynamic modeling of the photosphere and chromosphere. These will enable a physical understanding of the connection between small-scale field dynamics and structure and the integrated spectral output. Many realizations of such models with varying flux density and geometry will be achievable on planned petascale computing platforms in the next decade. These will allow a statistical understanding of the physical causes of spectral variability and the construction of radiative diagnostics to disentangle thermodynamic changes from unresolved magnetic flux elements in the "quiet-sun".

- Coupling of the measured or modeled solar spectral irradiance to comprehensive numerical models of the photochemical, radiative, physical and dynamical processes in the Earth's atmosphere and ocean. Such multi-component models have reached a state of maturity that makes this coupling realizable and meaningful, with major ongoing efforts aimed at efficient implementation of these codes on next generation petascale computing platforms.

The coming decade thus promises a convergence of capabilities that if focused and supported can contribute significantly to our understanding of the radiative coupling between the Sun and the Earth. Through these efforts we will gain enhanced abilities to successfully model both the Earth's regional climate variations and short and long term solar fluctuations.

Brief review of the scientific background

A critical component in understanding the causes of the Earth's climate variability is a thorough description of the Sun's role in this complex process (Lean 2000; Rind et al. 2008). Over the course of a solar cycle, as well as over centennial and millennial time scales, solar magnetic activity modulates its irradiance. Understanding the evolution and radiant output of solar magnetic features is thus at the heart of understanding the Sun-Earth connection, and radiometric imaging of the Sun is critical for both solar physics and Earth climate studies.

While it is still an open question whether global structural changes in the Sun play a role in its variability[8], particularly over long time scales, the appearance and disappearance of small scale magnetic structures is known to account for most of the variation observed over the course of a solar cycle. Sunspots and pores have negative contrast against the background disk, while smaller scale magnetic fields generally contribute positively. Changes in the fractional areas covered by structures of different intensities modulate the solar irradiance. The magnitude and even the sign of the contribution of any given magnetic structure, however, depends not only on its size, as suggested above, but also critically on the wavelength being observed and the structure's location on the solar disk. To understand solar irradiance variations three interlocking scientific activities are thus needed: 1) accurate and precise observations of the evolution of active regions (such as sunspots and bright facular regions) on the Sun through global solar imaging. This requirement for global imaging to determine the disk position of magnetic structures sets a limit on their resolution. 2) models and high resolution observations of solar magnetic structures that explain their radiant emission and the variance in this emission. This variance is caused by differences in the unresolved (sub-pixel) contributions underlying the global scale observations. 3) the

[8] The proposed effort will make a significant contribution to the resolution of this issue by measuring the radiometric contributions of magnetic structures and "quiet-sun" to solar spectral variability.

measurement of disk integrated solar spectral irradiance against which models can be tested. It is this disk integrated spectral irradiance, either observed (when available) or modeled (based on the level of solar activity), that is then used as radiative input to climate models to ultimately allow an understanding of the Sun's role in climate variability.

Harder et al. (2009) reported an important example of the effectiveness of this approach. They combined observations of solar spectral variability measured by SORCE SIM (Solar Radiation and Climate Experiment, Spectral Irradiance Monitor) and full disk images showing the evolution of solar active regions as observed by the HAO/NCAR Mauna Loa Solar Observatory's Precision Solar Photometric Telescope (MLSO PSPT; Rast et al. 2008) with theoretical interpretation using SRPM (Solar Radiation Physical Modeling; Fontenla et al. 2009) to show that both increasing and decreasing irradiance trends occurred as a function of wavelength during the descending phase of Solar Cycle 23. These offsetting trends summed to produce the observed decrease in total solar irradiance over the same period. Terrestrial climate modelers are now just starting to assess the ramifications of this result (Cahalan et al. 2010, Haigh et al. 2010).

To date all precision irradiance imaging of the Sun has been photometric, not radiometric, measuring the intensity of any pixel on the solar disk relative to some predefined background. By careful determination of the detector gain (flat-fielding) and the quiet-sun center-to-limb variation, the PSPT can achieve high precision (0.1%) relative photometry of the solar disk. *There are however no measurements of the absolute radiometric contributions of magnetic structures to the disk integrated irradiance*. This means that all observations and all modeling efforts must define a "quiet-sun" intensity, against which the contrast of the magnetic elements is determined. Since the "quiet-sun" intensity of the Sun varies from disk center to the limb, and this profile is wavelength dependent, all contrast measurements and magnetic structure identifications are made relative to some arbitrary definition of the "quiet-sun" center-to-limb profile. Moreover, the "quiet-sun" is indeed not completely quiet, and over the solar cycle the number of unresolved magnetic flux elements contributing to pixel-integrated intensities likely varies. Put starkly, we as yet have no idea whether the trends discussed above (or other variations observed) are due to changes in magnetic structure number densities, structure intensities, the "quiet-sun" reference state, or more likely, some combination of these; the possibilities cannot be unambiguously untangled without absolute radiometric imaging.

We suggest that this shortcoming can be readily overcome via a PSPT-like instrument with radiometric capabilities (the RSI). The effort would combine extant expertise in radiometry and imaging, to design an instrument capable of producing absolutely calibrated solar images at specific critical wavelengths (Table 1 summarizes possible filter bands of interest). Such an instrument would provide absolutely calibrated center-to-limb profiles of magnetic structures, and contribute the next critical step along a path that will ultimately yield a radiometric imaging spectrometer for full spatial and spectral monitoring of solar variability.

Radiometric Solar Imager (RSI)

The essential new instrumental component of solar irradiance studies over the next decade is an instrument that images the Sun at carefully chosen wavelengths while simultaneously measuring the integrated output along the same light path to produce *absolutely calibrated radiometric images*.

Table 1. Possible filter bands for the RSI instrument

Wavelength (nm)	FWHM (nm)	Use with F-P filter	Purpose/Function
CaIIK 393.4	0.02	Yes	Ca II K core and wing intensity (tunability requirements) Determination of faculae and plage regions Scattering polarization source characterization
607.1	0.5	No	Standard red PSPT continuum (instrument continuity) Determination of sunspot umbra and penumbra
409.4	0.02	Yes	Standard PSPT blue continuum (instrument continuity) Brightness temperature, relative photometry
HeI 587.6, 1083 NaID 589.0,589.6	0.2	No	Scattering polarization source characterization Chromospheric contributions
Ha 656.3	0.02	Yes	Prominence analysis Scattering polarization source characterization
CaII triplet 855	0.02	Yes	Scattering polarization source characterization
~600, ~700, 862, ~900 , ~1000, 1650	0.5	No	Photospheric temperature structure (relative photometry) Opacity minimum

What design features/requirements are needed for a compelling Radiometric Solar Imager?

- Diffraction limited images of the full solar disk.
  - This requirement sets the size of solar image to be ~0.6°, or ~0.5 arc seconds per detector pixel. The solar image must under fill the detector for scattered light determination.
- The detector relative pixel-to-pixel precision of 1 part in 10^3 is needed and must be maintained over the course of a solar cycle.
  - This requirement translates into being able to determine the brightness temperature of a 1 arc-second scene on the Sun with about 1-2 degrees Kelvin accuracy. This level of relative precision is currently achieve in the best PSPT images, and should be attainable using similar techniques from space.
- The instrument requires scientifically optimized broad wavelength band-pass filters to isolate different layers in the solar atmosphere.
  - The current PSPT instrument uses 5 filters (two Ca II K core and one wing, 607nm red continuum, 409nm blue quasi-continuum). Additional filters in the near infrared, in the neighborhood of the H^- opacity maximum, would be able to determine changes in the photospheric thermal structure.
  - Augmentation in specific chromospheric lines would enable radiometric calibration of the illumination source in the presence of magnetic homogeneities for scattering polarization measurements.
- Narrower and more uniform filter functions than can be achieved with band-pass filters are needed, leading to consideration of a tunable Fabry-Perot filter (F-P in Table 1).
  - Wavelength tuning through the Ca II K feature would allow scanning with height and careful understanding of core to wing ratio proxies.
  - Similarly, the anisotropy of the radiation will change significantly across both the Ca II K the Hα line profiles, and scanning would contribute to better understanding of the chromospheric polarization signal with implications for field measurements.
- Design must provide calibrated radiometric measurement of the light following the same light path as the beam through the optical elements.

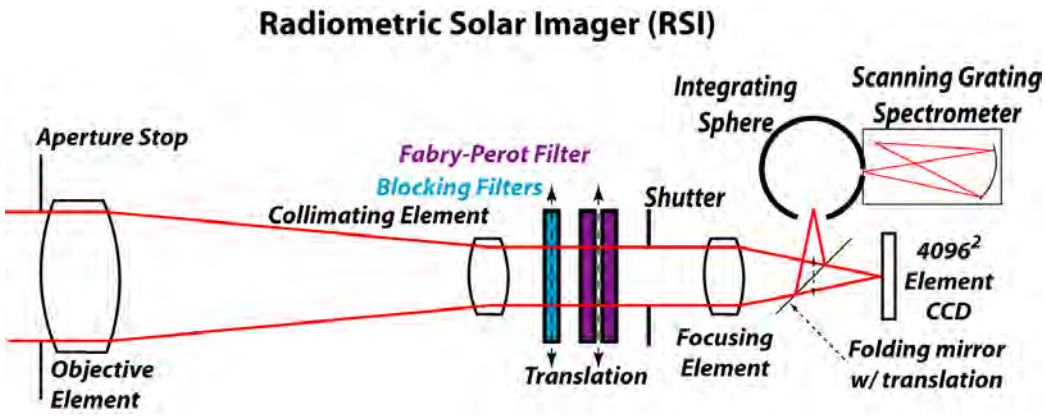


Figure 1: A concept drawing for this instrument, detailed optical design, component specification, and calibration methodology are required for a phase A instrument design.

  - *It is this part of the design that sets this instrument apart from other systems that image the Sun.* The RSI will provide absolutely calibrated solar images by coupling high precision relative photometry across the CCD to a radiometric measurement of the integrated light through the same pass-band.
  - In-flight radiometric calibrations to correct for sensitivity degradations are thus essential. Both detector-based corrections (such as the SORCE SIM, Harder et al. 2005) and light source-based calibrations (Thuillier et al. 2009) need to be considered in the design.
- The absolute preflight radiometric calibration of the instrument must be traceable NIST standards, as in the calibration plan for the upcoming SIM instrument for the NOAA TSIS mission using the NIST SIRCUS (Brown et al. 2006).
- Because of the broad wavelength coverage needed for this experiment, reflective designs are favored, but refractive equivalent designs have not been ruled out.
- This instrument could stand-alone and be flown on a mission of opportunity, or it could be flown in conjunction with other instruments on a larger mission or Small Explore Mission (SMEX). Moreover, the RSI instrument capabilities could be combined with other imager or magnetogram needs.

Importance of both global and high resolution observations

Advancing our understanding of solar spectral irradiance will require, over the next decade, the integration of global-scale relatively low- and small-scale high-resolution observations. Global imagery is required for solar spectral irradiance modeling and reconstruction because the intensity of magnetic features on the disk depends on their disk position, and so a measurement of their position is required to assess their contribution to the integrated disk irradiance. High-resolution observations are required to assess the unresolved structure and dynamics underlying the global scale observations.

Observations to date do not agree about the center-to-limb brightness of magnetic structures at global image resolution. This is not only because of wavelength dependencies and resolution differences between them, but more fundamentally because only contrast measurements at 0.1% precision have been possible to date (no radiometric imagery of that accuracy has been achieved), and there is no universal definition of the background “quiet-sun” intensity against which the contrast is to be measured. As a consequence, observations

of magnetic structure center-to-limb profiles vary widely (Ermolli et al. 2007), and models attempting to reproduce those profiles have no accepted measurement against which to assess their accuracy. The RSI will for the first time provide the absolute center-to-limb intensity of any disk pixel at critical wavelengths.

Moreover, the underlying magnetic structures are poorly constrained at full disk resolution. Pixels of the same magnetic flux density can have very different intensities depending on their underlying unresolved flux distribution (Criscuoli and Rast 2009). To understand the variance in absolute intensity observed for any one magnetic structure type, low-resolution global observations must be accompanied by high-resolution images of the same locations. These do not need to be radiometric, but they must be numerous, a collection of highly resolved tiles for each global structure type, to build up, as a function of disk position the statistics of the underlying intensity and dynamics. This is the only way that global modeling of the integrated solar spectral irradiance can move beyond one-dimensional static atmospheric models and begin to incorporate unresolved complexities.

Radiative MHD modeling underpinning

That in turn will require high-resolution radiative magnetohydrodynamic modeling of the photosphere and chromosphere that can be used to assess the fidelity of the irradiance models. Such high resolution modeling efforts will be increasingly possible with increasingly powerful computing resources in the coming decade. These should be used in conjunction with the high-resolution observations to inform the spectral irradiance modeling efforts (Domingo et al. 2009). In particular one wants to ask, what is the range and distribution of pixel intensities at the global image resolution that can be expected for a given wavelength and disk position. Many realizations of the dynamics in these simulations and the integrated output as a function of wavelength could be used to reconstruct the underlying variance of the lower resolution pixels. Global vector magnetic field observations from the Helioseismic Magnetic Imager (HMI) on the Solar Dynamics Observatory (SDO) should be used in close conjunction with the modeling to constrain the field topologies.

While the global run of temperature with height in the solar photosphere can be accurately determined from the center-to-limb variation of intensity at a given wavelength, the local thermal structure, which is needed to disentangle the irradiance contributions from different magnetic structures within a pixel, is much harder to determine and involves combining intensity measurements at different wavelengths to determine the temperature gradient. Such relative photometry (Foukal and Duvall 1985) is extremely difficult using ground based observations because of seeing induced misalignment and lack of an absolute photometric reference. RSI solves these problems. In principle, the temperature at any height can be determined from measurement of absolute intensity at a single wavelength, but multi-wavelength determination would be much preferred for accuracy. Unfortunately, even in the continuum, different wavelengths sample different atmospheric heights because of the wavelength dependence of the main opacity source in the photosphere, the H^- ion, making multi-wavelength measurement of intensity coming from a single height difficult. However, since H^- opacity increases with wavelength in the blue, reaches a maximum at 862 nm, decreases to a minimum near 1.6 μm, and then increases again towards longer wavelengths, careful selection of pass-bands at different wavelengths corresponding to equal opacity values on either side of the maximum allows a two-wavelength measurement of temperature at specific heights in the atmosphere. This procedure would ensure an accurate determination of

the thermal gradient with good height resolution, and thus potentially untangle thermal from magnetic contributions to each pixel. It would lead directly to an assessment of the role and cause of the “quiet-sun” variability.

Coupling to terrestrial climate modeling

The variability of the terrestrial climate and space environment are closely related to the variability of the solar irradiance. The extreme ultraviolet (EUV) and X-ray, which are absorbed by the thermosphere and ionosphere, can vary by tens to 100s%, while the visible bands vary by ~0.1% or less. As a result, the thermosphere and ionosphere vary significantly over a solar cycle, with the neutral temperature increasing by ~60% from solar minimum to solar maximum conditions (Schunk and Nagy 2009). Moreover, the stratosphere (van Loon and Labitzke 2000) and even the troposphere display 11-year variability on both global scales and regional scales (Haigh 1996, van Loon et al. 2007). The change, particularly in the troposphere, is stronger than expected from a linear response to the solar total irradiance change. Numerical studies have suggested that the stratosphere ozone heating and ocean-atmosphere interactions may both act to enhance the tropospheric climate response to solar signal (Kodera and Kuroda 2002, van Loon et al. 2007, Haigh et al 2005). Therefore, to correctly quantify the response of different atmospheric regions to solar variability, and to understand the mechanisms responsible for coupling these responses, it is critical to have both an accurate description of the solar spectral irradiance and a comprehensive atmospheric model that accounts for coupled photochemical, radiative, and dynamical processes in the upper and lower atmosphere and the ocean.

One example of such a model, pushing the capabilities as we enter the petascale-computing decade, is NCAR’s Community Earth System Model with the Whole Atmosphere Community Climate Model (WACCM) as its atmosphere component. The CESM consists of five components for the atmosphere, ocean, sea ice, land surface, and land ice linked through a coupler that exchanges fluxes and state information among these components. Details of CESM can be found from the NCAR CESM website. WACCM is a climate-chemistry general circulation model spanning from the Earth's surface to the thermosphere. Of particular interest for the solar variability, WACCM provides the infrastructure to incorporate solar spectral irradiance information. It is written to take advantage of modern machine architectures, with hybrid parallelization using shared memory techniques (OpenMP) within a node and a distributed memory approach (MPI) across nodes. NCAR is actively engaged in transitioning these models to machine architectures composed of hundreds of thousands of processors, and it is models like these, using accurate solar irradiance inputs, that will in the next decade untangle the link between solar and climate variability.

Synergistic science

While the primary aim of the Radiometric Solar Telescope is irradiance and climate coupling science, if appropriately designed the instrument could have significant impact on other important outstanding problems:

- Inversion of scattering spectropolarimetric data requires characterization of the illuminating source at the observation wavelength. Radiometric observations of the solar disk intensity in the presence of magnetic inhomogeneities would provide this.

The magnetic diagnostics of the chromosphere relies on the modeling of the resonance scattering polarization in many lines (some listed in Table 1). This requires the proper characterization of the illuminating source at the various wavelengths involved in the excitation of the corresponding atomic systems. In particular, the center-to-limb variation at those wavelengths is a critical quantity for the determination of the degree of radiation anisotropy. Along with the ambient magnetic field, this anisotropy is ultimately responsible for the observed polarization signature of the scattered radiation. For this reason, radiometric observations of the solar disk intensity in the presence of magnetic inhomogeneities are essential for the modeling of polarized radiation transport in the chromosphere and thus for deducing the magnetic field properties there (see McIntosh et al. White Papers).

- Large-scale flows within the solar convection zone, either associated with global scale motions or active regions should induce thermal fluctuations (Spruit 2003, Rempel 2005). Latitudinal variations in intensity have been reported in PSPT data (Rast et al. 2008) at the very limit of what can be achieved via ground-based observations. With RSI these could be tightly constrained as a function of wavelength, with implications for global dynamics and dynamo modeling efforts.

- Careful measurement of Doppler and intensity fluctuations as a function of wavelength can be used to characterize the source of the solar p-modes (Severino et al. 2008). The RSI could do this at multiple wavelengths to quantify the source properties, with implications for understanding global mode line-shapes and local helioseismic kernels.

REFERENCES:
(NOTE: These references are representative of recent work, they are not complete.)

Brown, S.W., Eppeldauer, G.P., and Lykke, K.R. (2006) Appl. Opt. **45**, 8218
Cahalan, R.F., Wen, G.Y., Harder, J.W., and Pilewskie, P. (2010) GRL **37**, L07705
Criscuoli, S. and Rast, M.P. (2009) A&A **495**, 621
Domingo, V., et al. (2009) SSRv **145**, 337
Ermolli, I., Criscuoli, S., Centrone, M., Giorgi, F., and Penza, V. (2007) A&A **465**, 305
Fontenla, J., Curdt, W., Haberreiter, M., Harder, J., and Tian, H. (2009) ApJ **707**, 482
Foukal, P. and Duvall, T., Jr. (1985) ApJ **296**, 739
Gray, L.J., et al. (2010) Rev. Geophys. **48**, RG4001
Haigh, J.D. (1996) Science **272**, 981
Haigh, J.D., Blackburn, M., and Day, R. (2005) J. Climate **18**, 3672– 3685.
Haigh, J.D., Winning, A.R., Toumi, R., and Harder, J.W. (2010) Nature **467**, 696.
Harder, J.W., Fontenla, J., Lawrence, G., Woods, T., and Rottman, G. (2005) Sol. Phys. **203**, 169
Harder J.W., Fontenla, J., Pilewskie, P., Richard E., and Woods T. (2009) GRL **36**, L07801
Kodera, K. and Kuroda, Y. (2002) JGR **107**, 4749
Lean, J. (2000) GRL **27**, 2425
Rast, M.P., Ortiz, A., and Meisner, R.W. (2008) ApJ **673**, 1209
Rempel, M. (2005) ApJ **622**, 1320.
Rind, D., Lean, J., Lerner, J., Lonergan, P., and Leboissitier, A. (2008) GRL **113**, D24103
Spruit, H.C. (2003) Sol. Phys. **213**, 1
Severino, G, Straus, T., and Steffen, M. (2008) Sol. Phys. **251**, 549
Schunk, R.W. and Nagy, A.F. (2009) Ionospheres, Cambridge University Press
Thuillier, G., Foujols, T., Bolsée, D., Gillotay, D., Hersé, M., Peetermans, W., et al. (2009) Sol. Phys. **257**, 185
van Loon, H. and Labitzke, K. (2000) GRL **27,** 2965
van Loon, H., Meehl, G.A., and Shea, D.J. (2007) JGR **112,** D02108

Approaches to Optimize Scientific Productivity of Ground-based Solar Telescopes

A White Paper Submitted to the 2012 Heliophysics Decadal Survey

K. Reardon (National Solar Observatory) [1]
S. Keil, T. Rimmele (National Solar Observatory)
K.D. Leka (Colorado Research Associates)
P. Judge, S. McIntosh (NCAR/High Altitude Observatory)
E. DeLuca (Harvard-Smithsonian Center for Astrophysics)
B. Abbett, H. Hudson (University of California, Berkeley)
G. Cauzzi (INAF – Arcetri Astrophysical Observatory)

Summary

This white paper seeks to emphasize the scientific and strategic value to the solar physics community of scientific programs that combine physical diagnostics from multiple regimes of the solar atmosphere. High-resolution, ground-based observations provide a range of observational details of the lower solar atmosphere, that could profitably be combined with simultaneous transition region and coronal observations, and compared to simulated observables from the increasingly realistic magneto-hydrodynamic (MHD) simulations. Notable advances could be made with existing instrumentation and would serve to prepare and foster the community for the Advanced Technology Solar Telescope (ATST). Support for development of robust, user-friendly data-reduction pipelines for ground-based instruments and service-mode operations are key elements in lowering the barriers of entry for non-traditional users of ground-based telescopes.

Recommendations

The challenge is to develop a strong US community, including multiple universities, that is able to make use of the unique capabilities of ground-based observations to perform innovative science covering a broad range of topics. This robust community will be crucial to producing the significant and sustained scientific output necessary for upcoming facilities such as ATST.

To achieve this, we encourage funding agencies to provide support for:

a) meetings, exchange programs, and outreach to the community with the goal to identify and foster scientific investigations that can contribute to a comprehensive picture of the dynamic solar atmosphere, including comparisons of observations with MHD simulation results;
b) data analysis workshops that will help familiarize the broader community, especially university students and young researchers, with the present capabilities and techniques in ground-based solar physics;
c) university-based studentships and post-doctoral fellowships that encourage the use of multi-wavelength diagnostics for transversal studies of the solar atmosphere;
d) the development of pathfinding observational and instrumental programs using existing telescopes and facilities (including NSO's Dunn Solar Telescope and NJIT's New Solar Telescope) in preparation for ATST;

[1] *corresponding author: reardonk@nso.edu*

e) the development of robust, well-calibrated data reduction pipelines to provide users with direct access to science-ready data;
f) experimentation with service-mode operations with existing high-resolution solar facilities, in preparation for its full implementation in the ATST operations.

The motivation for supporting such efforts as a viable approach to achieving new scientific insights and increased productivity from existing and future ground-based solar telescopes is outlined below.

Systems approach to the solar atmosphere

The solar atmosphere is characterized by multiple transitions in physical conditions as the density, temperature, and ionization degree vary by orders of magnitude over distances of only a few thousand kilometers or less. The pervasive solar magnetic field, present even in the quieter areas, interacts with convection and plasma motions to structure and dynamically heat the outer regions of the solar atmosphere (chromosphere, transition region, and corona) and accelerate the solar wind, in ways still not fully understood. Coupling between several mechanisms, including wave–mode conversion (Bogdan et al. 2003, Straus et al. 2008), non-linear processes caused by the density gradients (e.g. magneto-acoustic shocks, Hansteen et al. 2006), or the release of magnetic energy at small scales (De Pontieu et al. 2009; Edmonson et al. 2010) results in an intricate chain of processes producing the upward propagation and dissipation of non-radiative energy in the Sun's outer atmosphere.

Real progress in understanding this complex environment will only come from a "systems approach", in which complementary observations diagnosing different physical regimes are simultaneously obtained in order to encompass the full complexity of the magnetized plasma and its dynamic behavior. Our ability to pursue such an approach is facilitated by the current availability of powerful observational and theoretical tools: space missions like Hinode and SDO have been designed to explore at once multiple regions of the solar atmosphere, and 3D, MHD numerical simulations spanning the top of the convection zone all the way to the corona are becoming increasingly realistic.

The large apertures and large data fluxes available with ground-based telescopes in the visible, infrared, and millimeter wavelengths represent a crucial asset for this system approach, as they allow the acquisition of the most accurate observations in terms of spectral diagnostics, polarimetric accuracy, and spatio-temporal resolution of the lower solar atmosphere. These properties are essential for example to understand magneto-convection phenomena operating in the solar photosphere, such as the small-scale surface dynamo (Voegler & Schessler 2007) or the formation of pores and sunspots (Leka & Steiner 2001; Rempel et al. 2009). Ground-based facilities also provide the most efficient access to diagnostics of the chromosphere (van Noort & Rouppe van der Voort 2006; Cauzzi et al. 2008; Rouppe van der Voort et al. 2009), an atmospheric region that is increasingly understood as a crucial interface between the photospheric "source" of energy and mass and the resulting coronal structure (see also the white paper by McIntosh et al 2010). In particular, the measure of the intensity and topology of the chromospheric magnetic field at high resolution, currently possible only with ground-based instrumentation (e.g. Socas Navarro 2005; Casini et al. 2009; Judge et al. 2010), holds promise of providing meaningful boundary conditions for the extrapolation of such fields in the corona, and

hence an understanding of the mechanisms leading to the heating and accelerating of plasma in the corona and the solar wind.

These capabilities will only be accentuated with the arrival of the Advanced Technology Solar Telescope (ATST, Keil et al 2010), which in addition will possess coronagraphic capabilities to make possible direct measurements of the coronal magnetic field and unmatched spatial resolution for observations of coronal loop structures. Co-temporary new radio and millimeter facilities (ALMA, FASR) also will provide further, complementary diagnostics on the upper solar atmosphere.

Community Involvement

As mentioned before, multi-telescope observations have often been adopted as a fruitful venue of investigation for the solar atmosphere (e.g. the Joint Observing Programs of SOHO). Yet, while satellite data has become more accessible with the years (e.g. through the SOHO, RHESSI, and Hinode data centers), the increasingly complicated observing programs and data reduction schemes for ground-based data have made it a sometimes daunting task to analyze such observations. The striking high-resolution observations now being produced from ground-based instruments are generally the result of significant post-processing image reconstruction (van Noort et al. 2005; Woeger et al. 2008).

In order to increase the accessibility of ground-based data, developing more robust data reduction pipelines for high-resolution instruments appears mandatory. On the one hand, such a development would enable more multi-wavelength, transversal studies of the solar atmosphere. On the other, and more crucially, it would help attract more users for these instruments by drastically lowering the barriers-of-entry for the large portion of the community who are not familiar with the intricacies of reducing ground-based data. Even though well-calibrated data reduction pipelines are not trivial to develop, the delivery of reduced data has certainly contributed to the great productivity of recent space missions (the most famous example being the magnetic maps produced by MDI/SOHO), as well as the open data policy. While the stochastic nature of the atmospheric distortions certainly complicate the analysis of ground-based data, so far only few resources have been available for producing robust software packages for ground-based instruments. Presently such efforts are largely unfunded for ground-based facilities, but a relatively small investment in this direction would produce a significant return from existing and future instruments.

It is our conviction that more readily available ground-based data would greatly strengthen collaborations between the now distinct "high-resolution" and "coronal" communities. This would have great scientific and strategic value as the community prepares for the arrival of ATST, which will be a major investment in the US scientific infrastructure. Achieving a significant and sustained scientific output from the facility will demonstrate the robustness of the underlying community. Performance metrics for astronomical facilities, typically related to the number of papers published using data from a given telescope scaled to operational costs (Trimble & Ceja 2008, Simons 2009) indicate the need attract a sizable user base for ATST and its data products. It is expected that ATST will attract international users as well, but a strong participation of US researchers will be crucial.

Research specialized on high-resolution, ground-based observations is currently concentrated in a limited number of universities and research centers in the US. The challenge is to develop a diverse group of users, in multiple locations, able to perform innovative science covering a broad range of topics that will in the future make full use of ATST's observational capabilities. University programs should be encouraged that train students and young researchers in the skills needed for the analysis of ground-based data, especially solar spectroscopy. The NSO User's Committee recently recommended that there be "a conscious effort to train/involve/encourage wider US participation in the user community for [ATST]." Given that most scientists and programs are already heavily overextended, this goal might be best achieved by identifying research areas in which ground-based data can contribute to ongoing scientific studies.

It is easy to find examples. Simultaneous chromospheric and coronal data would help better identify the source of transition region explosive events and coronal jets (Reardon et al. 2008; McIntosh & De Pontieu 2009). Solar flares are best studied with information about both the chromospheric mass reservoir and the coronal energy release (Fletcher 2010). Ground-based observations are crucial in the validation and refinement of MHD simulations (e.g. Leenaarts et al 2009; Fang et al. 2010) and their extension into the plasma-physics domain. Chromospheric magnetic field measurements, currently available only from the ground, may provide one of the best sources of information on the coronal field configurations (see white paper by Judge, 2010). Measurements of the changes in the chromospheric magnetic fields before and after flares can be compared to theoretical models of reconnection invoking modifications in the field configuration.

Service-mode operations is a model in which observations are scheduled on a flexible, daily basis in order to optimally match the science programs to the current observing conditions, and thus does not require the PI's presence at the telescope. This approach is increasingly common at nighttime facilities (Jorgensen et al. 2010), but is not routinely offered at ground-based solar telescopes. Yet, its implementation will lower the investment of individual researchers in getting their desired datasets while increasing the overall efficiency and scientific productivity of the facility. Such an observing mode could be implemented and experimented with using existing ground-based facilities (e.g. Reardon et al. 2009) to increase the community's familiarity with this type of observation in preparation for ATST.

Modeling of magnetic flux emergence across scales

Matthias Rempel[1], Yuhong Fan[1], Aaron Birch[2], Douglas Braun[2],
Mark Cheung[3], William Abbett[4]

November 12, 2010

Abstract

Magnetic flux emergence is the key process responsible for solar magnetic variability and associated impact on the heliosphere on short time scales of up to a few days. The underlying physical process couples the solar interior, photosphere, chromosphere, corona and heliosphere and encompasses therefore a wide range of scales and distinct physical regimes that make building a comprehensive model a very challenging task. In recent years substantial progress has been made in individual component models, however the comprehensive coupled approach is still outstanding. The availability of high performance computing resources combined with high resolution data available from the NASA-SDO mission will allow for substantial progress in this field, it requires however a large scale coordinated effort in model development, model coupling and data assimilation that goes far beyond past investigations.

1 Introduction

Some of the most fundamental questions of solar magnetism are related to the solar dynamo and the flux emergence process, which transports magnetic field from the solar interior toward the photosphere and beyond. Due to the intrinsic complexity and vast scale separation it is currently not feasible to model the entire system, instead components are modeled with varying degrees of sophistication. Future progress requires coupling of component models and we feel that in particular models of flux emergence reached a degree of sophistication where this can be achieved on a 5-10 year time scale. A comprehensive model of flux emergence will address the following primary science questions that relate strongly to our understanding of solar variability and its heliospheric impact:

1. What are the physical processes that lead to active region and sunspot formation?
2. How do flare and CME preducing coronal structures form and how are they connected to the interior magnetic fields?
3. Which aspects of the subsurface evolution and structure can be constrained and predicted through helioseismic inversions?

Addressing these questions will require models of flux emergence that start in the solar interior, couple to detailed models of photospheric dynamics and continue the emergence process into the solar corona. The foundation for such an effort is given by recent developments we summarize in this paper.

[1]National Center for Atmospheric Research
[2]Colorado Research Associates, NorthWest Research Associates Inc.
[3]Lockheed Martin Solar and Astrophysics Lab
[4]University of California, Berkeley

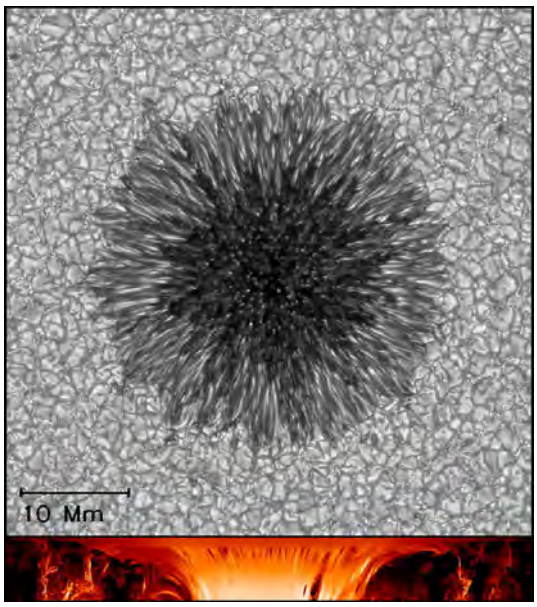


Figure 1: Numerical sunspot simulation including a realistic equation of state and 3D radiative transfer. The domain size is $49 \times 49 \times 6.1$ Mm, the grid resolution $16 \times 16 \times 12$ km ($3072 \times 3072 \times 512$ grid points). The simulation resolves sunspot fine structure from umbral dots to penumbral filaments. An Evershed flow with about 5 km/s average flow velocity is driven self-consistently by magneto-convection in the sunspot penumbra.

2 Recent progress

Due to the vast range of length and time scales as well as different physical regimes associated with the flux emergence process, different models and numerical approaches are used to describe the flux emergence process within the convection zone, the photosphere and corona:

2.1 Flux emergence in bulk of the convection zone

It is generally accepted that sunspots form from magnetic field rising from the base of the convection zone to the surface (see reviews by Moreno-Insertis, 1997; Fisher et al., 2000; Fan, 2009a, and further references therein). The first global models of the flux emergence process in the convection zone were based on the thin flux tube approximation (Choudhuri & Gilman, 1987; Fan et al., 1993, 1994; Moreno-Insertis et al., 1994; Schüssler et al., 1994; Caligari et al., 1995). These models demonstrated that the global properties of active regions (low latitude of emergence, latitudinal trend in tilt angles as well as asymmetries between leading and following spots) can be understood on basis of this picture, provided the initial field strength at the base of the CZ is sufficiently large. It was also early realized by Schüssler (1979) that untwisted magnetic flux tubes cannot rise coherently and fragment. Moreno-Insertis & Emonet (1996); Emonet & Moreno-Insertis (1998) demonstrated later that this problem can be alleviated in the case of twisted flux tubes. More recently 3D simulations of rising flux bundles in the convection zone became possible (Fan, 2008) that basically support previous results, however, it was also found that there is a very delicate balance between the amount of twist required for a coherent rise and the amount of twist allowed to be in agreement with observations of sunspot tilt angles. A major shortcoming of models up to this point was the absence of convective motions in background stratification. Using a global convection zone model including connection and differential rotation Jouve & Brun (2007, 2009) recently addressed the interaction between convection and rising flux, however, due to the focus on the global scale there was not enough resolution to properly address details.

2.2 Sunspot fine structure and flux emergence in photosphere

While most numerical simulations in the deep convection zone are based on the anelastic approximation and use a simple ideal gas equation of state, photospheric MHD simulations are fully compressible, include a realistic equation of state and use 3D radiative transfer. Over the past

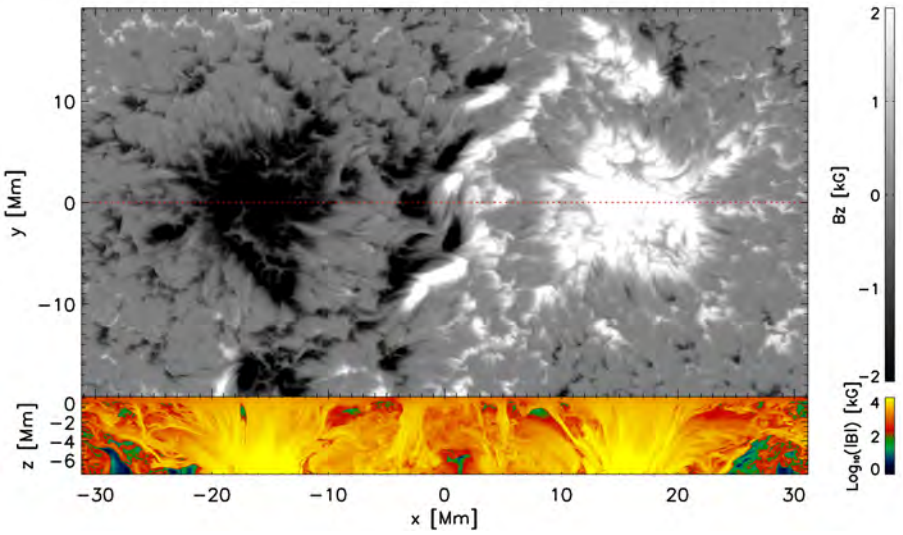


Figure 2: Flux emergence simulation reproduced from Cheung et al. (2010). The upper panel shows a synthetic magnetogram (sampled at $\tau_R = 0.1$) of a simulated emerging flux region $t = 15.3$ hrs after simulation initialization. The lower panel shows the vertical cross-section of the magnetic field strength ($\log_{10} B$) along $y = 0$ (delineated by red dashed line).

few years there has been a tremendous progress leading to a new realism in theoretical modeling of sunspot structure and a unified understanding of magneto-convection in strong field regions (Schüssler & Vögler, 2006; Heinemann et al., 2007; Rempel et al., 2009a,b; Kitiashvili et al., 2009). Over a time span of only 5 years models evolved from umbral dot simulations (Schüssler & Vögler, 2006) in about 5 Mm wide domains to simulations of entire sunspots in 50 Mm wide domains providing a comprehensive understanding of magneto-convection in sunspots. A result from the most recent high resolution simulation (Rempel, 2010) is presented in Fig. 1. In parallel, simulations of magnetic flux emergence in the photosphere were progressed (Cheung et al., 2007, 2008, 2010). Cheung et al. (2010) demonstrated for the first time how magnetic flux that was initially dispersed due to the emergence through the highly stratified convection zone can be re-amplified into strong ~ 3 kG spots in the photosphere. The simulation utilized a domain size of $100 \times 50 \times 8$ Mm coming close to a small active region. A photospheric magnetogram and the subsurface field strength at 15.3 hrs after simulation initialization is presented in Fig. 2. The compact sunspots forming have a flux of about $3 \cdot 10^{21}$ Mx. Currently work is in progress that expands these simulations to 150×75 Mm wide and 16 Mm deep domains, reaching a total flux exceeding 10^{22} Mx. The vertical density stratification beneath the photosphere captured by this setup will be about 10^4, which is a substantial fraction of the convection zone (10^6).

2.3 Emergence of active region flux into the solar atmosphere and the corona

Understanding how magnetic flux, energy, and helicity are transported from the interior into the corona in emerging active regions is a key for understanding the development of precursor structures for solar eruptions. Eruptive events such as CMEs and flares, among the principal drivers of space weather events here at earth, are believed to be manifestations of a buildup and release of free magnetic energy in the solar corona. The physics of CME initiation is not yet well understood, thus hampering efforts to reliably predict their occurrence. Whatever mechanisms are responsible, the ultimate source of magnetic energy and helicity must originate in the turbulent interior where the solar magnetic field is generated. NASA missions such as SDO are showing us that magnetic fields from the interior emerge through the surface, and energize the dynamic chromosphere and

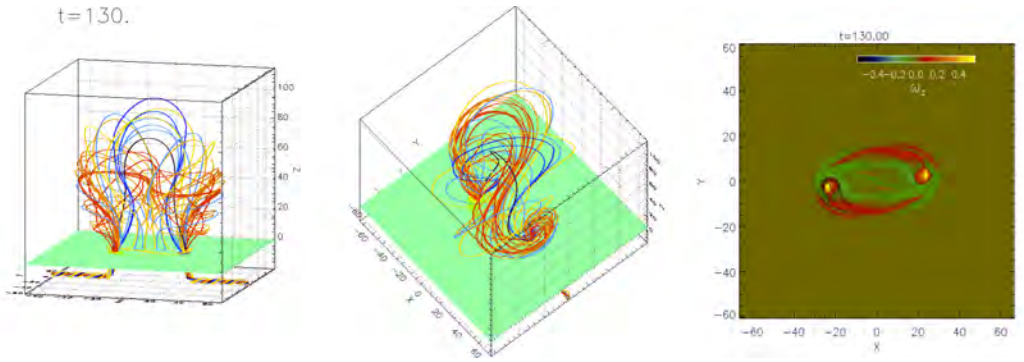


Figure 3: The left and middle panels show two perspective views of the 3D subsurface and coronal magnetic field resulting from a 3D MHD simulation of the emergence of a twisted magnetic flux tube from the interior into the corona (Fan , 2009b). The right panel shows the vertical component of vorticity on the photosphere over laid with contours of vertical magnetic field. It shows counter-clockwise vortical motions concentrated in the two polarity concentrations. The simulation domain spans $20\times19\times20$ Mm^3 at a resolution of $320\times300\times392$. Vertically the domain spans from 3 Mm below the photosphere to 17 Mm above the photosphere, into the corona.

corona over a wide range of spatial and temporal scales.

There have been a large body of 3D MHD simulations with highly simplified treatment of the thermodynamics, aimed to study the dynamic emergence of a twisted magnetic flux tube from the top layer of the solar interior into the solar atmosphere and the corona (e.g. Fan , 2001; Magara & Longcope , 2001; Magara , 2004; Manchester et al. , 2004; Archontis et al. , 2004; Murray et al. , 2006; Magara , 2006; Archontis & Toëroëk , 2008; Archontis et al. , 2009; Fan , 2009b; Archontis & Hood , 2010). These simulations demonstrated that magnetic flux reaching the photosphere can undergo a dynamic expansion into the atmosphere as a result of the non-linear growth of the magnetic buoyancy instability. These simulations also suggest that it is difficult for a subsurface twisted flux tube to rise bodily into the corona as a whole due to the heavy plasma that is trapped at the bottom concave (or U-shaped) portions of the winding field lines. Shear and rotational flows on the photosphere driven by the Lorentz force of the twisted flux tube during flux emergence are the crucial means whereby twist is transported from the interior into the solar corona, driving the formation of a coronal flux rope with sigmoid-shaped, dipped core fields (e.g. Magara & Longcope , 2003; Manchester et al. , 2004; Magara , 2006; Manchester , 2007; Fan , 2009b; Archontis et al. , 2009; Archontis & Hood , 2010, see also Figure 3). The models suggest that rotating sunspots may be a manifestation of nonlinear torsional Alfvén waves propagating along the emerging flux tube, transporting twist from the tube's interior portion, where the rate of twist is high, towards its expanded coronal portion, where the rate of twist is lower (Longcope & Welsch , 2000; Fan , 2009b).

Progress is being made towards building a realistic radiation MHD model of magnetic flux emergence, encompassing both the magneto-convection process of the top layer of the convection zone, and the emergence of the active region flux into the stably stratified solar atmosphere and the corona with realistic treatment of the thermodynamics. Examples of such numerical models include, Abbett (2007, 2010) and Martínez-Sykora et al. (2008, 2009). Recent observations show that the interplay between granular scale magnetic features, and large-scale structures from both emerging and decaying active regions can affect global coronal magnetic topology in complex and fascinating ways. Being able to model this physics in a manner that admits to direct comparison with observations would be of enormous utility in our effort to understand the physics of coronal heating, the solar wind, and magnetic eruptions (among other phenomena). Figure 4 shows a radiative-MHD simulation of granular convection in a coronal hole region. Although this particular

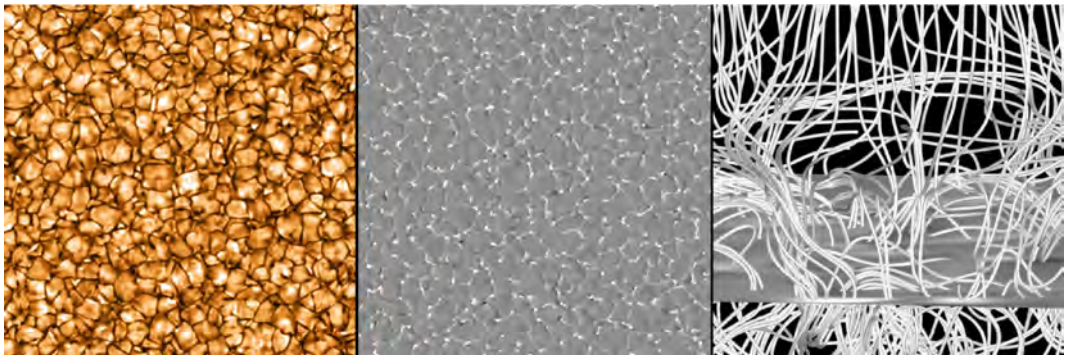


Figure 4: *Left:* The temperature from a RADMHD (Abbett, 2007, 2010) simulation of a coronal hole region on the Sun. *Middle:* The vertical component of the magnetic field (note that there is a net positive vertical component since this particular simulation was constructed to study the interactions of large and small scale magnetic structures in an open field region). *Right:* Magnetic fieldlines threading the convection zone and atmosphere in a small sub-domain showing the complex connectivity of the Sun's magnetic field. The simulation domain spans $24\times24\times12$ Mm^3 at a resolution of $512\times512\times256$. The model convection zone is 2.5 Mm deep, the domain includes a chromosphere, transition region, and corona

simulation is designed to study the energetics of the solar wind in a quiescent region, it shows an example of recent progress in the effort to simulate the convection zone-to-corona system within a single computational domain.

Can we expect to use numerical models to make sense of the complex physics of the transition layers between the turbulent convection zone and the magnetically-dominated corona over large, active region spatial scales? Realistically simulating the combined system at large scales over extended periods of time presents a number of daunting challenges. While the distance separating the low corona and visible surface is a tiny fraction of the Sun's radius, physical conditions change drastically in these layers. The highly-stratified solar atmosphere spans a ten to twelve order of magnitude change in density and transitions from an environment where convective energy transport is the dominant process and magnetic fields are confined by gas pressure, to an environment where magnetic pressure dominates gas pressure, and energy transport is controlled by radiative processes. Characteristic time scales vary from weeks of solar time (as active region magnetic fields evolve) to fractions of a second (as eruptive events restructure the corona).

In addition, the radiation field transitions from an optically-thick regime in the solar photosphere, to a transparent, optically-thin regime in the high-temperature corona. This transition is particularly challenging to model in a realistic fashion, since in the solar chromosphere, many of the energetically important radiative transitions are decoupled from the local temperature of the plasma (a state of non-local thermodynamic equilibrium), and other physical processes such as cross-field diffusion likely play a significant role.

2.4 Helioseismology on scale of sunspots and active regions

An intrinsic problem of flux emergence simulations is the unknown initial state of the rising magnetic flux rope beneath the photosphere. Ultimately this uncertainty will be addressed once a fully consistent 3D solar dynamo model becomes available. In the meantime simulations have to make assumptions about the subsurface field and explore the parameter range to find the most realistic precursor structure for emerging flux. In this situation additional information from helioseismology on either the field or flow velocities just before or during emergence of magnetic field are extremely valuable and necessary to constrain models.

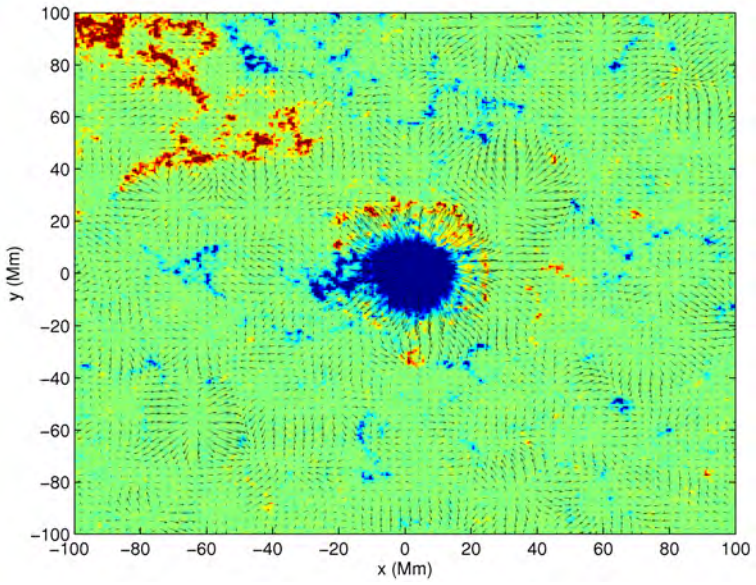


Figure 5: Horizontal flows (arrows) inferred from helioseismic holography measurements made using HMI/SDO data. The color image in the background is a HMI line-of-sight magnetogram, with the color scale saturated at ±300 Gauss. The moat flow from the sunspot is visible as are the outflow associated with the supergranulation.

Kosovichev (2009) applied time-distance helioseismology to a few very large emerging active regions. In these case studies, the emerging active regions were associated with outflows and upflows. These results suggested that a large sample of emerging AR may be needed to draw statistically significant conclusions. Birch et al. (2010) computed the signal-to-noise ratio that would be expected for time-distance measurements of the flow field from the model "LNT" (Fan, 2008) for a flux concentration rising from the base of the convection zone. This work showed that a sample of 150 emerging AR would be sufficient to detect emerging AR at up to 5 days before emergence. Hartlep et al. (2010) claimed that the emergence of magnetic flux is preceded by a reduction in the power in the observed acoustic wave field and suggested that this reduction may be associated with changes in the subsurface wave speed. It will be important to determine if this effect is generally associated with emerging active regions, and to what extent it is consistent with numerical models of flux emergence.

HMI/SDO data is very promising for the helioseismic study of the emergence of magnetic flux as it provides almost continuous full-disk coverage at high resolution. As a result, essentially all active regions that emerge on the visible disk can be studied using local helioseismology. Figure 5 shows an example inversion from helioseismic holography for flows in the depth range of one to two Mm. Subsurface flow maps promise to help constrain models of emerging magnetic flux.

2.5 Code coupling and data-driven simulations

By carrying out numerical MHD simulations in controlled settings, the studies cited above have shed light on the physical mechanisms responsible for the transport of magnetic flux from the solar interior to the solar atmosphere. Such numerical models have been successful in the sense that many synthetic observational diagnostics from the models are, in a statistical sense, similar to their counterparts from actual solar observations. The next logical step in the verification of the models is to perform data-driven simulations to facilitate direct comparisons between models and observations. This pursuit is particularly timely in light of the high-cadence photospheric magnetograms and EUV images being delivered by the Solar Dynamics Observatory (SDO), which may well serve the community with high quality data for years beyond the standard mission length. The complexity of the magnetic structures revealed by SDO compells us to construct numerical models to help interpret the observations.

To develop techniques that will allow us to perform numerical simulations mimicking observations to a high-degree of fidelity in a data-driven fashion, attempts should be made to couple codes

that treat the convection zone/photospheric layer with codes that treat the higher atmosphere. The coupling of such codes has been attempted in the past (Abbett & Fisher , 2003). With the recent availability of codes which have a more comprehensive treatment of the physics essential for the convection zone and atmosphere, renewed attempts should be encouraged. An approach which may advance the maturity of data-assimilation techniques is the following. Consider a model which simultaneously treats the convection zone, photosphere, chromosphere and corona (e.g., Abbett, 2007). From such a model, synthetic photospheric observables (such as vector magnetograms) can be produced and used as bottom boundary conditions for simulations which include only regions above the photosphere in their computational domains. This is not a trivial problem since photospheric observables do not necessarily constrain the bottom boundary conditions (e.g., consider the 180 deg ambiguity of vector magnetograms). Once such models can successfully reproduce the atmospheric evolution of the original reference model, they can perhaps be used for data-driven simulations using real solar data.

3 Future steps, conclusions

We summarized above recent developments of individual component models for the flux emergence process on the sun. The next steps toward a comprehensive model requires improvement of these individual component models and coupling codes with different physical ingredients, numerical schemes and domain sizes. Such model will address fundamental questions of solar magnetism (active region and sunspot formation) and the physical mechanism underlying eruptive events on the sun (CMEs, flares).

The challenge, then, is to (1) develop and refine the theoretical and computational techniques necessary to model these physically, spatially, and temporally disparate regimes in a way that retains the essential physics well enough to interpret current and future ground and space-based observations; and (2) couple models of the complex transition between the upper convection zone and corona to large scale models of flux emergence deep in the convection zone, high-resolution models of active region sunspots, and large-scale models of the solar corona and heliosphere. The implementation of these models has to happen such that they can efficiently utilize current and future computing resources with 100,000+ CPUs.

While incremental progress has been made in each of these areas, a program of focused and sustained research is necessary to make significant forward progress in the future. Such program would need to encourage large teams with expertise in modelling, computational science and data analysis. The outcome of such an activity would strongly support current and future space missions by providing data mining resources and through sufficiently realistic modeling activities that enhance the physical interpretation of observational data. Due to the required collaborative effort over a long time frame support within GI, TR&T, HTP, or SR&T programs is currently not sufficient. We anticipate funding requirements on a $2M per year level for a period of up to 5 years.

The next decade will provide the stage for such an effort, since (1) computing resources have reached the threshold at which realistic flux emergence models at the scale of active regions are feasible and (2) missions such as SDO will provide data with an unprecedented quality to constrain and drive such models.

The need for large scale model development and computing resources goes far beyond the example of flux emergence we highlighted in this white paper. Modeling throughout the heliosphere as well as the overwhelming data rate from current missions such as SDO requires coordinated modeling and data mining efforts that are currently not present. The latter was recently expressed in the ACCEHS workshop (`http://www.hao.ucar.edu/ACCEHS/index.php`). See also the separate white paper on this topic for further information.

References

Abbett, W. P. 2007, ApJ, 665, 1469

Abbett, W.P., & Fisher, G.H., 2003, ApJ, 582, 475

Abbett, W. P. & Fisher, G. H., 2010, *Mem. Soc. Astron. Ital.*, in press

Archontis, V., Moreno-Insertis, F., Galsgaard, K., Hood, A., & O'Shea, E., 2004, ApJ, 426, 1047

Archontis, V., & Toëroëk T., 2008, A&A, 492, L35

Archontis, V., Hood, A.W., Savcheva, A., Golub, L., Deluca, E., 2009, On the Structure and Evolution of Complexity in Sigmoids: A Flux Emergence Model, ApJ, 691, 1276

Archnotis, V., & Hood, A. W. 2010, "Flux emergence and coronal eruption", A&A, 514, A56

Birch, A. C., Braun, D. C., & Fan, Y. 2010, ApJ, 723, L190

Caligari, P., Moreno-Insertis, F., & Schüssler, M. 1995, ApJ, 441, 886

Cheung, M. C. M., Rempel, M., Title, A. M., & Schüssler, M. 2010, ApJ, 720, 233

Cheung, M. C. M., Schüssler, M., & Moreno-Insertis, F. 2007, A & A, 461, 1163

Cheung, M. C. M., Schüssler, M., Tarbell, T. D., & Title, A. M. 2008, ApJ, 687, 1373

Choudhuri, A. R. & Gilman, P. A. 1987, ApJ, 316, 788

Emonet, T. & Moreno-Insertis, F. 1998, ApJ, 492, 804

Fan, Y. 2001, ApJ, 554, L111

Fan, Y. 2008, ApJ, 676, 680

—. 2009a, Living Reviews in Solar Physics, 6, 4 http://solarphysics.livingreviews.org/lrsp-2009-4

—. 2009b, ApJ, 697, 1529

Fan, Y., Fisher, G. H., & Deluca, E. E. 1993, ApJ, 405, 390

Fan, Y., Fisher, G. H., & McClymont, A. N. 1994, ApJ, 436, 907

Fisher, G. H., Fan, Y., Longcope, D. W., Linton, M. G., & Pevtsov, A. A. 2000, Sol. Phys., 192, 119

Hartlep, T., Kosovichev, A. G., Zhao, J., & Mansour, N. N. 2010, ArXiv e-prints 1003.4305

Heinemann, T., Nordlund, Å., Scharmer, G. B., & Spruit, H. C. 2007, ApJ, 669, 1390

Jouve, L. & Brun, A. S. 2007, Astronomische Nachrichten, 328, 1104

—. 2009, ApJ, 701, 1300

Kitiashvili, I. N., Kosovichev, A. G., Wray, A. A., & Mansour, N. N. 2009, ApJ, 700, L178

Kosovichev, A. G. 2009, Space Sci. Rev., 144, 175

Longcope, D.W., & Welsch, B.T., 2000, ApJ, 545, 1089

Manchester IV, W., 2007, ApJ, 666, 532

Manchester IV, W., Gombosi, T., DeZeeuw, D., Fan, Y., 2004, ApJ610, 588

Magara, T., 2004, ApJ, 605, 480

Magara, T., 2006, ApJ, 653, 1499

Magara, T., & Longcope, D.W., 2001, ApJ, 559, L55

Magara, T., & Longcope, D.W., 2003, ApJ, 586, 630

Martínez-Sykora, J., Hansteen, V., & Carlsson, M. 2008, ApJ, 679, 2008

Martínez-Sykora, J., Hansteen, V., & Carlsson, M. 2009, ApJ, 702, 129

Moreno-Insertis, F., Caligari, P., & Schuessler, M. 1994, Sol. Phys., 153, 449

Moreno-Insertis, F. & Emonet, T. 1996, ApJ, 472, L53+

Moreno-Insertis, F. 1997, Memorie della Societa Astronomica Italiana, 68, 429

Murray, M.J., Hood, A.W., Moreno-Insertis, F., Galsgaard, K., & Archontis, V., 2006, ApJ, 460, 909

Rempel, M. 2010, Proceedings IAU Symp. 273, ArXiv e-prints 1011.0981

Rempel, M., Schüssler, M., & Knölker, M. 2009a, ApJ, 691, 640

Rempel, M., Schüssler, M., Cameron, R. H., & Knölker, M. 2009b, Science, 325, 171

Schüssler, M. 1979, A & A, 71, 79

Schüssler, M., Caligari, P., Ferriz-Mas, A., & Moreno-Insertis, F. 1994, A & A, 281, L69

Schüssler, M. & Vögler, A. 2006, ApJ, 641, L73
Vögler, A., Shelyag, S., Schüssler, M., Cattaneo, F., Emonet, T., & Linde, T. 2005, A & A, 429, 335

THE HELIOSPHERIC INTERACTION WITH THE LISM: OBSERVATIONS AND MODELS

John D. Richardson[1], John Belcher, Leonard Burlaga, John Cooper,Alan Cummings, Robert Decker, Joe Giacalone, George Gloeckler, Don Gurnett, Matthew Hill, Randy Jokipii, William Kurth, Ralph McNutt, Norman Ness, Merav Opher, Nikolai Pogorelov, Edmond Roelof, Gary Zank, Ming Zhang

[1]*Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology*

OVERVIEW

A key science goal of the last decadal survey and of the NASA Heliophysics Roadmap is to understand the interaction between the heliosphere and the local interstellar medium (LISM). Voyagers 1 (V1)and 2 (V2) are observing the heliosheath in situ and are on their way to the LISM. The Interstellar Boundary Explorer (IBEX) is observing energetic neutral atoms (ENAs) generated by charge exchange in the heliosheath, giving a global view of the heliosphere. And the theory community has developed more sophisticated tools for modeling these data. This paper proposes to continue the Voyager Interstellar Mission and to support modeling efforts to make progress toward this goal.

INTRODUCTION

Figure 1 shows a model of the heliosphere with the Voyager trajectories superposed (courtesy of S. Borovikov and N. Pogorelov). The model is run for solar minimum conditions with the HCS tilted 20° from the solar equator. The HCS varies sinusoidally in the supersonic solar wind upstream of the termination shock. At 80-100 AU, the solar wind encounters the termination shock (TS) and becomes subsonic. The flow speed in the heliosheath beyond the TS is much smaller than in the solar wind, so the HCS is compressed. The outward speed further slows as the plasma approaches the heliopause (HP), the boundary between the solar wind and interstellar medium.

The flow in the heliosheath diverts to carry plasma down the heliotail; the flow in this model has a large poleward component in the northern hemisphere which carries the HCS northward toward the V1 location. The TS and HP are asymmetric, closer to the Sun in the south (V2 direction) than the north (V1 direction) due to the direction of the LISM magnetic field. The HP is expected to be several tens of AU beyond the TS. Beyond this boundary may be a bow shock where the incoming LISM plasma would undergo a shock transition if it were supersonic. The region between the HP and bow shock is known as the outer heliosheath.

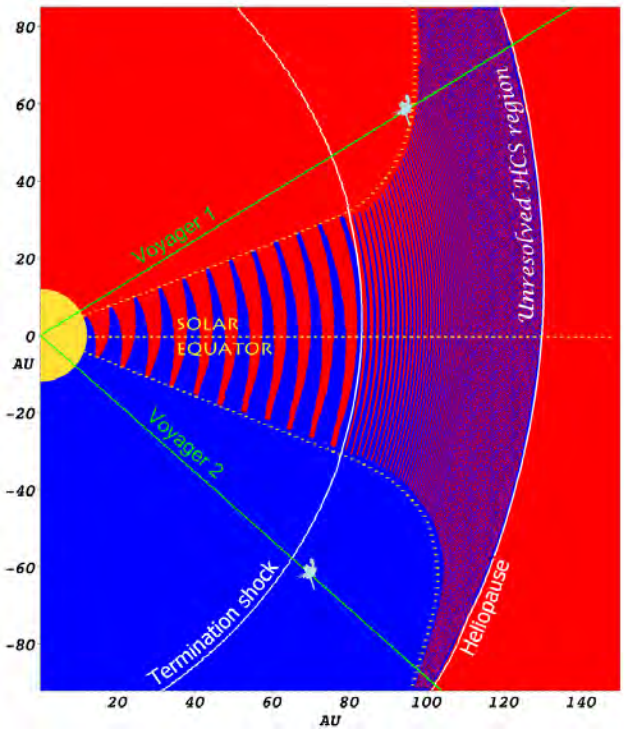


FIGURE 1. The HCS in the heliospheric slice made by the V1/V2 plane.

The Voyager crossings of the TS provided the first concrete information on the scale size and the shape of

the heliosphere. Voyager 1, in the northern hemisphere of the heliosphere, crossed the TS at 94 Astronomical Units (AU) while V2, in the southern hemisphere, crossed it at 84 AU. Based on these TS distances and model predictions, the heliopause (HP) and LISM are probably 30-50 AU further out. The asymmetry in the TS crossing distances verifies that the southern hemisphere of the heliosphere is pushed inward, probably by the interstellar magnetic field.

MISSION PROPOSAL

The purpose of this paper is to emphasize the scientific progress to be made by continuation of the Voyager mission. The Voyagers are making the first measurements in the heliosheath and are likely to make the first in situ measurements of the LISM. Although the uncertainties in the HP position are large, models suggest that the Voyager spacecraft have a good chance of reaching this boundary in their operational lifetimes. The observed asymmetry in the heliosphere may allow V1 and V2 to cross the HP at roughly the same time and provide simultaneous observations of the LISM.

The Voyager trajectories are shown in Figure 2. Distance in AU is plotted vs. heliolatitude and times are marked on the trajectory traces. Also marked are the locations of the V1 and V2 TS crossings and the V1 and V2 HP crossings based on model results. Both the V1 and V2 HP crossings are likely to occur between 2015 and 2020. The spacecraft have sufficient power to operate all instruments until at least 2016; after this time, power-sharing can extend the useful life of the spacecraft beyond 2023, well beyond the predicted HP location.

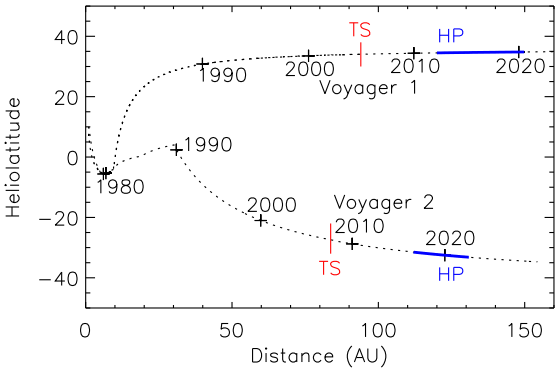


FIGURE 2. Trajectories of the Voyager spacecraft.

On their way to the LISM, the Voyager spacecraft are exploring an entirely new region, the heliosheath. The data show that this region of subsonic flow is very active, with large fluctuations in the plasma and magnetic field over times scales of hours to days. These variations likely result both from structures entrained in the solar wind and from those generated by the motions of the TS and possibly also the HP. Large fluctuations in MeV particle fluxes are observed but not yet understood. This region contributes to the modulation of the galactic cosmic rays (GCRs) [McDonald et al., 2002]. Unanticipated attributes are being discovered, such as much lower than predicted plasma temperatures, very different plasma speeds at V1 and V2, and an increase of anomalous cosmic ray (ACR) intensities with distance.

OVERVIEW OF SCIENCE QUESTIONS AND GOALS

Basic questions in NASA's roadmap relevant to this proposal are

"How are plasmas and charged particles heated and accelerated?"

"How do solar wind disturbances propagate and evolve through the solar system?"

"How do the heliosphere and interstellar medium interact?"

"What is the magnetic structure of the Sun-Heliosphere system?"

The Voyagers are the only spacecraft positioned to directly observe the boundaries of the heliosphere and to directly sample the LISM. These boundaries are the largest structures in the heliosphere and allow us to study physical processes such as magnetic reconnection, particle acceleration and transport, and the interaction of the solar wind plasma with the LISM neutrals in a system with scale size ∼100 AU. The addition of the IBEX spacecraft to the heliospheric observatory has made the role of Voyager even more critical. IBEX measures properties of heliospheric boundary regions by observing energetic neutral atoms. These observations have provided a global picture of the Sun-LISM interaction which is complementary to the simultaneous Voyager in situ measurements. Voyager directly measures the ions that are the neutral source population. Knowledge of the ENA source is critical for understanding the IBEX data. The global picture obtained by IBEX will be of great help for understanding the in situ observations and vice versa.

Particle Acceleration and ACRs

Particle acceleration and transport are uniquely addressed by the Voyager spacecraft. The pre-Voyager paradigm that ACRs were pickup ions accelerated at the TS was challenged by the observation that the ACR source was not observed at the V1 or V2 TS

crossings. The Voyagers did discover an energetic particle population with energies from tens of keV to a few MeV which is accelerated at the TS called termination shock particles (TSPs), but we still do not know the source of the ACRs. We note that the TSP spectra of H at V1 and V2 are similar, indicating that the acceleration conditions on the shock at locations 120 AU apart are similar.

At least five different theoretical models for the origin of ACRs have been advanced. McComas and Schwadron [2006] suggest that the blunt shape of the termination shock combined with a classical Parker spiral magnetic field causes acceleration to be more efficient in the flanks of the heliosphere. This hypothesis still relies on the termination shock as the accelerator of the ACRs, but not at the locations where V1 and V2 crossed the TS. Variants of this idea, that ACR acceleration occurs primarily in the tail region of the TS or at hot spots along the TS, were also proposed by Kota [2008], Kota and Jokipii [2008], and Jokipii and Kota [2008].

Fisk and Gloeckler [2009] applied their combined theoretical and observational work related to suprathermal ions having phase space densities f that closely match a velocity power law spectrum with a -5 spectral index, i.e., $f \propto V^{-5}$ [Gloeckler et al., 2008] to ACR acceleration [Gloeckler et al., 2009; Fisk and Gloeckler, 2009]. Their theory uses a particular type of stochastic acceleration in the heliosheath and makes the prediction that the large majority of ACRs are accelerated near the heliopause, because that is where the pre-accelerated particles spend most of their time.

Two other papers discuss mechanisms which accelerate ACRs preferentially near the heliopause [Drake et al., 2010; Lazarian and Opher, 2009]. Lazarian and Opher [2009] suggest that fast magnetic reconnection is responsible for ACR acceleration. The ACRs gain energy at the expense of the magnetic energy that is lost during the field annihilation that takes place at magnetic field reversals. These reversals take place throughout the heliosheath, and in the solar wind as well, but close to the heliopause where the reversals become squeezed together the conditions for fast reconnection are improved.

Drake et al. [2010] ultimately rely on the same global magnetic field reversals to provide the energy for their ACR acceleration mechanism. Their theory includes self-consistent acceleration of energetic particles within thinning current sheets where collisionless magnetic reconnection takes place. The acceleration of energetic particles arises from the contraction of magnetic islands and is a first order Fermi acceleration process, but occurs near the heliopause rather than at the TS.

The Voyager measurements as they approach the HP will be crucial to discriminate amongst these theories or to provide evidence for an entirely different explanation. We can separate the effects of temporal changes and radial gradients by comparing the observed intensity gradient between the two Voyagers with the observed change at each spacecraft as they move outward at >3 AU/year.

Global heliosphere and ENAs

Ion-neutral interactions are a key to understanding Voyager observations and the interaction of the heliosphere with the interstellar medium. Voyager first revealed the importance of the pickup ions for balancing the pressure in solar wind structures and the significant slowing of the solar wind from pickup of the ISM neutrals. At the TS, most of the solar wind flow energy went into heating these pickup ions. IBEX is now measuring the energetic neutral atoms (ENAs) generated by charge exchange in the heliosheath, giving a global view of the heliosphere. The ion distributions observed by the Voyager spacecraft are crucial for interpreting the IBEX results and the IBEX results allow the Voyager results to be put into a global context.

Our understanding of the way in which the interstellar plasma, neutrals, and magnetic field control the global structure of the heliosheath is rapidly evolving. The IBEX mission and the independent measurements from the INCA imager on Cassini showed that the ENA flux was dominated by a near circular ribbon. Figure 2 shows the emission detected by IBEX and the V1 and V2 locations [McComas et al., 2009]. The first two IBEX all-sky images revealed that dramatic temporal changes in prominent ENA features occur on time scales of only six months. These changes in the ENA features must be related to temporal changes in the 0.4-6 keV protons in the inner heliosheath or changes in the neutrals formed in the supersonic solar wind because of the long transit times of the ENAs to 1 AU ($\sim$1 year for 1 keV H). These regions of space are being observed in situ by Voyager.

These evolving populations are related to the temporal changes observed in the thermal and non-thermal protons measured by Voyager, because all three components (thermal, non-thermal pickup ions, and energetic ions) are required to satisfy the generalized Rankine-Hugoniot conditions for the conservation of mass, momentum, and energy across the termination shock. Furthermore, the Voyager measurements have demonstrated that temporal changes in the total ion pressure can be identified directly and causally with reconfigurations of the global velocity dependence of the solar wind emanating from changing patterns in the polar coronal holes on the Sun. The Voyagers play an invaluable role in establishing "ground truth" at two points within the heliosheath. The IBEX and

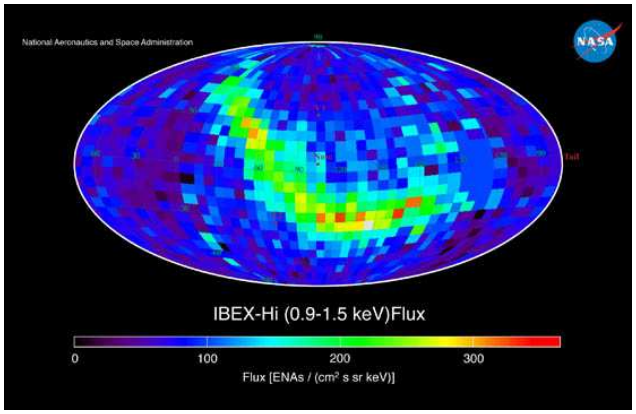


FIGURE 3. Map of ∼1.1 keV H emission observed by IBEX. The locations of V1, V2, and the nose of the heliosphere are also shown.

Cassini/INCA ENA images integrate the suprathermal proton intensities (weighted by the neutral H-atom densities) along lines of sight (LOS) many tens of AU in length. Only the Voyager spacecraft can provide guidance as to the distribution of the energetic protons along these ENA lines of sight.

The IBEX results in Figure 2 provide a new view of the heliosphere; the ribbon region may represent the symmetry plane for the higher pressure regions and thus flow may be away from the ribbon, not the nose, which could better fit the data. Alternatively, the ribbon may be due to a line-of-sight selection effect for secondary ENAs born in the outer heliosheath from solar wind pickup ions gyrating around interstellar magnetic field lines [Heerikhuisen et al., 2010].

Solar Cycle Variations

The minimum of solar activity at the end of solar cycle 23 has finally arrived. The average magnetic field strength in the solar wind at 1 AU started to recover from the lowest annual value ever measured, 4 nT. The interval of weak magnetic fields was broad, from 2007 into 2010. Since the propagation time of the solar wind from one AU to V1 and V2 is a few years, V1 and V2 have observed the heliosheath in its "ground state" (i.e. free of major disturbances related to solar activity) since the beginning of 2008. The maximum latitudinal extent of the HCS has been below the latitude of Voyager 1 and close to the latitude of Voyager 2 since the beginning of 2008, and the HCS tilt is now decreasing. During 2011, both Voyager spacecraft will probably be above the HCS and influenced primarily by the fast solar wind. The effects of the increasing solar activity in solar cycle 24 may be seen in the heliosheath in late 2011 or soon after. For at least a few years after that one expects to observe significant effects of solar activity in the heliosheath.

Although the position of the heliopause is not known, the time for the Voyager spacecraft to move through the heliosheath will almost certainly be of order of the duration of the solar cycle. Thus, we must consider the effects of the solar cycle on the structure and dynamics of the heliosheath. We are fortunate to have the opportunity to first study the heliosheath in the simplest conditions associated with solar minimum and the onset of solar activity before we confront the additional difficulties associated with major solar activity. Since the propagation time of the solar wind from one AU to V1 and V2 is a few years, V1 and V2 are still observing solar minimum conditions and in 2011 both spacecraft should be above the latitude of the HCS and observe fast polar coronal hole solar wind. The effects of the increasing solar activity in solar cycle 24 may be seen in the heliosheath in late 2012 or soon after. For at least a few years after that one expects to observe significant effects of solar activity in the heliosheath, including ICMEs and MIRs..

The Solar Wind in the Heliosheath

The plasma experiment on Voyager 2 (V2) is providing the first measurements of the thermal plasma in the heliosheath. The average speed was about 150 km/s until 2009.5 and then started to decrease. The density fell by a factor of 2 between the TS crossing and 2008.7 and has remained near 0.001 cm^{-3} through 2010. The density variations are smaller after 2008.7 than near the TS crossing. The temperature also decreased from the TS to 2008.5, then remained constant at about 50,000 K, roughly 1/2 the post-TS value.

The Voyager MAG team reports that the average magnetic field strength is 0.11 nT with a standard deviation of 0.05 nT. Models of the heliosheath predict that the plasma speed will decrease with distance which will cause B to increase as V1 and V2 approach the heliopause [Parker, 1961, Axford, 1972; Nerney et al., 1991; 1993; Pauls and Zank, 1996; 1997]. In particular, the radial speed must go to zero at the nose of the heliopause. These stronger magnetic pressures might in turn modify the flow by driving an azimuthal (field-aligned) flow of plasma away from the nose.

One of the surprises observed by V1 in the heliosheath is the large-amplitude fluctuations in B at scales of several hours down to minutes. This variability at small scales is often referred to as "turbulence" owing to its complexity. However, basic turbulence properties such as the inverse cascade and the spectral structure and exponents have not been established for the heliosheath turbulence owing to the large daily data gaps.

Complicating this picture is the large latitudinal speed gradient associated with solar minimum, with high-speed, low-density flows at higher latitudes. We think that V2 was in the transition region between fast polar coronal hole flow and slow equatorial solar wind beginning just before the TS crossing [Richardson and Wang, 2010] and that corotating interaction regions (CIRs) were driving the variability in this region [Roelof et al., 2010], although TS motions may also play a role. This hypothesis explains a 30% decrease in solar wind flux observed before the TS and could explain differences among V1 flows, V2 flows, and model results. Ulysses observed the transition from low to high-speed flows at latitude 28°S in 2008, so V2 (now at 29°S) should be near this transition region. If V2 were moving into regions of higher-speed flow, the constant V_R observed could result from a balance between the latitudinal speed increase and the expected slowdown across the heliosheath. Models which incorporate time dependent changes in the solar wind source and HCS tilt can reproduce the constant V_R observed by V2 because of this effect [Pogorelov et al., 2010].

The 3D MHD simulations of Linde et al. [1998] and Opher et al. [2003; 2004] predict that magnetic ridges form near the heliopause in conjunction with a slowing of the plasma flow. If these ridges were present, they would affect the dynamics of the flow and provide scattering sites for particle acceleration. The Voyager spacecraft will observe whether these magnetic ridges exist and if they were associated with a deceleration of the heliosheath flow.

A puzzling feature of the LECP 0.08-0.14 MeV ion data in the heliosheath is that intensity variations measured at V1 and V2 are very different. Although V1 and V2 intensity levels are comparable, the time scales of intensity fluctuations are not. V2 ion intensities and partial pressures have had continuous quasi-periodic (12-35 day) variations with amplitudes from 5-25% since crossing the TS [Decker et al., 2009; Roelof et al., 2010]. These variations were not seen at V1. These variations may extend to the core of the heliosheath pickup ion population (lower in energy by a factor of $\sim$10 relative to the lowest measured energy of 28 keV) which carries the bulk of the plasma pressure.

These quasi-recurrent heliosheath ion variations could be produced by solar wind variations causing enhanced TS injection or acceleration, by re-acceleration of low-intensity suprathermal seed ions embedded in remnant CIRs at the TS, or some combination thereof. Populations of energetic ions and electrons accelerated or trapped when CIRs were closer to the sun may provide a quasi-recurrent seed population that can be re-accelerated at the TS and in the heliosheath [Rice 2000, Decker et al., 2001]. The quasi-periodic variations may be driven by large-scale, quasi-recurrent pressure variations in the solar wind at the heliolatitude (29°S) of V2. If the quasi-periodic V2 ion intensity variations were driven by solar wind pressure variations at 29°S, then the likely sources are high-speed solar wind streams from the equatorward extensions of polar coronal holes [Nolte et al., 1976]. The resulting stream-stream interactions produce pressure enhancements. Propagation of the CIRs and evolved merged CIRs out to 100 AU is a complex process [Burlaga et al., 2003a;b]. From an observational point of view, the existence of V2 ion intensity variations may provide evidence that quasi-recurrent solar wind pressure variations disturb the TS, even though velocity variations in the solar wind have damped out.

Future Science

The recent senior review produced a listing of science questions we expect to answer over the next decade if these programs continue. These are listed below.

• Discover how the heliosheath plasma and flows evolve as V2 moves through the heliosheath. We expect

a) continued decrease of the radial speed and increased flow deflection as V2 nears the HP

b) A decrease in radial flux due to plasma moving tailward; the density may decrease significantly if a plasma depletion layer were observed ahead of the HP.

• Observe the effect of solar cycle changes in the heliosheath; collaborate with modellers to understand how the solar minimum conditions affect the observed speed profiles.

• Compare these data with model predictions to understand the flow pattern in the heliosheath and the direction of the interstellar magnetic field.

• Compare Voyager in situ data with the global IBEX ENA observations to determine the TS and heliosphere shape.

• Examine the characteristics of the B in the heliosheath under solar minimum conditions when V2 and V1 are measuring flows from the polar coronal hole regions.

• Investigate the radial gradients of B in the heliosheath.

• Determine if magnetic ridges form in the heliosheath.

• Describe the large-scale fluctuations of B in the heliosheath and how they change with the solar cycle.

• Compare model predictions for the location of the HCS with V1 and V2 observations of HCS crossings.

• Study how the sector structure and boundaries change as solar activity and the HCS tilt increase.

• Document how the suprathermal tails change with distance from the TS along two widely separated paths through the heliosheath.

• Compare the ENA-derived pickup ion and low-energy tail distributions at the V1 and V2 locations to pro-

vide information on the spatial variations of pickup ion pressure and the thickness of the heliosheath.
• Determine how different solar wind flows are modified by the TS and by propagation within the heliosheath.
• Determine the effects of extensions of the polar coronal holes on flows, magnetic fields, and energetic particles in the heliosheath.
• Determine the nature and origins of fluctuations of the intensities of energetic particles as a function of changing solar activity
• Observe the evolution of the energy spectra of ACRs at V1 and V2, the intensity gradients between V1 and V2, and the gradients at each spacecraft as they move deeper into the heliosheath, and compare with models for the ACR source.
• Compare ACR spectra to estimate the rigidity dependence of the mean free pathlength in the heliosheath.
• Determine the composition of ACRs from measurements of their source spectra.
• Observe the effects of the transition from solar minimum to the onset of solar maximum on particle intensities and modulation in the heliosheath.
• Determine the nature, location, and frequency of magnetic reconnection in the heliosheath and near the heliopause and discover if it contributes to ACR acceleration.
• Determine whether fast magnetic reconnection plays a role in the production of radio emission by looking for reconnection signatures near the HP.
• Find a self-consistent explanation of the IBEX ribbon, the H deflection, and all the asymmetries detected by the Voyager spacecraft.
• When V1 and V2 cross the HP, directly measure the in situ local interstellar magnetic field.
• When V2 crosses the HP, observe the LISM plasma.
• When V1 and V2 cross the HP, measure the unmodulated low-energy GCR spectrum.
• Observe the HP locations in the V1 and V2 directions and determine the asymmetry.

COSTS

Continuation of the Voyager data production and analysis with sufficient funds do science analysis would cost roughly $5 million/year.

SUMMARY

The Voyager spacecraft continue their epic journey of discovery, traveling through a vast unknown region of our heliosphere on their way to the interstellar medium. Both Voyagers are now traversing the heliosheath, with the first crossings of the HP and the first in situ observations of the interstellar medium still to come. These encounters could answer many basic, long-standing questions about the plasma and magnetic properties of the LISM, the nature of the TS and its role in the acceleration of the ACRs, the role of the heliosheath in GCR modulation, the spectra of low-energy interstellar GCRs, and the source and location of the heliospheric radio emissions.

Exploratory missions such as Voyager provide key tests of physical theories and also provide observational surprises which often lead to major advances in physical understanding. The continued ACR modulation in the heliosheath, the pre-shock slowdown and lack of heating of the thermal plasma at the TS, the low shock strength of the TS, and increases in the 6-14 MeV galactic electron intensities in the heliosheath are a few examples of observations forcing revisions to long-term hypotheses on particle acceleration. The combination of the IBEX ENA mapping and the Voyager in situ observations should lead to many advances on these fronts as well as to the understanding of the heliospheric-LISM interaction in a global sense.

The longevity of the Voyagers makes them ideal platforms for studying solar cycle variations in the heliosheath. Their distance make them ideal for studying the evolution of the solar wind, shocks, and cosmic rays. The interpretation of Voyager data is greatly enhanced by the ability to compare with data from Earth-orbiting spacecraft (Wind, ACE, SAMPEX, STEREO), Ulysses, and IBEX. These data make deconvolution of solar cycle, distance, and latitude effects possible. To further this intercomparison of data sets and to provide opportunity for the community to provide new insight into these observations, we strongly endorse Guest Investigator and Theory programs focusing on the outer heliosphere. Theory and multi-spacecraft comparisons are needed to provide the best understanding of the data Voyager provides. The Voyager Interstellar Mission provides a unique chance to study both temporal and spatial variations in the outer regions of our heliosphere and beyond and we should fully capitalize on this opportunity.

In summary, the Voyagers are invaluable and irreplaceable scientific resources for developing our understanding of this most recently discovered and novel plasma regime—the heliosheath and its transition to the interstellar medium.

8. REFERENCES

Axford, W.I. (1972) in Solar Wind, edited by C.P. Sonett, J. Coleman, P. J., and J.M. Wilcox, pp. 609, NASA Spec. Publ.

Burlaga, L. F., et al. (2003a) Astrophys. J., 590, 554 .

Burlaga, L. F., et al. (2003b) J. Geophys. Res., 108, 8028.

Decker, R. B., et al. (2001) Proc. COSPAR, Pergamon, New York, 11, 321.

Decker, R. B., et al. (2009) Shock Waves in Space and Astrophysical Environments, AIP Proc., in press.

Drake, J., et al. (2010) Astrophys. J., 709, 963.

Fisk. L. A., and G. Gloeckler (2009) Adv. Space Res. 43, 1471.

Frisch, P. C. (1996) Space Sci. Rev. 78, 213.

Frisch, P. C. et al. (2009) Space Sci. Rev., DOI 10.1007/s11214-009-9502-0.

Giacalone, J., and R. B. Decker (2010), Astrophys. J. 710, 91.

Gloeckler, G., et al. (2009) Space Sci. Rev., 143, 163.

Heerikhuisen et al. (1010), Astrophys. J. Lett., 708, L126.

Izmodenov, V., et al. (2008) Adv. Space Res. 41, 318.

Jokipii, J. R., and J. Kota (2008) in Particle Acceleration and Transport in the Heliosheath and Beyond?7th International Astrophysics Conference, G. Li et al., eds., American Institute of Physics, CP1039, 390.

Kota, J. (2008) Proc. 30th Internat., Cosmic Ray Conf., Merida, Vol. 1, 853.

Kota, J., and J. R. Jokipii (2008) in Particle Acceleration and Transport in the Heliosheath and Beyond - 7th International Astrophysics Conference, G. Li et al., eds., American Institute of Physics, CP1039, 397.

Lazarian, A., and M. Opher (2009), Astrophys. J. 703, 8.

Le Roux. J. A., and G. M. Webb (2009) Astrophys. J., 603, 534.

Linde, T. J., et al. (1998) J. Geophys. Res. 103, 1898.

McComas, D. J., and N. A. Schwadron (2006) Geophys. Res. Lett., 33, L04102.

McComas, D. J. et al. (2009) Science, 326, 959.

Nerney, S., et al. (1991) Astron. Astrophy., 250, 556.

Nerney, S., et al. (1993) J. Geophys. Res., 98, 15169.

Nolte, J. T., et al. (1976) Solar Phys., 46, 303.

Opher, M., et al. (2003) Astrophys. J., 591, L61.

Opher, M., et al. (2004) Astrophys. J. , 611, 575.

Parker, E. N. (1961) Astrophys. J., 134, 20.

Pauls, H. L. and G. P. Zank, (1996) J. Geophys. Res., 101, 17081.

Pauls, H. L. and G. P. Zank, (1997) J. Geophys. Res., 102, 19779.

Pogorelov et al. (2010), in Proc. Solar Wind-12.

Richardson, J. D, and Wang, C. (2010) Astrophys. J. Lett, 711, L44-L47.

Rice, W. K., et al. (2000) Geophys. Res. Lett., 27, 509.

Roelof, E. C., et al. (2010) Solar Wind 12, AIP Proc., in press

The Long Wavelength Array (LWA): A Large HF/VHF Array for Solar Physics, Ionospheric Science, and Solar Radar

A Ground-Based Instrument Paper for the 2010 NRC Decadal Survey of Solar and Space Physics

Lee J Rickard, Joseph Craig, Gregory B. Taylor, Steven Tremblay, and Christopher Watts
University of New Mexico, Albuquerque, NM 87131

Namir E. Kassim, Tracy Clarke, Clayton Coker, Kenneth Dymond, Joseph Helmboldt, Brian C. Hicks, Paul Ray, and Kenneth P. Stewart
Naval Research Laboratory, Washington, DC 20375

Jacob M. Hartman
NRC Research Associate, in residence at Naval Research Laboratory

Stephen M. White
AFRL, Kirtland AFB, Albuquerque, NM 87117

Paul Rodriguez
Consultant, Washington, DC 20375

Steve Ellingson and Chris Wolfe
Virginia Tech, Blacksburg, VA 24060

Robert Navarro and Joseph Lazio
NASA Jet Propulsion Laboratory, Pasadena, CA 91109

Charles Cormier and Van Romero
New Mexico Institute of Mining & Technology, Socorro, NM 87801

Fredrick Jenet
University of Texas, Brownsville, TX 78520

and the LWA Consortium

Executive Summary

The Long Wavelength Array (LWA), currently under construction in New Mexico, will be an imaging HF/VHF interferometer providing a new approach for studying the Sun-Earth environment from the surface of the sun through the Earth's ionosphere [1]. The LWA will be a powerful tool for solar physics and space weather investigations, through its ability to characterize a diverse range of low-frequency, solar-related emissions, thereby increasing our understanding of particle acceleration and shocks in the solar atmosphere along with their impact on the Sun-Earth environment. As a passive receiver the LWA will directly detect Coronal Mass Ejections (CMEs) in emission, and indirectly through the scattering of cosmic background sources. If coupled with a suitable transmitter, the LWA would be an excellent receiver for solar radar, potentially demonstrating accurate geomagnetic storm prediction from the Earth's surface. Both radar and passive receiving techniques could monitor the Sun-Earth environment during daytime as a complement to nighttime space weather remote sensing techniques.

The LWA will also naturally provide a measure of small-scale spatial and temporal ionospheric structure, a prerequisite for accurate calibration and imaging of solar and space weather phenomena. As a sensitive monitor of differences in total electron content (TEC) through the ionosphere, the LWA will provide an unprecedented characterization of ionospheric turbulence and waves, capable of testing predictions of global ionospheric models with an aim towards improving their accuracy through input to physics-based models. As a fully digital, multi-beaming instrument, the LWA can monitor the Sun daily with a dedicated solar beam, while simultaneously pursuing ionospheric and astrophysics science programs both day and night.

The planned LWA will consist of 53 phased-array "stations," each consisting of 256 pairs of crossed-dipole antennas, operating with Galactic noise-limited sensitivity over the frequency range 20–80 MHz. The stations will be distributed over the state of New Mexico, with maximum baselines (distances between stations) of up to 400 km, and nominally 16 stations in a 'core' within the central 10 km. Beam-forming (up to 4 simultaneous beams) will be done at the individual stations with the data then transmitted to a central correlator for image formation. The LWA will image wide fields of view with sufficient diversity of baselines to study both compact and complex sources in interferometric mode. The planned sensitivity in each beam will be a few mJy in 8 hours with a resolution of 8" to 2" (20 to 80 MHz). With a user-oriented operational concept and the goal to help stimulate University-based radio astronomy and hands-on student training, the LWA project embodies a 'science every step of the way' approach.

The first LWA station, LWA-1, will be complete at the start of 2011. LWA-1 will form four independent (in both frequency and pointing) beams on the sky, with instantaneous bandwidths of 8 MHz per beam, spectral resolutions down to 100 Hz, and temporal resolutions to 0.1 ms. Its homodyne receiver architecture will allow direct beam-formation over the entire LWA bandwidth. Operating as a fully electronic, phased array, the beams can be cycled rapidly among many sources on millisecond timescales. The science plan for the coming year includes several solar and ionospheric studies.

1. Introduction

The Long Wavelength Array (LWA, [2]) will be optimized for exploring the lowest electromagnetic window (~10 – 88 MHz) available for ground-based interferometric imaging, between the ionospheric cutoff and the FM bands. The LWA follows the heritage of the 74 MHz system developed by the Naval Research Laboratory and the National Radio Astronomy Observatory (NRAO) for the Very Large Array (VLA) radio-telescope [3]. That system was the first connected element, imaging interferometer to demonstrate sub-arcminute resolution imaging below 100 MHz. This breakthrough was achieved based on innovations in HF/VHF adaptive optics required to accurately measure and model small scale (temporal and spatial) ionospheric-induced phase variations that become severe on antenna separations (or interferometer baselines) > 5 km. The LWA is designed to extend the maximum baselines

available with the original 74 MHz VLA system (~35 km) by approximately one order of magnitude, and the available sensitivity by more than two orders of magnitude.

In the following sections we provide (§§ 2 and 3) a discussion of solar physics and heliospheric physics applications of the LWA, respectively; (§4) a review of LWA ionospheric remote sensing applications, afforded naturally through its requirement to calibrate against thousands of natural cosmic background sources; and (§5) a consideration of the LWA as a radar imaging receiver for detecting Earthward-bound Coronal Mass Ejections (CMEs) for geomagnetic storm prediction. In addition, we provide (§6) a brief technical overview of the LWA and its current construction status; and (§7) a summary of programmatic issues.

2. Solar Physics

The LWA frequency range, optimized for 20-80 MHz, is the domain of the classic solar radio burst types: Type II bursts, generated by shocks in the corona, with their characteristic fundamental-harmonic split-band structure; Type III bursts, drifting rapidly in frequency as the electron beams that generate them propagate into the solar wind; and Type IV bursts, from trapped electrons producing broadband emission in long-duration flares. Examples of Type II and IV emission are shown in Fig. 1; on a plot like this a Type III burst appears as a brief nearly vertical feature. With its sensitivity, flexible frequency coverage and multibeam capability, the LWA will be a powerful tool for the study of these bursts. Each of them is believed to emit by the mechanism of plasma radiation, *i.e.*, the conversion of electrostatic Langmuir waves into electromagnetic waves at the fundamental and second harmonic of the electron plasma frequency 9000 $n_e^{0.5}$ Hz, where n_e is the electron density (cm^{-3}). The frequency of plasma emission thus reveals the density in the source. Some examples of possible solar science topics are described below.

An important advance provided by the LWA is the ability to image at many frequencies across a significant bandwidth simultaneously. Previous imaging of, *e.g.*, Type II bursts has been at fixed frequencies, and one then only has access to the phenomena that happen to be drifting through the observing frequency at any given instant. This is a disadvantage for Type II bursts because, as fig. 1 shows, they typically consist of fundamental and harmonic bands, each of which is split. It is believed that the splitting may be due to emission coming from the upstream and downstream regions of a shock, with the differing densities in the two locations leading to different plasma emission frequencies. Previously it has never been possible to make an image of the two split bands simultaneously: the LWA, with its large instantaneous bandwidth and the ability to image anywhere in that bandwidth, will make that possible and thus provide a test of models that use the splitting to infer the shock Mach number.

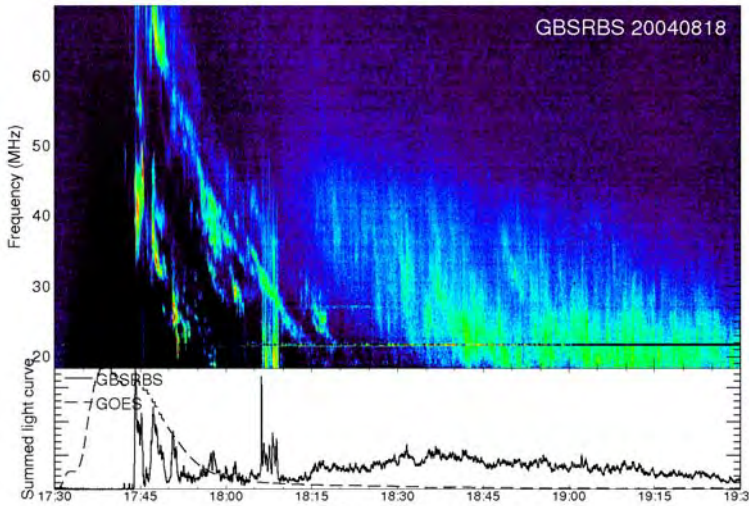


Figure 1: Radio emission from a solar flare in the LWA frequency range, observed with the Green Bank Solar Radio Burst Spectrometer. The upper panel shows a dynamic spectrum (frequency vs. time plot) with harmonic bands of a Type II burst dominating from 17:45-18:00 UT, and the broadband continuum of a Type IV burst drifting gradually lower in frequency from 18:20-19:30 UT.

Similarly, fig. 1 shows that Type IV bursts exhibit strongly modulated emission. It is suggested that such emission comes from plasma emission that is broadband due to the range of densities experienced by

electrons trapped in post-flare loops as they propagate along the length of the loops. The modulations then represent simultaneous injections of additional nonthermal electrons onto the loops. An alternative explanation is that the Type IV emission is composed of individual electron beams that drift rapidly over a finite height range. LWA-1, with its combination of exceptional sensitivity and time resolution, will be capable of looking for extreme drift rates in Type IV substructure across 8 or 16 MHz of bandwidth, while the full LWA system will be able to measure changes in source position with frequency and trace out the locus of the burst source.

In events in which the bright plasma emission sources are absent in some regions of the LWA's frequency coverage, it should be possible to detect CMEs directly via the synchrotron emission from electrons accelerated in the CME-driven shock above the solar corona. This has been demonstrated with Nançay Radioheliograph observations at 164 MHz by [4]. The emission by these electrons is generally weak relative to the bright coherent plasma emission of the classic large burst types, and does not show up on a dynamic spectrum such as fig. 1, but imaging reveals a quasicircular arc of emission co-located with the CME front. The spectrum of this emission close to the Sun [4] appears to be consistent with a nonthermal spectrum suppressed at low frequencies by the Razin effect. The LWA will be able to trace such emission to much greater heights above the Sun thanks to its lower frequency range; Razin suppression will be less effective at greater heights due to the lower densities in the atmosphere there, and the emission should be very bright and easily observed in the LWA frequency range provided that it is not swamped by plasma emission sources. These observations have the potential to address a large number of important questions in shock physics, including:

- What is the relative effectiveness of quasi-perpendicular and quasi-parallel shock acceleration?
- What is the magnetic field geometry of the expanding CME structure?
- How does the rate of electron acceleration vary with height and what is their energy distribution?
- What is the energy distribution of the accelerated electrons?

Related to this topic is the observation of "moving Type IV" bursts. These are huge, rare broadband disturbances in the LWA frequency range observed historically by the Culgoora Radioheliograph and (on a few occasions) by Clark Lake Radio Observatory; they cannot easily be distinguished from normal Type IV bursts from dynamic spectra alone, and there is debate as to whether their emission mechanism is plasma radiation or synchrotron. They are seen to move out to great heights above the solar surface at speeds similar to fast CMEs. However, their exact relationship to CMEs is poorly understood because there was only a brief overlap between modern space-borne coronagraph imaging of CMEs and the operation of Culgoora: in the one well-studied case, the moving Type IV burst appeared to be lagging behind the leading edge of the associated CME, lying closer to the densest material in the ejection (probably the erupting filament) [5]. The LWA will permit the observation of moving Type IV bursts again, and the study of their relationship to other erupting phenomena, for the first time in 25 years.

The LWA will also be capable of mapping the quiet solar atmosphere once sufficient stations in the core are available to provide good *u-v* coverage for a source of dimension 1° such as the Sun. Observing the variation in the structure of the solar atmosphere in the 1-3 R_{sun} height range will be an important adjunct to ongoing studies of the formation and acceleration of the solar wind, which is believed to occur in that height range. Currently a large emphasis is placed on MHD modeling and magnetic field extrapolation techniques to study these topics due to the difficulty of obtaining suitable measurements: The LWA's frequency flexibility will permit mapping of the solar atmosphere from 20 to 80 MHz and reveal changes in structure with height that will strongly constrain these studies.

3. Heliospheric Physics

The LWA will be able to employ the well-established technique of interplanetary scintillation (IPS) to track regions of increased density associated with CMEs as they propagate through the interplanetary medium. This technique is being used by several groups around the world (e.g., see http://ips.ucsd.edu/). The method is to observe the level of fluctuations ("scintillations") in the signal from a bright background radio source as it passes through the solar wind: the fluctuations are produced by scattering and diffraction of the signal in the density fluctuations of the solar wind. The level of density fluctuations is generally found to be well correlated with the absolute density; a turbulent feature such as a CME-driven shock may well increase the level of scattering further.

IPS radio systems (e.g., STELab in Japan and Ooty in India, both at 327 MHz) generally employ large telescopes operated as single frequency 'transit instruments', where a set of strong compact radio sources is observed daily as the source transits the local meridian at the observing site. The advantage that the LWA possesses is that it has no moving parts: "pointing" is achieved electronically by phasing the signals from each of the 256 elements so that, when combined, they form a beam pointing at the source of interest. With up to 4 beams and its large collecting area, the LWA can observe 4 sources simultaneously with a time resolution of order milliseconds, opening up the possibility for unique data analyses. Since the LWA can switch to a new set of sources essentially instantaneously, it will be able to monitor the level of fluctuations in the solar wind with much denser time-and-space sampling than can be achieved with current instruments, with a consequent improvement in the ability to track the movement of discrete features in the solar wind such as CMEs. Furthermore, the LWA's ability to make measurements over a significant frequency range simultaneously allows the use of frequency-correlation techniques that can yield velocity information even with a single station.

Initial work on IPS measurements can be carried out with LWA-1 using the classical IPS and frequency-correlation methods; once more stations are available, it will be possible both to add measurements of solar wind and density-structure velocities using the multi-station technique employed currently by STELab, and to use a denser array of scintillation sources thanks to the smaller beam of the multi-station array.

The ability to measure polarization adds an extra dimension to both IPS and other observations. The polarization of the CME synchrotron emission observed by [4] will be of great interest for revealing the structure of the magnetic field in an expanding CME front. Further into the heliosphere, Faraday rotation measurements can be used to infer magnetic fields: such measurements require measuring changes in the full polarization state of background radio sources as a CME passes in front of them, and the variation of the polarization with frequency. These are difficult measurements and this is not a prime goal of the LWA, but this technique will be explored once the LWA has sufficient sensitivity and has measured the polarization characteristics of its stations.

4. Ionospheric Physics

The distribution of electric charge in the Earth's ionosphere perturbs the refractive index for radio waves passing through it, and these refractive index fluctuations change the phase velocity of the rays and thus introduce delays between different paths. This effect makes a multi-element low frequency array extremely sensitive to small *differences* in total electron content (ΔTEC) through the ionosphere, as defined by the separation of interferometer pairs sampling it [1,2]. Typical ionospheric remote sensing techniques sample ΔTEC at the 0.1 TECU level (1 TECU = 10^{12} electrons*cm^{-2}) and on spatial scales of $\geq$ 100 km. The VLA can measure ΔTEC at the milli-TECU level and on spatial scales approaching 100 m. The VLA has been used for 330 MHz-based studies of travelling ionospheric disturbances (TIDs) [6], while more recent studies have extended measurements to lower frequencies [7].

The LWA will be a far more powerful ionospheric probe than the VLA, coupling the technical advantages of much greater collecting area, spatial sampling, and broadband low frequency spectral

coverage with electronic multi-beaming and continuous monitoring. While on the one hand the ionosphere will be a significant problem for the LWA astrophysical observations, this "problem" offers a tremendous opportunity for ionospheric science studies and serves as a major justification for the instrument. The electron content of the ionosphere above the LWA, varying with position, height and time, introduces phase delays in the signal paths from the sky to the telescope. These delays show up to lowest order as a simple position shift of a source on the sky, and to second order as a smearing of the source, but over the long baselines planned eventually for the LWA this is an inadequate description of ionospheric effects: they will be different in different parts of the sky, and for the same source viewed from different stations. The LWA can therefore provide unparalleled measurements of TEC fluctuations simultaneously from small (~1 km; ~milli-TECU) to large (~400 km; ~TECU) scales, a combination of synoptic and detailed views that no other instrument can approach. Furthermore, it will do it as a matter of course, as a pre-requisite for accurate calibration and imaging of the Sun and other cosmic radio sources. And, as a real-time ionospheric monitor, the LWA support tests of the predictions of global ionospheric models, with an aim towards improving their accuracy through input to physics-based models.

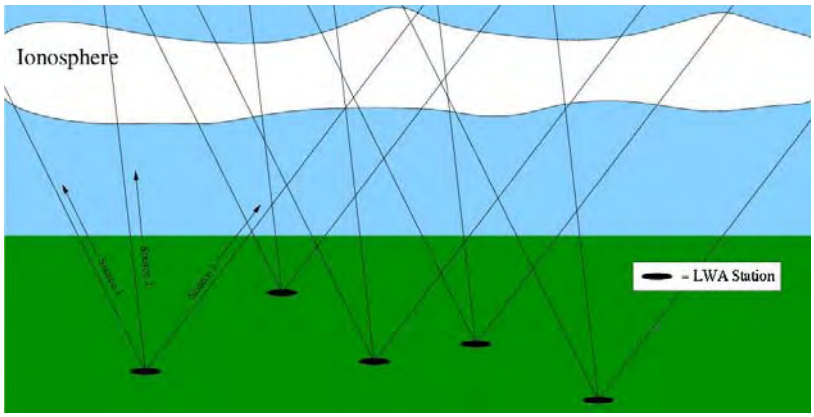


Figure 2: A schematic illustration of LWA ionospheric tomography. With electronic steering, multiple beams and high sensitivity, the LWA can image 100 cosmic background sources within 5 seconds. Here, 5 stations and 3 sources yield 15 lines of sight. 53 LWA stations will provide 5,300 pierce point samplings through the ionosphere.

From a single LWA station, the ionosphere can be studied by rapidly cycling through a set of strong compact sources that cover the sky down to some elevation and measuring the delay for each line of sight by established methods (fig. 2). As for the IPS application, the fact that the LWA has multiple beams and no moving parts, so that repointing the beams can be achieved almost instantaneously, allows such studies to be carried out with unprecedented time sampling. Ionospheric perturbations such as "sudden ionospheric disturbances" can be detected as position-varying phase shifts depending on the electron content along each line of sight. Mapping these changes in electron content with time allows one to see the movement of features overhead.

With multiple LWA stations, measuring the electron content towards the same source in the sky from different stations introduces 3-dimensional information: the height distribution of the electron content can be studied from the differences of different lines of sight (fig. 2). The baseline range of the full LWA, 400 km, is ideal for sampling the full height range of the ionosphere by this tomographic technique, potentially providing thousands of measurements through the ionosphere on timescales of seconds or less. The combination of measuring small and large spatial scales simultaneously, coupled with three-dimensional information and high time resolution, means that the LWA has the potential to improve greatly our understanding of the dynamics of the fine-scale structure of the ionosphere.

Ionospheric studies with just a single LWA station can contribute indirectly to magnetospheric physics. Recent experimental evidence shows that wave-particle interactions are significant factors in both beta acceleration and beta loss in the natural radiation belts [8]. Among the natural processes that are now known to induce electron precipitation are lightning, through generation of whistler-mode waves [9], earthquakes, by generation of ultra-low-frequency (ULF) and very-low-frequency (VLF) waves [10], and the direct distortion of the magnetosphere by solar flares and coronal mass ejections (CMEs) [11].

The precipitating electrons cause localized ionization and changes of ionospheric plasma conductivity [12]. This results in enhanced absorption of the cosmic high-frequency (typically 10 – 60

MHz) radio background, originating primarily in the D-region. Such enhanced, or anomalous, absorption measurements have been the subject of ground-based studies since before the original International Geophysical Year [13]. The method currently used for such studies is riometry, and has remained relatively unchanged for almost as long [14]. A long time-series of observations is used to determine a 'quiet background level', which is a convolution of the system response and the (strong) Galactic background, and varies slowly on an hourly and yearly basis. Anomalous absorption measurements are detected as shorter term variations in which the observed flux drops, perhaps by tens of dB.

The standard riometric measurement is made at a single, relatively narrow bandwidth. With the LWA, however, we can provide measurements over a broad range of frequencies, in the 10 – 80 MHz range. Furthermore, the system can be calibrated so that, rather than having to derive a 'quiet background level' from the data, the absolute sky brightness against which the absorption is measured can be determined. These are both significant enhancements of the methods for studying anomalous absorptions. Additional information can be gleaned from the spatial distribution of the absorption event. To date, this has been done by using small arrays of antennas (typically 4x4 or 8x8) and analog beam formation (via Butler matrices), to generate a set of beams on the sky at fairly coarse spatial resolution.

5. Solar Radar

The LWA could potentially provide an important new capability as a receiving array for radar studies of the solar corona. The solar radar concept was first proposed as an approach to investigating the ionized structure of the solar corona [15, 16]. Among the early experiments, echoes were reported both from Arecibo [17] and El Campo [18]. The El Campo campaign extended from 1961-1969, and its published reports are the main comprehensive description existing of solar radar echoes.

The El Campo solar radar operated at 38 MHz with a total power of 500 kW and used an array of antennas of about 18,000 m^2 in size, both to transmit and to receive. The fan-shaped beam of the El Campo radar was about 1° x 6° in the NS x EW direction, allowing only one transit experiment of the sun per day to be conducted. An experiment consisted of a coded waveform transmitted toward the sun for 16 minutes (the round-trip travel time of light) followed by 16 minutes of reception. The data recorded during the reception time interval were then correlated with the transmitted code to search for signals at time delays and frequency shifts expected for a solar echo at the reflection level of 38 MHz, expected at about 1.4 R_o (where R_o is the solar radius). From those cases in which an echo could be identified at some level of statistical probability, a radar cross section of the sun was calculated. The majority of radar cross sections had values about equal to or slightly larger than the solar geometric cross section. However, over the years of the El Campo experiments, there emerged a subset of cross sections that appeared at earlier time delays. It was suggested that these "high corona echoes" might be related to dense irregularities possibly associated with shock fronts in the corona [19].

Today we recognize that these high corona echoes could have been produced by large scale coronal disturbances such as Coronal Mass Ejections, a phenomenon that was unknown to solar physicists of the 1960s, but which today is recognized as a major driver of space weather at the earth. Thus, interest in the use of solar radars was revived in the 1990s [20] when it was recognized as a possible technique for early detection of earthward-moving CMEs.

Fig. 3 illustrates the concept of using a solar radar to detect an earthward-moving CME while it is still near the sun. The leading front of the CME is expected to cause a frequency shift in the radar echo proportional to the velocity of the CME, from which the travel time to the earth can be derived. However, today there is no extant solar radar comparable to El Campo, which was de-commissioned long ago (perhaps a memorial to being ahead of its time). Several high power transmitter facilities are in use today that approximate some characteristics of a solar radar. However, because these facilities were designed and built for a different purpose, they all lack some crucial aspect needed for a true solar radar.

Among the existing facilities are Over-The-Horizon (OTH) radars, which have power levels and frequencies comparable to those of the El Campo radar, but have antenna beams that are hard-pointed at

low elevation angles and azimuths and unable to illuminate the sun except on a few days of the year. Another class is the high power ionospheric heating transmitters, which have powers in some cases exceeding the power of the El Campo radar, but operate only at the low frequencies designed to maximize interaction with the earth's ionosphere. Only at special times when the ionospheric cutoff frequency is low enough are these transmitters able to illuminate the sun. Other high power ionospheric radars have fixed pointing directions or operate at frequencies too high for solar radar. Although these facilities were not designed to be solar radars, several experiments have been done with them anyway, in configurations that are as close as possible to the El Campo radar. Among these experiments, frequencies of 9 MHz, 25 MHz, 40 MHz, and 50 MHz have been used. However, the results have been somewhat disappointing, with no evidence of a well-defined echo. The most recent solar radar experiments have been unable to clearly detect solar echoes [21].

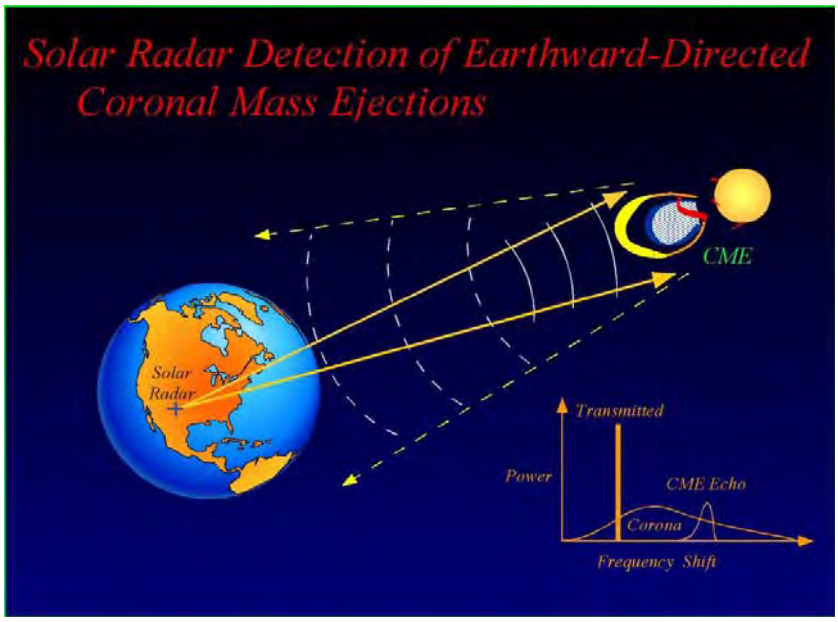


Figure 3: The concept of using solar radar to measure CME velocity and travel time to Earth.

Thus it remains an important task to understand the limitations that may have been significant in these various attempts to reproduce the El Campo solar radar experiments. It is possible that solar radar experiments, already intrinsically difficult, are not adequately addressed with existing facilities and that a true solar radar is required.

While a solar radar demands high power transmission (radiated power up to 3 MW is desirable) due to the great distance to the sun and to the natural radio noise environment of the sun, it is just as important that the receiving array be well-matched to the angular size of the target and be able to resolve spatial and temporal variations in the echo and background emissions.

The LWA will receive frequencies in the range of 20 to 80 MHz and can provide several beams to receive several frequencies simultaneously, providing a significant improvement to the receiving component of a solar radar; all previous solar radars involved only one frequency. The LWA is being designed with solar studies in mind and will use modern techniques of digital signal processing of the data. The array will have electronic phasing of antennas in the various sub-arrays (or stations) situated over a large area. The spatial extent and phasing of the LWA bring about two important characteristics needed for a solar radar: beam sizes comparable to and smaller than the angular size of the solar disk (~ 30 arc-min) and the agility to form and track one or more beams to follow the sun across the sky. Neither of these capabilities existed in previous solar radars, including those of El Campo. All previous solar radar experiments have been done in transit mode in which the sun drifted across a stationary beam. The small beam size of the LWA will allow it to receive its signals from a solid angle that includes only the sun or to utilize smaller beam widths to map structure in the echo signal across the solar disk. The tracking capability expands the solar radar experiment to several hours, allowing more experiment cycles on a given day. The basic cycle is about one-half hour, composed of 16 minutes of transmission followed by 16 minutes of reception. As mentioned above, El Campo conducted only one experiment cycle per day. The experiment cycle time corresponds to an angular displacement of the sun by about 8°, requiring a stationary beam width to be broad enough to maintain sufficient antenna gain on the sun. The beam of

the LWA phased array can be electronically steered to track the sun for several hours per day, thus making possible several experiment cycles per day. In addition, the LWA beam maximum response can be kept centered on the sun throughout the time of tracking. The additional cycles possible with the LWA can be used to help discriminate against solar radio noise that may resemble a radar echo.

6. Technical Overview

The LWA will take advantage of emerging technologies in calibration, imaging and digital processing to surpass previous instruments by several orders of magnitude in resolution and sensitivity below 100 MHz. The LWA will consist of 53 phased-array "stations," each consisting of 256 pairs of crossed-dipole antennas, operating with Galactic noise-limited sensitivity over the frequency range 20–80 MHz. The stations will be distributed over the state of New Mexico, with maximum baselines (distances between stations) of up to 400 km, and nominally 16 stations in a 'core' within the central 10 km (Fig 4). Beam-forming (with up to 4 independent – both in frequency and pointing – simultaneous beams) will be done at the individual stations, with the data then transmitted to a central correlator for image formation. Its homodyne receiver architecture will allow direct beam-formation over the entire LWA bandwidth, with instantaneous bandwidths of 8 MHz per beam, spectral resolutions down to 100 Hz, and temporal resolutions to 0.1 ms. The LWA will image wide fields of view with sufficient diversity of baselines to study both compact and complex sources in interferometric mode. The planned sensitivity in each beam will be a few mJy in 8 hours with a resolution of 8" to 2" (20 to 80 MHz).



Figure 4: Nominal LWA station locations shown in yellow; location of the VLA is shown in blue.

The first station (LWA-1) will be complete at the start of 2011. Tests with the analog system already show that there are no unanticipated problems in meeting performance expectations. It is anticipated that the design of the additional stations will be similar to that for LWA-1, as described below.



Figure 5: Aerial view of LWA-1, showing stand layout and (lower right) electronics shelter.

Receiving Elements and Station Layout: The primary receiving element is a fixed stand that incorporates two broadband, crossed, linearly-polarized dipoles. A group of 256 stands will form a roughly 100-meter diameter station, with a pseudorandom distribution (fig. 5) to prevent aliasing of the main lobe and to mitigate large sidelobes at higher frequencies where the aperture will be undersampled. The station beam can be steered to any point in the sky by adjusting the digital delays of the individual elements. Beam steering will be entirely electronic.

Each dipole (Fig. 6) feeds a balanced-input active-balun to produce a single-polarization, Galactic-noise-dominated output signal. The front end electronics (FEE; Fig. 7) uses commercial InGaP HBT MMIC amplifiers (Mini-Circuits GALI-74) in a differential configuration presenting a 100 Ω balanced load to the antenna. This is followed by a 180° hybrid, a fifth order RFI rejection filter, and an additional 12 dB gain stage (Mini-Circuits Gali-6) to produce a 50 Ω single-ended signal suitable for transmission over coaxial cable. The gain, noise temperature, and input 1-dB compression point are approximately 36 dB, 250 K, and −18 dBm respectively, roughly independent of frequency over 10–88 MHz. The FEE circuit board has been designed to enable identical signal polarization units to be assembled back-to-back, forming a dual polarization FEE. A coarse-mesh ground screen underneath the stand will reduce ground losses and stabilize the antenna impedance.



Figure 6: Antenna stand in field.

The choice of 256 stands distributed over ~ 100m is driven by the anticipated need to rapidly image several hundred sources across the sky, to calibrate for ionospheric effects within the station field of view (FWHM ~ 8°[20 MHz/ν]). It meets the sensitivity goals of the science requirements while achieving a balance with classical and sidelobe confusion for 400-km baselines over plausible integration times. It also balances cost against quality of image calibration over a broad range of frequencies and zenith angles.

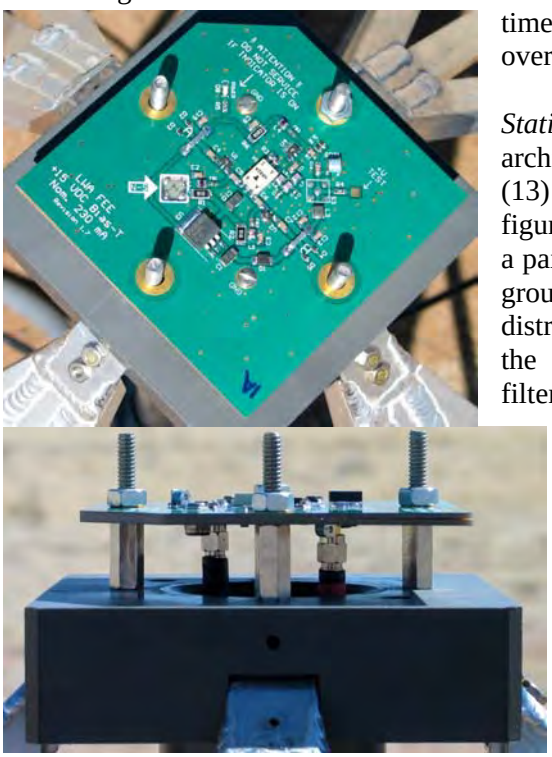


Figure 7: (Top) FEE viewed from above; (Bottom) FEE pair viewed from side.

Station Architecture and Electronics: Figures 8-9 summarize the architecture of the station; (10-12) are the basic signal path, and (13) is the common monitor and control system (MCS). In figure 5, signals from the 256 stands (STDs), each consisting of a pair of antennas (ANT) and front-end electronics (FEE) over a ground screen (GND), go by way of a common RF and power distribution network (RPD) to the signal-entry panel (SEP) of the electronics shelter (SHL). Inside, each polarization is filtered and amplified by an analog receiver (ARX) and then direct-sampled at 196 Msps by the A/D converter in figure 9. We are adopting a digitization scheme that enables us to fold the FM band (88-108 MHZ) over on itself, substantially reducing its impact. This fixes a high-frequency limit of 88 MHz and a low-frequency limit of 10 MHz.

In figure 9, the beam-forming units (BFUs) employ a time-domain delay-and-sum architecture, which allows the entire 10-88 MHz passband from each antenna to be processed as a single wideband data stream. Delays are implemented in two stages: A coarse delay is implemented using a first-in first-out (FIFO) buffer operating on the A/D output samples, followed by a finite

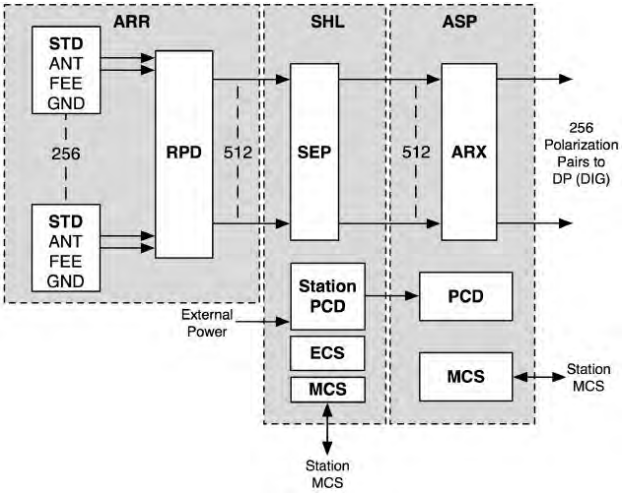


Figure 8: Architectural schematic of LWA from the antennas through the analog signal processing.

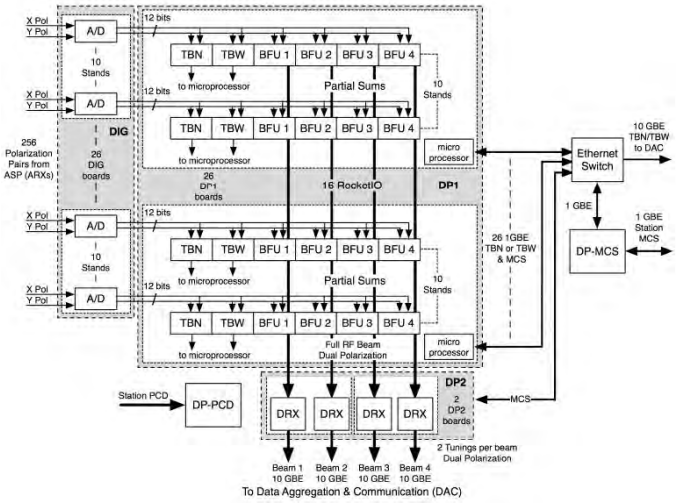


Figure 9: Architectural schematic of LWA electronics, from digitization through digital processing.

impulse response (FIR) filter for sub-sample delay. The signals are then added to the signals from other antennas processed similarly. Four dual-polarization beams of bandwidth 78 MHz, each capable of fully independent pointing over the visible sky, will be constructed in this fashion. Each full-bandwidth beam is down-converted by a digital receiver (DRX) and sent to a polyphase filter bank to channelize the beams into spectral channels. With 4096 channels, beam bandwidths up to 8 MHz and spectral resolution down to 61 Hz are achieved.

These beams will be available for various "backends" implemented at the station level, such as data recorders, wideband spectrometers, and pulsar machines. For interferometric imaging, two "tunings" will be extracted from any frequency in the 78 MHz-wide passband for each beam. These beams are the output to the LWA correlator. In addition, stations in the LWA core will output a wideband beam derived from one of the full-RF beams.

To facilitate commissioning activities, diagnostics, and certain types of science observations requiring all-sky fields of view, the station electronics will also have the capability to coherently capture and record the output of all A/Ds, where each A/D corresponds to one antenna. This will occur in two modes: the "transient buffer – wideband" (TBW) allows the raw output of the A/Ds to be collected continuously for 57 ms at a time. The "transient buffer – narrowband" (TBN) allows a single tuning of 1 kHz to 100 kHz bandwidth to be streamed continuously.



Figure 10: Analog signal processing board

Data Handling and Correlator: For the first LWA station, the data streams will be handled by a set of data recorders (MCS-DR), as shown in figure 12. The MCS-DR consists of 5 identical computers, four of which receive data streams from the digital processor (DP) corresponding to the output of each of the four station beams. The fifth computer receives the DP output data streams corresponding to TBW/TBN outputs. All computers contain internal arrays of large hard drives, to which data can be streamed at the rate received from the DP.



Figure 11: Digital processing board

The data aggregation and communication system (DAC) for the full LWA will be a Gigabit ethernet over optical fiber connecting stations to the correlator at a central location. For most stations, the output to the correlator will be both polarizations of 3 beams of 8-MHz bandwidth each, resampled to 8 bits at 1.5 times the Nyquist rate. This results in a data rate of 576 Mb/s. One of these beams will always be available to assist in measuring the ionosphere, as part of image calibration, whereas the other two can be used for simultaneous independent observing programs. Practical limitations in data transmission from the stations to the correlator constrain the immediate plan for multibeam operation to a bandwidth of 8 MHz (selectable within the tuning range). However beams will be formed at the station as "full RF"; i.e., with bandwidth equal to the tuning range. LWA core stations will also transmit a wideband beam (~56 MHz) to the correlator simultaneously with the three 8 MHz beams. This beam will be used primarily for solar science during the day and for early universe ("Dark Ages") studies at night.

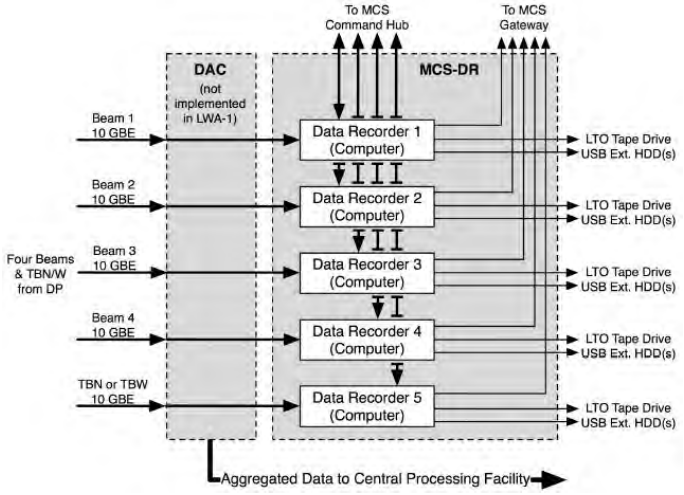


Figure 12: Architecture of LWA data recording.

The large number of high data-rate signals involved makes correlation extremely computationally-intensive, requiring dedicated equipment running in real time for the complete LWA. This is not the case for LWA-1. The early LWA is rather in the regime of VLBI – a small number of stations and a moderate bandwidth to be correlated. Thus, we need merely capture the station outputs at a central location using disk buffers, and perform correlation in software using general-purpose computers, using a system like DiFX [22]. We intend to build a small computer cluster, using 3 AMTF MTG15400Duo servers to host a software correlator for the few-station LWA. These units can easily handle the standard station data rates of 576 Mb/s, and can even cope with the full bandwidth output of a single station: 1 Gb/s for each of four beams, plus 1 Gb/s for the narrowband transient buffer.

Control Systems: The station MCS (Fig. 13) is essentially the set of computers that controls the station, and provides status information. Various subsystems including the analog and digital processors also have MCSs, which are embedded computers subordinate to the station MCS. The subsystem MCSs are implemented to facilitate modularity in the station design and to facilitate independent development of subsystems. The Scheduler is a computer whose primary function is to issue commands and receive status from other LWA subsystems. It handles tasks that are extremely time sensitive and that must be coordinated on timescales down to milliseconds. The Executive is the computer which exercises top-level control over MCS as well as the station. It is responsible for interpreting observation requests and, from these, generating the data which becomes the content of command messages issued by the Scheduler. This includes numerically intensive operations such as computation of FIR-filter coefficients. The Executive manages tasks that are moderately time-sensitive and that must be coordinated on timescales down to

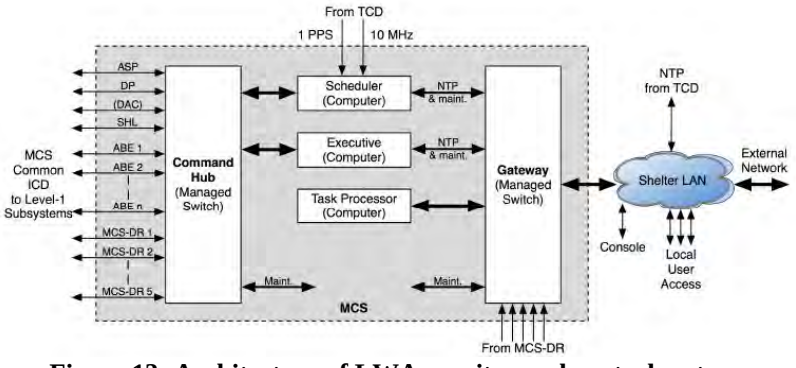


Figure 13: Architecture of LWA monitor and control system.

seconds. The Task Processor is a computer which exists primarily to host applications that are not time critical (time resolutions > 1 sec), and therefore can be offloaded to reduce the processing burden of the Executive. It is the primary interface with users, managing command line and GUI interactions. The Task Processor is also responsible for the scheduling and interpretation of internal diagnostics (either automatic or user-directed), and manages MCS-DR.

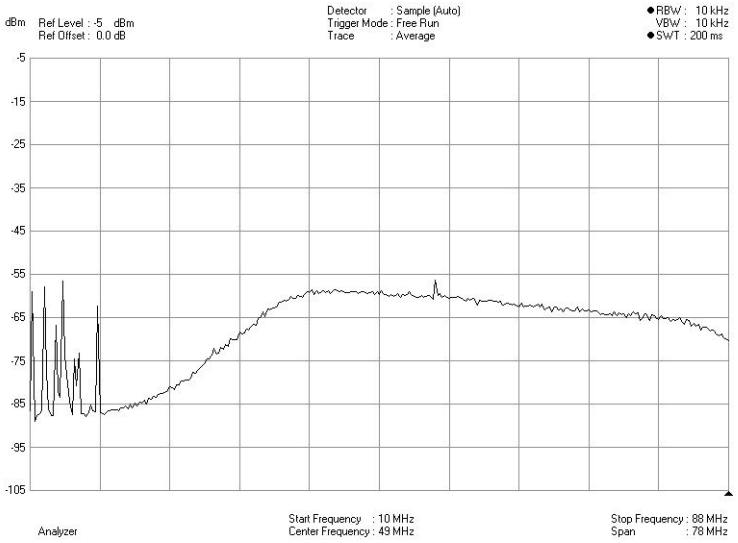


Figure 14: Spectrum acquired in the field using the LWA active antenna, FEE, and ARX, as described in the text. The gain for the lower part of the band has been reduced to minimize the impact of RFI at the lowest frequencies.

Radio Frequency Interference: A sense of the RFI environment is conveyed by figure 14, which shows the observed sky signal obtained in the field with the actual antenna, front-end electronics, and analog receiver. The result is sky-noise dominated by at least 6 dB over the range 20–80 MHz, and by at least 10 dB over 28–47 MHz. The impact of very strong RFI at the lowest part of the band has been minimized through use of a split-gain configuration of the ARX. Note that the major RFI features usually seen above 40 MHz, from analog TV signals, have now essentially vanished at the site, thanks to their move to digital TV carriers at higher frequencies.

Experience from the 74-MHz VLA system and other instruments has demonstrated that RFI from external as well as internal sources will be present at all levels throughout the spectrum. Having demonstrated that the strongest local RFI does not threaten linearity, we can be confident that the primary difficulty will be the increase in the amount of manual effort required to reduce data. A variety of solutions for automatic real-time mitigation of RFI are being considered for implementation. In the station electronics, this may include the ability to modify the responses of digital filters to suppress narrowband RFI, and pulse blanking to remove strong, bursty interference. Spatial or space-frequency nulling can be supported by the electronics architecture. For spectrometry, time-frequency blanking to resolutions of a few ms × a few kHz is supported. Other devices and backends may use additional application-specific methods, and the specific mix of techniques employed will depend on the observing mode and RFI present.

Ionospheric Calibration: The early LWA stations will have baselines of tens of km. These are comparable to baselines that have been used at the VLA at 74 MHz, so that we have reasonable expectation of being affected by the same phenomena as seen at the VLA. Ionospheric calibration will be done with a 'flickering' method (cf. §5 below), which consists of rapidly cycling through and imaging the several hundred celestial radio emitters (sources) that we predict will be isolated enough to be imaged with standard calibration techniques. Such sources can therefore function as ionospheric calibrators, because calibrating to them determines the differences in ionospheric column density (Total Electron Content differences, or ΔTEC) along the lines of sight from each LWA station towards each calibrator source. An analysis based on VLA Low-frequency Sky Survey (VLSS, [23]) fluxes and LWA antenna simulations shows that, at 74 MHz, roughly 100 calibrator sources can be observed at any single time, which is far greater than the number of instantaneously available GPS or similar satellite beacons. This process can be done within a cycle time of 6 to 7 seconds, with an average of 50 msec per source. In addition to providing calibration, it provides the essential data to generate a dynamic, all-sky model of the ionosphere rich in phenomenology.

The key issue preventing tests of this method with LWA-1 is the confusion noise limit for a single station, which is reached in ~ 1 msec. By providing 'outrigger' baselines, to antennas significantly outside LWA-1, this confusion noise can be reduced significantly.

An indication of the scale of ionospheric fluctuations has been derived from VLA 74 MHz observations of Virgo A [24]. The amplitudes and spatial extents of the apparently turbulent structures seen during the night are similar to those of tropospheric phase fluctuations seen in the visible by optical interferometers [25]. That, together with the scanty sampling of station spacings, puts the full LWA imaging more in the regime of the optical interferometer than more densely sampled systems like the VLA. In this regime, phase closure 'imaging' (really more like model-fitting) is possible on bright sources [26,27], suggesting that the early phases of the LWA can operate in this mode.

The ultimate operational characteristics of the LWA are summarized in Table 1.

Table 1: LWA Operational Characteristics

Frequency Range	10-88 MHz (20-80 MHz optimized)
Effective Collecting Area	10-30 MHz: ~10^5 m^2; 30-88 MHz: ~10^5 (30 MHz/ν)^2 m^2
Dipole Elements Total / Dipoles Per Station	~10^4 / 256
Number of Stations / Station Diameter	~53 / ~100 m
Station Interferometer Baseline Range	0.1-400 km
Point Source Sensitivity (dual polarization, 1 hour integration, 4 MHz BW)	1.0 mJy @20 MHz, 0.5 mJy @80 MHz
Angular Resolution	15" @ 10 MHz; 5" @ 30 MHz; 2" @ 80 MHz
Field of View	~2º @ 80 MHz (proportional to ν)
Number of Independent FOV (beams)	4
Maximum Observable Bandwidth	~20 MHz/beam
Spectral Resolution	≤ 1 KHz
Image Dynamic Range	≥ 10^4
Digitized Bandwidth	Full RF

7. Programmatic Issues

The LWA Project is funded through a contract from the Office of Naval Research to the University of New Mexico, which partners directly with the Naval Research Laboratory, Virginia Tech, the Jet Propulsion Laboratory, Los Alamos National Laboratory, and the University of Iowa. In addition, the National Radio Astronomy Observatory, the Air Force Research Laboratory, and New Mexico Tech participate on the LWA Executive Committee.

The Project is now moving into two concurrent phases. The first is commissioning, validation, and scientific operation of LWA-1. This phase will include validation of performance; determination of methods for calibration, correlation, and imaging; and application to a suite of scientific problems, including studies of solar bursts, Jupiter bursts, and ionospheric turbulence and absorption events. The second phase is construction of new stations based on the LWA-1 design, incorporating design changes driven by LWA-1 commissioning results, and operation of the growing array. After validating the station systems in LWA-1, we will construct stations according to the optimal configuration [28, 29]. After completing the first three stations, we will test closure phase methods. There is a planned pause at completion of the first ~ 13 stations, with which we will be able to execute a full-scale proof of principle of complex imaging through the ionosphere and with simultaneous extraction of useful ionospheric data. There is, however, no intrinsic reason why, with subsequent funding, we cannot grow the array to the final 53-station instrument while operating it as an active scientific instrument. Nor is there any

insuperable barrier to using more funds during the first four years to build more than the 13 stations in this plan. A detailed analysis of this case was made for the Astro2010 Decadal Review [30].

The cost of completion depends on the pace at which additional stations are built. Given the advanced state of construction of LWA-1, we have a fairly confident estimated cost of parts and ground preparation of ~$750K/station, and 5,000 hours (2.5 FTE) of construction labor. At a pace of 4 stations/year, and including the cost of making connections to power and data fiber, we arrive at a rough estimate of $1M/station for costs outside construction labor, and ~ $1.2M/station with labor.

The operating cost is driven by power and communications; our estimate is $51K/station/year. The maintenance cost is $38K/station/year, based on very rough reliability figures for the station components. However, at least one major real-world issue – failure rates due to lightning – has not yet been factored in. This component obviously grows with the number of operating stations.

The location of the LWA in New Mexico enhances synergies with two major players in RF studies of space physics. Because the LWA will provide a unique data source for information on the structure and dynamics of the ionosphere at scales from hundreds of kilometers down to hundreds of meters, on timescales of minutes, and with a precision exceeding that of other current measurement techniques, AFRL recognizes that it can be an important asset to studies of ionospheric scintillation and refraction, affecting a wide variety of space-based systems. The LWA is also seen as an important asset for solar studies. The AFRL Space Weather group is currently providing support for students and postdoctoral fellows, establishing joint research projects between its personnel and UNM faculty, and represented on the LWA Executive Committee.

In addition, the NSF/AUI National Radio Astronomy Observatory is also represented on the LWA Executive Committee. NRAO hosts the first LWA station at the Very Large Array and works with UNM on low frequency radio astronomical techniques. NRAO has been an important source of advice, both in the technical arena (receiver development, site development, radio-frequency interference, data communications) and in the science.

Finally, we note the potential utility of LWA data for the requirements of defense, national, civil, and commercial space systems. Navigation (GPS, WAAS) and communications systems suffer deterioration and even outages because of small-scale ionospheric structures; thus, the ability to predict their occurrence and extent is needed. DoD global assimilating ionospheric models (comparable to modern numerical weather prediction models) require estimates of sub-grid structure that few instruments other than the LWA could provide. Space-based attempts to geolocate tactical radio emitters (many in the HF band) require corrections for refractive offsets due to these small-scale structures. Space-based radars operating in the VHF (for foliage penetration) and in the HF (for ground penetration) require similar refractive corrections, and must also be able to make coherent two-dimensional images. Over-the-horizon radars for very precise long-range target detection and tracking, as well as for large-scale ocean surface current measurements, need to calibrate the ionospheric mirror at fine levels. Operationally responsive space assets that require good ionospheric propagation conditions (for tactical geolocation, communications, and other aspects discussed above) will need to predict ionospheric 'weather' sufficiently well to support mission planning; this requires an understanding of how ionospheric dynamics leads to small-scale structure, as will be studied by LWA.

8. References

1. Kassim *et al*. 2010, 'The Long Wavelength Array (LWA): A Large HF/VHF Array for Solar Physics, Ionospheric Science, and Solar Radar', *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*, Sept 14-17, 2010, Maui, Hawaii. (http://www.amostech.com/TechnicalPapers/2010/Posters/Kassim.pdf)
2. Ellingson *et al.*, 'The Long Wavelength Array,' *Proceedings of the IEEE*, **97**, 1421-1430, 2009.
3. Kassim *et al.*, 'The 74 MHz System on the Very Large Array,' *The Astrophysical Journal Supplement Series*, **172**, 686-719, 2007.
4. Bastian *et al.*, 'The Coronal Mass Ejection of 1998 April 20: Direct Imaging at Radio Wavelengths,'

The Astrophysical Journal, **558**, L65-L69, 2001.
5. Stewart *et al.*, 'Visible light observations of a dense plasmoid associated with a moving Type IV solar radio burst,' *Astronomy and Astrophysics*, **116**, 217-223, 1982.
6. Jacobson and Erickson, 'Observations of electron density irregularities in the plasmasphere using the VLA radio interferometer,' *Annales Geophysique*, **11**, 869-888, 1993.
7. Cohen and Röttgering, 'Probing Fine-Scale Ionospheric Structure with the Very Large Array Radio Telescope,' *Astronomical Journal*, **138**, 439-447, 2009.
8. Vainio *et al.*, 'Dynamics of the Earth's Particle Radiation Environment,' *Space Science Reviews*, **147**, 187-231, 2009.
9. Bucik *et al.*, 'Satellite observations of lightning-induced hard X-ray flux enhancements in the conjugate region,' *Annales Geophysique*, **24**, 1969-1976, 2006.
10. Pulinets, *et al.*, 'Main phenomenological features of ionospheric precursors of strong earthquakes,' *Journal of Atmospheric and Solar-Terrestrial Physics*, **65**, 1337–1347, 2003.
11. Orsolini *et al.*, 'An upper stratospheric layer of enhanced HNO_3 following exceptional solar storms,' *Geophysical Research Letters*, **32**, L12S01, doi:10.1029/2004GL021588), 2005.
12. Demirkol *et al.*, 'Ionospheric effects of relativistic electron enhancement events,' *Geophysical Research Letters*, **26**, 3557, 1999.
13. Mitra and Shain 1953, *J. Atm. Terr. Physics*, **4**, 204.
14. Little and Leinbach 1959, *Proc. IRE*, **47**, 315
15. Kerr, 'On the possibility of obtaining radar echoes from the sun and planets,' *Proceedings of the IRE*, **40**, 660, 1952.
16. Bass and Braude, 'On the question of reflecting radar signals from the sun,' *Ukrainian Journal of Physics*, **2**, 149, 1957.
17. Parrish, 'Solar Radar Experiments,' *Arecibo Observatory Internal Report*, CRSR 300, (D. Campbell, private communication), 1967.
18. James, 'Radar studies of the sun,' in *Radar Astronomy*, edited by J. V. Evans and T. Hagfors, McGraw-Hill Book Company, New York, 1968.
19. James, 'Some observed characteristics of solar radar echoes & their implications,' *Solar Physics*, **12**, 143-162, 1970.
20. Rodriguez, P., 'Radar studies of the solar corona: A review of experiments using HF wavelengths,' in *Radio Astronomy at Long Wavelengths*, *Geophysics Monograph Series*, **119**, 155, 2000.
21. Coles, W. A., J. K. Harmon, M. P. Sulzer, J. L. Chau, and R. F. Woodman, 'An upper bound on the solar radar cross section at 50 MHz,' *Journal of Geophysical Research*, **111**, A04102, doi: 10.1029/2005JA011416, 2006.
22. Deller et al., 'DiFX: A Software Correlator for Very Long Baseline Interferometry Using Multiprocessor Computing Environments,' *Publications of the Astronomical Society of the Pacific*, **119**, 318-336, 2007.
23. Cohen et al., 'The VLA Low-Frequency Sky Survey,' *Astronomical Journal*, **134**, 1245-1262, 2007.
25. Perley and Bust, 'Probing the Ionosphere with the Very Large Array,' URSI General Assembly; 2002, LWA Memo #8 (http://www.ece.vt.edu/swe/lwa/memo/lwa0008.pdf).
25. Saha, 'Modern optical astronomy: technology and impact of interferometry,' *Reviews of Modern Physics*, **74**, 551-600, 200.
26. Baldwin et al. 'The first images from an optical aperture synthesis array: mapping of Capella with COAST at two epochs,' *Astronomy and Astrophysics*, **306L**, 13-16, 1996.
27. Hummel et al., 'First Observations with a Co-phased Six-Station Optical Long-Baseline Array: Application to the Triple Star η Virginis,' *Astronomical Journal*, **125**, 2630-2644, 2003.
28. Cohen, 'A Potential Array Configuration for the LWA,' LWA Memo #55 (http://www.ece.vt.edu/swe/lwa/memo/lwa0055a.pdf), 2006.
29. Cohen and Taylor, 'Array Configuration for the Long Wavelength Intermediate Array (LWIA),' LWA Memo #113 (http://www.ece.vt.edu/swe/lwa/memo/lwa0113.pdf), 2007.
30. Taylor *et al.* 2009, LWA Memo #162 (http://www.ece.vt.edu/swe/lwa/memo/lwa0162.pdf)

The Heliophysics Data Environment as an Enabler of HP Science of the Next Decade

D. Aaron Roberts
NASA GSFC

Co-Authors/Supporters:

Robert Bentley, University College, London
Reiner Friedel, Los Alamos National Lab
Eric Donovan, University of Calgary
Joseph Gurman, NASA GSFC
Frank Hill, National Solar Observatory
Eric Kihn, NOAA/National Geophysical Data Center
Janet Luhmann, University of California, Berkeley
John Manuel, Canadian Space Agency
Zoran Mikic, Predictive Science, Inc.
Robert McGuire, NASA GSFC
Larry Paxton, Johns Hopkins University, Applied Physics Lab
Christopher Perry, STFC Rutherford Appleton Laboratory
Robert Rankin, University of Alberta
Robert Weigel, George Mason University
Janet Kozyra, University of Michigan
Jim Thieman, NASA GSFC
Raymond Walker, University of California, Los Angeles

Contents:

Summary
Introduction and Goals
Historical and Scientific Motivation
Current Status of the Heliophysics Data Environment
Plans
Needs
Acronyms and Glossary

Summary

Heliophysics science requires efficient, long-term access to well-maintained repositories of carefully prepared, documented, and preserved data to "develop an integrated research strategy that will present means to address [high-priority scientific] targets," as required in the charter for this Decadal Survey. Because of this, we strongly urge the decadal committee to:

(1) reaffirm support for the "Solar and Space Physics Information System" recommended in the last decadal report, and that has been moving forward due to the efforts of many countries and agencies;
(2) assert the importance to the accomplishment of science goals of adequate and sustained funding for agency efforts to maintain and further improve long-term archive and distribution mechanisms; and
(3) strongly support the need for general standards for archiving and distribution of data as exemplified by those contained in the NASA Heliophysics Science Data Management Policy, and starting with the endorsement of an open data policy by all relevant agencies.

Introduction and Goals

Research in Heliophysics (HP) now routinely involves working with many instruments, often in different space- or ground-based observatories, correlated with various time lags and interpreted or predicted by simulations and theory. Data from decades ago, such as from the Helios spacecraft to the inner heliosphere or the CRRES mission through the radiation belts, are still frequently used to examine questions of intrinsic importance and for planning of new missions such as Solar Probe Plus or the Radiation Belt Storm Probes. The testing and improvement of models of the magnetosphere require data from DoD spacecraft (such as DMSP and TWINS), NSF and international collections of ground-based instruments (e.g., the magnetometer network seen in SuperMAG), and from a large fleet of spacecraft from Japan (e.g., Geotail), Europe (the four Cluster spacecraft, for example), NASA (Wind, Polar, THEMIS, etc.), and other nations. Studies of the Sun using the rich new data from the Solar Dynamics Observatory only realize their full potential when used in conjunction with data from Hinode, STEREO, other spacecraft, and a variety of ground based telescopes. **Thus, we have a strong need for a coordinated, easy to use, international, long-term data environment to exploit the available resources to solve the problems facing Heliophysics in the next decade. Recent developments have both brought us much closer to this goal and point the way toward more powerful techniques to come.**

At the time of the previous decadal survey, Heliophysics data systems were improving in a piecemeal fashion, but there was no overall plan for going forward. At NASA, the open data policy and the increasing emphasis on data plans in Senior Reviews of extended missions were generally making data more available and easier to use. Data were increasingly becoming available online from US and international space- and earth-based observatories. However, there were no explicit guidelines for what data and metadata were to be captured and how they were to be served and preserved. Since then, much has been achieved in attaining a goal of the previous Decadal Survey, as part of the Vitality Program, that stated a requirement for a

> "Solar and Space Physics Information System: Multiagency program for integration of multiple data sets and models in a system accessible by the entire solar and Space Physics community."

The goal of this White Paper is to argue for the continuing need to support a Heliophysics Information System as essential to HP research, and to indicate how recent developments are making Heliophysics data increasingly easy to use and supportive of the needs of current research. **Given the many diverse groups involved in this system, the integration of the HP Data Environment must be based on shared standards, rather than on direct consolidation. The most important "standard" is the acceptance of an open data policy by all providers of HP data.**

The above-recommended Vitality Program activity has been advanced by a number of initiatives. One early significant effort was the establishment of the international Space Physics Archive Search and Extract consortium in 2001 to encourage interoperability and develop a common "data model" to provide uniform descriptions of Heliophysics data products. This effort continues to be supported by NASA. This was followed by one of the most important initiatives, the institution in 2007 of a NASA Heliophysics Science Data Management Policy and its subsequent significant revision in 2009 (http://hpde.gsfc.nasa.gov/Heliophysics_Data_Policy_2009Apr12.html). This Policy codified many *de facto* practices and provided the basis for clarification and advancement of the HP Science Information System as envisioned in the Decadal Survey. The flow of data from spacecraft to final archive and from server to end-user is described. The Data Policy specifies the requirements and expectations for Project Data Management Plans, Mission Archive Plans, Resident and Final Archives, Virtual Observatories, and data/metadata standards. The basic set of groups and functions have been or were already established within NASA to accomplish the goals of the Data Policy. New missions are using the Policy to guide their efforts, and older missions are now largely compliant with it. We have yet to realize the full potential of new approaches such as Virtual Observatories, but the groundwork has been laid. The Data Policy has facilitated interactions with other agencies (especially NOAA and NSF) and with international partners in Japan, Canada, Europe, and other countries, and it has been used for guidance in the planning of similar policies outside of NASA.

NASA Heliophysics has been committed to stable funding of the NASA HP Data Environment, as described in the Data Policy, and it is crucial for the success of essential multifaceted and long-term scientific studies that this support and that of other agencies be continued.

The rest of this White Paper provides information on the need for an HP Information System, and the current status, plans, and needs of the Heliophysics Data Environment (HPDE; see also http://hpde.gsfc.nasa.gov), with the emphasis on NASA's role but with indications of the essential and coordinated roles played by many others.

Historical and Scientific Motivation

Initial observational efforts in Space Physics consisted of a team flying their own experiment in space. The results would become known to the rest of the community and the general public through meetings, press conferences, and publications, but the data were not generally shared, and there were no standards for formats if they were. Thus we learned of the existence of radiation belts, the solar wind, the Earth's bow shock, solar coronal holes, and the other now familiar aspects of the heliosphere through a series of carefully designed instruments sent out to a wide variety of locations with no clear sense of how or when the data might be shared. It was recognized early on that the data thus gathered were unique and irreplaceable, so provisions were made for a National Space Science Data Center to save them and to document their formats, but often, while the bits were preserved, the utility of the data declined rapidly after the end of the mission due to the difficulty of reading and using what were essentially proprietary datasets. A number of these datasets have been reformatted for general use over the decades since their production. When the expertise was available, they have been reprocessed and improved, and they are now easily web-accessible, typically as ASCII or CDF files. One goal of the current HP Data Environment is to avoid the need to recover data, but rather to have them seamlessly pass to long-term active archives.

Subsequent missions gathered time series and images that have become increasingly like long-term weather station and imagery records of the Earth's weather and climate, but now for the history and dynamics of the Sun, heliosphere, magnetospheres, and the near-planet plasma environments. Scientific questions have progressed from determining the elements of the heliophysical system to determining both the small-scale details of the physics (e.g., how magnetic fields reconnect to release energy in flows, waves, and energetic particles) on the one hand, and the global interactions of the various components (such as how Coronal Mass Ejections form, propagate through space producing energetic particles, and interact with the Earth to generate magnetic and particle storms) on the other. Both of these kinds of investigation require substantial coordination of data from multiple instruments, multiple spacecraft, and comparisons of these with simulation results and theory. For example, much deeper understanding of reconnection and its role in the magnetosphere has come from simulations combined with detailed studies of the data from many

instruments on the four Cluster spacecraft for the microphysics, and from the many instruments on the five THEMIS spacecraft combined with other spacecraft studies and ground-based observations of magnetic fields, aurora, and auroral winds for the large-scale picture. In this way we have come to appreciate, for example, the role of electrons in the small-scale processes and the sequences of reconnection and energization of the near-Earth fields and particles.

Establishing a vibrant, robust and agile data environment is an essential element of an integrated research strategy capable of addressing the science targets identified by the Decadal Survey. The creation of the Heliophysics Information System is guided by efforts in other science domains. Progress in the Earth Sciences has been greatly aided by the use of a uniform data format for the large missions, namely HDF (which now incorporates NetCDF) and the provision of data from distributed but uniformly accessible sites. The connections to data have been enhanced by the adoption of OPeNDAP as a protocol for data transfer that allows users to bring data directly into their favorite analysis applications so they can apply their own tools or those from libraries developed by others. A similar development has occurred in Astronomy where the US Virtual Astronomical Observatory and the International Virtual Observatory Alliance have set standards for data, metadata, transfer protocols, and other aspects of the data system that now allow astronomers to call up previous observations of any object in the sky or to do complex correlations over large sets of objects defined by user-selected criteria. This promises to revolutionize the types of studies possible, and thus to greatly deepen insight on, for example, the evolution of stars and galaxies. These types of capabilities, with similar potential for aiding research, are now on the horizon for Heliophysics, with core capabilities already in place.

Current Status of the Heliophysics Data Environment

The HPDE consists of a set of related data and modeling systems that are becoming increasingly interconnected. Large, active repositories of data from old and new observatories and programs are held at NOAA's NGDC; NASA's SPDF, SDAC, and NSSDC (data) and CCMC (models); JAXA's DARTS; the European CDPP; various NSF locations such as the SuperMAG and CEDAR/VSTO sites; and at a large set of active mission repositories in many countries. An increasingly complete set of these data sources is being registered in uniform terms that can be easily searched and accessed (see below). There has been considerable convergence on data formats, making data use much easier. Data are flowing in a natural way into post-mission archives, although in some cases clearer paths need to be mapped out. The current HP Information System is functioning well, delivering large volumes of data to researchers in useful ways.

In addition to the formulation of a NASA HP Data Policy, a number of complementary developments have considerably advanced the Data Environment. The last two NASA HP mission Senior Reviews required each mission to provide a Mission Archive Plan (specified in the Data Policy) that detailed the products, documentation, formats, and dissemination methods of the mission data for the long term. In each review, new datasets were produced and better documented, and the plans have become progressively clearer. Thus, the active missions now are, for the most part, producing highly useable products for the long term. New NASA missions, such as RBSP and MMS are taking the Data Policy to heart, following the PDMP guidelines and planning for appropriate description and archiving of data. Other datasets are becoming more unified, such as the NSF, NASA, and internationally sponsored ground-based magnetometer chains that are now largely united by the SuperMAG project that is sponsored by both NSF and NASA.

The uniformity required for data descriptions to be useful is now being supplied by the "SPASE" Data Model (http://www.spase-group.org) that consists of a uniform set of terms and an implementation (in XML) that allows us to describe and access products and services from all Heliophysics disciplines, wherever they are located throughout the world. This Data Model is the result of many years of collaboration by a representative group from many US and foreign organizations including NASA, NOAA, and many universities; there are now SPASE data model adopters in Europe, Japan, and Canada in addition to NASA Heliophysics Virtual Observatories and Data Centers. Within the next two years, NASA's Space Physics Data Facility, in its role as a NASA Active Archive, plans to have an essentially complete set of high-level product and service descriptions, made available through the Virtual Space Physics Observatory (http://vspo.gsfc.nasa.gov) as a registry. It is already possible to locate and access data from most of the current and many past Heliophysics missions. The NASA HP Virtual Observatories work with missions

and archives, and in particular providing a contact point for new missions, to provide more detailed descriptions of the complete set of products to assure that products will be scientifically useable based on the SPASE metadata. Some NSF-based products (such as ground magnetometers) are now registered in SPASE terms, and informal arrangements have been made for NOAA and NASA to help NSF to serve and preserve datasets produced by NSF Space Physics projects. On the solar physics side, the Solar Data Analysis Center has been working with recent missions to provide access to data both directly and through the Virtual Solar Observatory that unites access to a very wide range of solar physics products, including all the products from the ESA/NASA SOHO mission, the two STEREO spacecraft, and the huge data volumes coming from the Solar Dynamics Observatory. Uniform VSO access complements mission specific access to data.

Plans

The distributed HP Information System expects to continue to archive and provide access to an increasingly complete set of data from observations and complementary models. Funding should continue to be provided so that older datasets will be upgraded to be better and more accessible when deemed important. As missions terminate, data will flow into post-mission archives, and standards for archival formats are being established more clearly (e.g., there is a plan to make CDF a true NASA archival standard, to join FITS and HDF/NetCDF). Datasets will continue to be registered in SPASE terms, and made accessible through Virtual Observatories as unifying portals. This general unification will enable a wider range of easy cross-dataset and data-model comparisons than now possible, as well as other services such as searching for data across many missions with a query based on time, location, event or feature lists, the values of particular quantities, or the similarity of target events to chosen events; the efficiency and depth of these sorts of services are potentially transformative in the science that is possible. The progress toward uniform access and the associated growth in possible tools will take place in many US agencies as well as in collaboration with international partners. An example of the latter is the work in progress involving the European "HELIO" project and the US HP Virtual Observatories. Much of what must be done consists of adopting common standards or of finding ways to do rapid translations between systems.

Needs

The Heliophysics Data Environment, now being unified through a partly formal, partly informal Heliophysics Information System, requires a basic level of support to realize its goals. The current level of NASA and other funding, if it remains stable as intended by the current NASA HP managers, is sufficient for the basic needs of the data system, but allows little room for much beyond that. New demands will come from the need to archive large datasets, such as those from recent solar physics missions and upcoming multispacecraft Space Physics missions, and from the related need to support tools for more sophisticated data mining, feature recognition, and other services that will be facilitated by the integration of the data environment. The total funding for HP Data Centers and related data system activities at NASA (outside of mission budgets) amounts to about 1% of the NASA Heliophysics budget, which seems a small price to pay for the benefits. Modest funding increases to preserve and make effective use of the rapidly growing data volumes are warranted to obtain the potential benefits from the data.

In terms of organization, it will be important to continue the work to establish interagency and international standards and collaboration. We should encourage DoD and DoE, along with all other agencies that produce data of importance to Heliophysics, to agree on a general data release policy that safeguards any national security concerns while adhering to an open data policy. Fledgling associations, such as the informal agreement for NASA and NOAA to aid NSF in archiving space physics data, should be recognized at the agency levels as being important and thus encouraged to flourish, and these should be expanded to include all relevant agencies responsible for data provision. Similarly, international efforts to unify standards and increase interoperability through agreements on data and metadata formats and communication protocols should be endorsed and made part of the working environment of each of the constituent groups.

CCMC: Community Coordinated Modeling Center, sponsored by NASA, NSF, the Air Force and other agencies; provides runs on demand and works to transition models to operational space weather use.

CDF: Common Data Format, a self-documenting format increasingly used for the storage and serving of Space Physics data, developed and maintained by the SPDF.

CDPP: Centre de Données de Physique des Plasmas, the French national active archive of Space Physics data.

CRRES: Combined Release and Radiation Effects Satellite, investigated the radiation environment of the inner and outer radiation belts, 1990-1991.

DARTS: Data ARchives and Transmission System, JAXA's active archive for Space Physics (and other) data.

HDF/NetCDF: Hierarchical Data Format, a self-documenting format used for storage and serving of Earth Science and related data, e.g., for ionospheric spacecraft data.

Helios: A pair of spacecraft that explored the inner heliosphere from 0.3 to 1 AU, 1975-1981.

Hinode: A Japanese/US spacecraft that provides high temporal and spatial resolution images of the Sun in a variety of wavelengths, launched in 2006.

DMSP: Defense Meteorological Satellite Program, a series of DoD meteorological satellites that have particle sensors relevant to Space Physics.

FITS: Flexible Image Transport System, a file format used for solar and astrophysical image data.

HELIO: HELiophysics Integrated Observatory, a European consortium (with US involvement) concerned with uniform access to HP data, based on both simple and complex search criteria.

HP: Heliophysics

HPDE: Heliophysics Data Environment, a collective term for the datasets, repositories, access methods, and other services associated with the scientifically relevant output from HP observatories and instruments.

JAXA: Japanese Aerospace Exploration Agency

NGDC: National Geophysical Data Center, NOAA's source for Heliophysics (and other) data.

NOAA: National Oceanic and Atmospheric Administration

NSSDC: National Space Science Data Center, a safe-keeping, "deep" archive for HP data; provides long-term backup services to, e.g., SPDF, among other functions.

Polar, Earth orbiting spacecraft, 1996-2008, that imaged the aurora and measured in situ particles and fields

SDAC: Solar Data Analysis Center, a NASA HP designated Active Final Archive for Solar data; serves solar data from many current and past missions.

SPASE: Space Physics Archive, Search, and Extract, an international collaboration that has produced a Heliophyics Data Model that allows the uniform registration and use of HP data and services.

SPDF: NASA's Space Physics Data Facility, the primary NASA Active Final Archive for Space Physics data, home to data systems such as CDAWeb and OMNIWeb.

STEREO: Solar TErrestrial RElations Observatory, a pair of Sun orbiting spacecraft that image and measure the particles from Coronal Mass Objections and other solar activity, lunched in 2007.

THEMIS: Time History of Events and Macroscale Interactions during Substorms, a set of five spacecraft studying magnetic and particle activity in the magnetosphere, launched in 2007.

TWINS: Two Wide-angle Imaging Neutral-atom Spectrometers, a pair of instruments on non-NASA spacecraft that use stereo energetic neutral atom images to study the Earth's ring current, launched in 2008.

Wind: A spacecraft that measures the solar wind fields and particles upstream of the Earth, launched in 1994

SuperMAG: A worldwide collaboration of ground-based magnetometers, supported by NSF and NASA.

Telemachus $^{(+)}$ *Redux*$^{(++)}$

$^{(+)}$ Faithful son of Odysseus (Ulysses) $^{(++)}$*Brought back*

Edmond C. Roelof and G. Bruce Andrews*
Johns Hopkins University/Applied Physics Laboratory, Laurel, MD 20723-6099

*For the Telemachus **NASA Roadmap** Teams[1]

Abstract: Telemachus is a polar solar-heliospheric mission first described in the NASA Sun–Earth Connections Roadmap (2003–2028). To continue the quest begun by Ulysses, Telemachus will be injected into a 0.2 AU-2.5 AU solar polar orbit (~90 deg inclination from the ecliptic) after gravitational encounters (VEEJGA) with Venus, Earth (twice) and Jupiter, followed by a maneuver at first perihelion. Spacecraft power is drawn from conventional solar panels. With an orbital period of 1.5 years, Telemachus will pass over the solar poles at a radial distance of 0.37 AU. It will alternate between Sun-pointing spinning beyond 1AU (optimum for *in situ* plasma, fields, and particles observations) and 3-axis stabilization inside 1 AU (optimum for solar polar imaging over a range of wave lengths). The relatively few technical concerns raised in the earlier formulations of the Telemachus mission design and scientific payload have been significantly mitigated by more recent flight experience and/or studies from relevant missions like MESSENGER, New Horizons, STEREO, Solar Probe Plus, and Beppi Colombo/Solar Orbiter

Approach: The main purposes of this white paper are to:

1. Describe briefly the Telemachus mission concept and its observational advantages as an out-of-the-ecliptic (OOE) mission.
2. Highlight the recent exciting scientific discoveries that raise Telemachus' priority.
3. Demonstrate that Telemachus is a well-studied, doable and cost-effective mission because it utilizes present-day technology for the spacecraft, propulsion, and instrumentation.

The material in this white paper will therefore show that Telemachus meets all of the following evaluation criteria explicitly specified in the *Decadal Study Request for Information from Community*.

a. *Identified* in NASA's 2003 Sun-Earth Connection and 2006 Heliophysics Roadmaps.
c. *Contributes to important scientific questions facing solar and space physics today, e.g.,* imaging polar helioseismology of interior dynamics as well as the link between the non-equatorial solar corona and the heliosheath beyond the termination shock through *in situ* measurements of the solar wind, magnetic field, plasma waves and suprathermal ions, and energetic particles
e. *Complements other observational systems or programs available, e.g.,* GONG Network, Solar Dynamics Observatory, Heliospheric Orbiter, Solar Probe Plus.
f. *Is affordable (cost benefit)*; Already studied as an LWS Low-to-Mid-Cost Multi-

Objective Mission, corresponding to a ~ small ($250M-$500M) to medium ($500M-750M) cost mission (in FY 10 dollars).

g. *Appropriate degree of readiness (technical, resources, people)*; See (f) above.
h. *Fits with other national and international plans and activities;* See (e) above.

Task: Telemachus addresses *compelling science challenges that have arisen from recent advances and accomplishments* in solar imaging as well as in the dynamics of the heliosheath beyond the termination shock that are driven by the global solar wind and the suprathermal particles within it.

Scope: Telemachus will measure the:

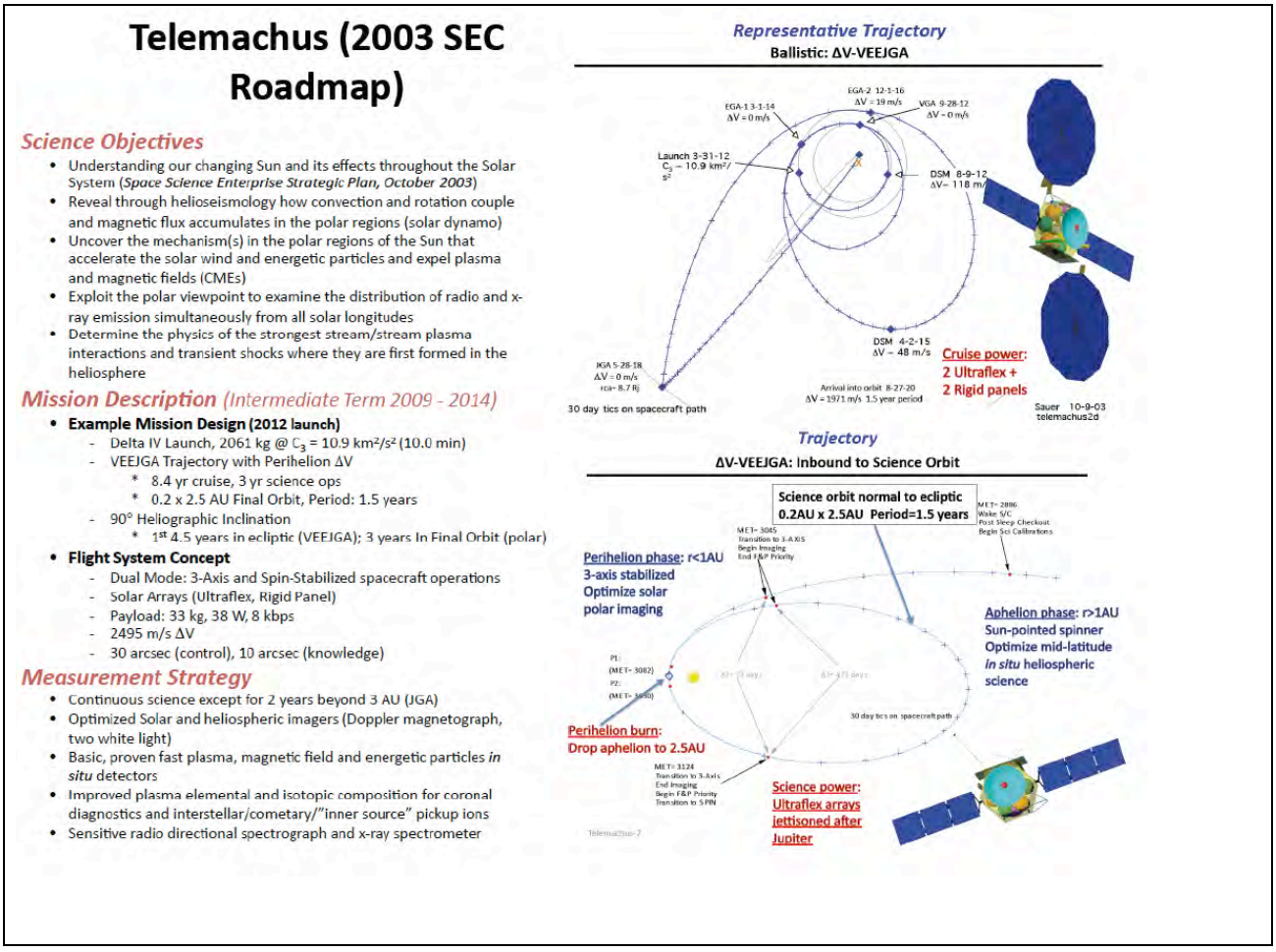
Structure of Sun and properties of its outer layers

Characteristics and physics of the interplanetary medium from the surface of the Sun to interstellar space beyond the boundary of the heliosphere

Theme: *Solar and Heliospheric Physics*

Telemachus Mission Design:

The scientific objectives require a polar orbit of Telemachus (see the Figure). The orbit is inclined ~90° to the ecliptic with a perihelion ~0.2AU and an aphelion of 2.5AU. This orbit passes over the solar polar caps at 0.37AU (80 R_S), compared to 2.4 AU for



Ulysses. The period is ~1.5 years which will yield 14 polar passes (north and south) within one solar cycle. The duration of each polar pass above 45° latitude lasts about 5 weeks. The period is ~1.5 yr which gives the advantage that, if the first perihelion is in quadrature with the Earth, then every perihelion pass would be off the solar limb, alternating east and west.

The orbit involves gravity assists (VEEJGA) using flybys of Venus, Earth (twice) and Jupiter (similar to Galileo and Cassini). A "quad" chart (modified from the 2003 NASA SEC Roadmap) is shown in the Figure. The first 4.8 yr are in the three ecliptic petals, during which time excellent imaging and *in situ* measurements can be made. Once en route to Jupiter, no science data will be transmitted when Telemachus is beyond 3 AU (outbound or inbound), because Telemachus the limited (but sufficient) power available from the conventional solar arrays and "flash" batteries (eliminating any dependence on non-conventional propulsion or RTGs). This beacon-mode phase will last 2.7 yr, followed by the beginning of high-latitude (16°) science 7.5 yr after launch. First perihelion will begin in the 8th yr after launch when a motor burn will inject Telemachus in its final 0.2x2.5 AU orbit inclined 90°.

Radius-dependent priorities between imaging and *in situ* science will be accommodated by having the spacecraft mainly spinning for r>1 AU (latitude<48°) but 3-axis stabilized for r<1 AU (latitude>48°). Of course, either of these dual modes can be commanded at any place along the orbit. Near perihelion, science will be second in priority to assuring the survival of the spacecraft through the thermal environment of each perihelion.

Instrumentation: The original 2003 NASA/NRC Roadmap instrument complement is presented in the first Table. It goes without saying that the payload should be re-examined in the light of the new science that has accumulated over the intervening 7 years. That necessary activity is beyond the scope of this white paper, but we offer some comments below on progress in the relevant technologies. The details of the *observational variables to be measured* and the *characteristics of measurement (accuracy, spatial and temporal resolution)*, and their *domain of observation* will be found in the second Table. It has been extracted from Appendix A of the JPL Team-X report for the 2003 NASA/SEC Roadmap.

Scientific Motivation: *(Importance of achieving the science objectives in the context of the state of the field)*

Although there are a myriad of solar and interplanetary scientific objectives can be examined with Telemachus (as has been evident from the scientific productivity of Ulysses and the ecliptic observations from SOHO, ACE, Wind, STEREO, Yokoh, TRACE, RHESSI, Hinode, *etc.*), we have only enough room here to highlight just two different realms of exciting new science that can be investigated with the dual-mode mission Telemachus:

Imaging helioseismology. Telemachus' imaging of solar polar photospheric flows (heretofore difficult to observe) will provide critical boundary conditions on the transport

of solar magnetic fields. These global flows and fields are direct signatures of the solar dynamo, leading to the emergence of magnetic fields and the formation of active regions and eventually all the manifestations of solar activity that Telemachus can image. These ultimately determine the dynamics of the entire heliosphere. The high-order ploidal modes of the solar dynamo are difficult to resolve from low-latitude viewing points, and these modes are sensitive to coupling dynamics. Telemachus will measure how magnetic flux accumulates in the polar regions. This is a critical test of flux transport theories. A recent review[2] summarizes the current status of helioseismology, the diverse recent results, and prospects for future investigations. The limitations of space in this white paper preclude further discussion, but our community already strongly affirms the compelling nature of polar helioseismology. We therefore devote more space to the very recent transformational science of the heliosheath and its dependence on the global solar wind structure and PUIs that Telemachus will measure in transit to the termination shock.

In situ characterization of the out-of-the-ecliptic solar wind and interstellar pickup ions. The ongoing Voyager Interstellar Mission (VIM) and recent observations from the energetic neutral atom (ENA) imagers of the Interstellar Boundary Explorer (IBEX) and Cassini/INCA missions have revealed the interaction of the heliosphere with the very local interstellar medium (VLISM) to be much more complex than heretofore assumed by our present day concepts. Voyager 1 and Voyager 2 continue to reveal significant fluxes of energetic particles in the heliosheath, including a well-defined suprathermal ion "tail" in which the differential intensities fall off $\sim E^{-1.5}$ above ~30keV[3]. At even higher energies (~100 MeV), there is no "unfolding" of the energy spectrum of the anomalous cosmic rays (ACRs), thus pointing to a more remote location for the modulation region and source[4]. Most strikingly, direct measurements of the shocked, solar-wind flow speed obtained from Voyager 2 revealed that the flow remains supersonic in the heliosheath beyond the termination shock[5].

All of these particle observations, taken together, unambiguously demand that the bulk of the energy density in the heliosheath plasma resides in a non-thermal component that extends to very high energies[6]. Strong implications, both quantitative and qualitative, follow from this fact for the overall heliosheath structure. We have never encountered a large-scale plasma regime in which the non-thermal ion pressure dominates the thermal pressure and overwhelms the magnetic field stresses. This is why even sophisticated MHD models failed to predict anything like the striking new features that have just been observed in the last two years in the all-sky energetic neutral atom (ENA) images from IBEX[7] and Cassini/INCA[8] of the remarkable (and as yet not satisfactorily explained) "ribbon" and "belt" produced by energetic protons (1-50 keV) that arc across the entire sky. The energy and mass sources for these huge structures must come from the solar wind and pickup ions and energetic particles within it. Ulysses revealed that the strongest long-lived stream-stream interactions occur in the mid-latitude solar wind[9], and that is also where the interstellar pickup ions (PUIs) are most efficiently accelerated[10] and develop their "ubiquitous" high-energy spectral tails $\sim E^{-1.5}$. These mid-latitude regions will be repeatedly sampled *at all heliolatitudes* by the *in situ* Telemachus measurements along its orbit every 1.5 years.

Telemachus

Instrument Summary

- **Representative instrument payload complement/characteristics provided by cognizant science advocates within SEC community (2002).**

Instrument	Mass (kg)	Power (w)	Data Rate (bps)
Magnetometer	*0.6*	*0.5*	*431*
Plasma Composition	*2.6*	*3.0*	*1360*
Fast Plasma	*1.6*	*0.8*	*1071*
Energetic Particles	*0.9*	*1.3*	*521*
Cosmic Ray Telescope	*1.0*	*0.8*	*200*
WAVES	*2.0*	*2.3*	*579*
White Light Coronagraph	*10.0*	*6.5*	*1200*
Heliospheric Imager	*5.5*	*10.0*	*1192*
Doppler Magnetometer	*7.0*	*10.0*	*1200*
X-ray Spectrometer	*1.5*	*2.5*	*273*
TOTAL	**32.7**	**37.7**	**8027**

Note (2010): In light of subsequent mission experience, most of these mass/power estimates were quite optimistic.

Instrument Specifications

← ***Remote***

In situ →

	WAVES (E and B fields)	White Light Coronagraph	Heliospheric Imager	Doppler Magnetometer	X-ray Spectrometer
Mass (kg)	2.0 Two sensor heads, low noise amplifiers, front end electronics, digital interface	10 7.5 kg instrument, 1.3 kg CEB 0.3 kg harness	5.5	7	1.5
Size (cm)	Sensor (E): (qty 3) 100 x 0.30 Sensor (B): 4x4x4 Electr: 10 x 10 x 4	Telescope: 130 x 120 Camera Electronics: 17 x 12 x 8	60 (dia) x 30	Sensor: 60 x 12 x 12 Electr: 10 x 10 x 8	Sensor: 13 x 13 x 5 Electr: 10 x 10 x 7
Power (W) [Peak/Avg/Stdby]	2.3/2.3/0.6	Op Heater Power: 2/2/0 CCD Camera Power: 5/5.5/0	10/10/1.5	10/10/3	2.5/2.5/1
Thermal					
Operational T (°C)	E field antennas (-120 to 1000) Other components (-40 to 50)	-40 to 50	-40 to 50	-40 to 50	-40 to 50
Survival T (°C)	-55 to 65	-55 to 65	-55 to 65	-55 to 65	-55 to 65
Power (mW)	<1000 during cruise	<1000 during cruise	<1000 during cruise	<1000 during cruise	<1000 during cruise
Cost (&M)	4.5	10	5, excluding integration	8	3

	Magnetometer	Plasma Composition (Ions)	Fast Plasma (Electrons & Ions)	Energetic Particles (Electrons & Ions)	Cosmic Ray Telescope (Electrons & Ions)
Mass (kg)	0.6 (sensor, front end electronics, cable, not including boom)	2.6 (one head, HVPS, front end electronics, digital interface)	1.6 (two heads, HVPS, front end electronics, digital interface)	0.85 (one head, HVPS, front end electronics, digital interface)	1.0 (one head, stacked detector config., front end electronics, bias, digital interface)
Size (cm)	Sensor: 3 x 6 x 4 Electr: 10 x 10 x 3	Sensor(i): 6 x 6 x 7 Electr: 10 x 10 x 6	Sensor(E): 6x 6 x 5 Sensor(i): 6 x 6 x 5 Electr: 10 x 10 x 5	Sensor: 16 x 10 x 2 Electr: 10 x 10 x 5	Sensor: 3 x 3 x 4 Electr: 10 x 10 x 3
Power (W) [Peak/Avg/Stdby]	.5/.5/.3	3/3/2	0.75/0.75/0.4	1.3/1.3/0.7	0.8/0.8/0.2
Thermal					
Operational T (°C)	-40 to 50	-40 to 50	-40 to 50	-40 to 50	-40 to 50
Survival T (°C)	-55 to 65	-55 to 65	-55 to 65	-55 to 65	-55 to 65
Power (mW)	<1000 during cruise	<1000 during cruise	<1000 during cruise	<1000 during cruise	<1000 during cruise
Cost (&M)	1.5	4	3	3.5	2.5

		WAVES (E and B fields)	White Light Coronagraph	Heliospheric Imager	Doppler Magnetometer	X-ray Spectrometer
Data Sampling (bits/sample)						
Digitization Prec		16	14	16	14	12
Sampling Freq (samples/s)						
Routine		1 Hz	1024 x 1024 image 10 sec exposure 20 minute cadence	1242 x 1152 image 2 minute cadence	1024 x 1024 image 10 sec exposure 20 minute cadence	5,000 events/sec
Burst 1		4KHz	2 minute cadence		2 minute cadence	
Burst 2						
Bus Xfr Rate (bps)	Total					
Routine	2,234	579	1200 (assuming 10:1 lossy compression)	1192 bps (co-add/ compressed)	1200 (assuming 10:1 lossy compression)	273
Burst 1	26,560	14684 (stored on SSR during perihelion pass)	12000 (assuming 10:1 lossy compression)		12000 (assuming 10:1 lossy compression)	
Burst 2	66,496					

	Magnetometer	Plasma Composition (Ions)	Fast Plasma (Electrons & Ions)	Energetic Particles (Electrons & Ions)	Cosmic Ray Telescope (Electrons & Ions)
Data Sampling (bits/sample)					
Digitization Prec	16	12	12	12	12
Sampling Freq (samples/s)					
Routine	16	Up to 50,000 events/sec	Up to 50,000 events/sec	Up to 50,000 events/ sec	Up to 50,000 events/sec
Burst 1	76				
Burst 2					
Bus Xfr Rate (bps) Total					
Routine 2,234	431	1360	1071	521	200
Burst 1 26,560	2045 (stored on SSR during perhelion pass)	15025 (stored on SSR during perhelion pass)	15025 (stored on SSR during perhelion pass)	2732 (stored on SSR during perhelion pass)	1200 (stored on SSR during perhelion pass)
Burst 2 66,496					

	WAVES (E and B fields)	White Light Coronagraph	Heliospheric Imager	Doppler Magnetometer	X-ray Spectrometer
FOV @ 0.2 AU FOV @ 0.6 AU	N/A	2 – 15 solar radii 6 45 solar radii	Hemisphere (180°)	Full disk of Sun	12°
Mounting: Method and Location	B field antennas on short stalk, E field antennas are S/C corner mounted	On S/C side, pointing sunward	On S/C side, pointing orthogonal to Sun, No intrusions within FOV	On S/C side, point sunward	Inside Deck, Pointing just off Sun
Covers	N/A	One time deployable	Optics covered during station keeping firing	One time deployable	N/A
Pointing: Control Accuracy (°)	30 arc-sec (best that is provided by S/C with bang-bang thruster control)	30 arc-sec (best that is provided by S/C with bang-bang thruster control)	.05 deg	30 arc-sec (best that is provided by S/C with bang-bang thruster control)	1 degree
Ctr'l Knowledge (°)	10 arc-sec	10 arc-sec		10 arc-sec	
Control Stability (°/s)			1.5 arc-sec/sec		

	Magnetometer	Plasma Composition (Ions)	Fast Plasma (Electrons & Ions)	Energetic Particles (Electrons & Ions)	Cosmic Ray Telescope (Electrons & Ions)
Clear FOV	N/A	Instantaneous FOV - 360° azimuth x 70° elevation; full sky when spinning	Instantaneous FOV - 90° azimuth x 50° elevation; 2x when spinning	Instantaneous FOV - 160° azimuth x 12° elevation, nearly full sky when spinning	Instantaneous FOV - 30° full cone
Mounting: Method and Location	On Boom – anti-sunward direction	Inside Deck Mounted, view sky through aperture in S/C wall	Inside Deck Mounted, view sky through aperture in S/C wall	Inside Deck Mounted, view sky through aperture in S/C wall	Inside Deck Mounted, view sky through aperture in S/C wall
Covers	N/A	Possibly not needed – verification by acoustic test	Recloseable Cover, Acoustic protection and thermal protection during perihelion pass	One time deployable	Possibly not needed – verification by acoustic test
Pointing:					
Control Accuracy (°)					
Ctr'l Knowledge (°)	1	1	1	1	1
Control Stability (°/s)					
Note: Payload is not powered during cis-Jupiter (i.e., cruise) phase due to lack of power					

Update on technology readiness levels for Telemachus:

The technology data presented in this white paper are drawn from the comprehensive JPL Team-X Study of the Telemachus mission produced for the 2003 NASA Sun-Earth Connections Roadmap. Details are given in its Summary and Appendices. The consensus was that the mission was essential technologically ready in 2002, but there were a few technical concerns. We now discuss how those concerns have been effectively mitigated by the rapid technological progress over the past 7 years.

The Telemachus orbit will have a perihelion of 0.2 AU. This imposes severe demands on the thermal protection system of the spacecraft. Fortunately, near-Sun missions have been a focus of intense technical development during the past decade. The currently-flying MESSENGER mission has performed three flybys of Mercury with a closest approach of 0.35 AU. The Bepi-Columbo/Solar Orbiter spacecraft are presently in development for near-Sun missions and are being designed for closest approaches of 0.28 AU. The now-approved Solar Probe Plus mission, diving to within 0.16 AU of the Sun, has completed several studies to elevate the TRL of the thermal protection system. In short, the technology now exists to lower the thermal environment risks of Telemachus to an acceptable level.

The unique vantage point of the Telemachus spacecraft during the 1.5 year polar orbit allows for both exciting solar polar imaging as well as *in situ* measurements. The remote sensing instruments require 3-axis stabilization while the *in situ* particles and fields instruments can best acquire data when the spacecraft spinning (although some imaging can be achieved while spinning and useful *in situ* data can be obtained when 3-axis stabilized). Consequently, the spacecraft is designed to be 3-axis stabilized near the sun to support remote sensing measurements, but in order to conserve fuel and minimize wear and tear on the momentum wheels, the spacecraft will be configured to spin during the large period of time when the orbit is outside 1 AU. The avionics to support this dual mode of operation have been developed and are in use in flight on the New Horizons mission. During planetary encounters (*e.g.* Jupiter) the New Horizons spacecraft is 3-axis stabilized while during cruise phase it is a spinner.

The swing-by of Jupiter is necessary in order to significantly change the inclination of the initial orbit from ecliptic to polar. If no further orbit adjustments were made after the Jupiter swingby, the orbit would be similar to that of Ulysses and the orbit period will be ~6 years. In order to maximize near-Sun science imaging while still sampling the mid-latitude heliosphere *in situ*, the Telemachus orbit will be reduced to a less eccentric elliptical orbit by performing a perihelion burn at 0.2 AU. The baseline method advocated in 2002 to achieve the required delta-V was to utilize the main thruster of the bi-propellant propulsion system. In order to maintain the position of the heat shield towards the sun, it was expected that the thruster needed to be gimbaled. However, this requirement for a gimbaled main thruster should now be revisited, as the MESSENGER mission is able to achieve Mercury orbit insertion, while keeping the heat shield toward the sun, without the need for gimbaling. Performing the perihelion burn with a separate

Star motor was also studied. A Star motor burn is more efficient, from a mass standpoint, because of the higher thrust of the motor when compared to a bi-propellant, but the spacecraft must be spun up to ~60 rpm prior to the burn. This imposes constraints on the design of the heat shield, which must be re-configured to surround the spacecraft during the brief period of time while the spacecraft is spun up and the motor is firing. A thermal analysis of this "rotisserie" mode was recently performed at JHU/APL for Telemachus with a MESSENGER-type heat shield, and it indicated that temperatures on the spacecraft could be held within reasonable limits during the orbit adjustment event.

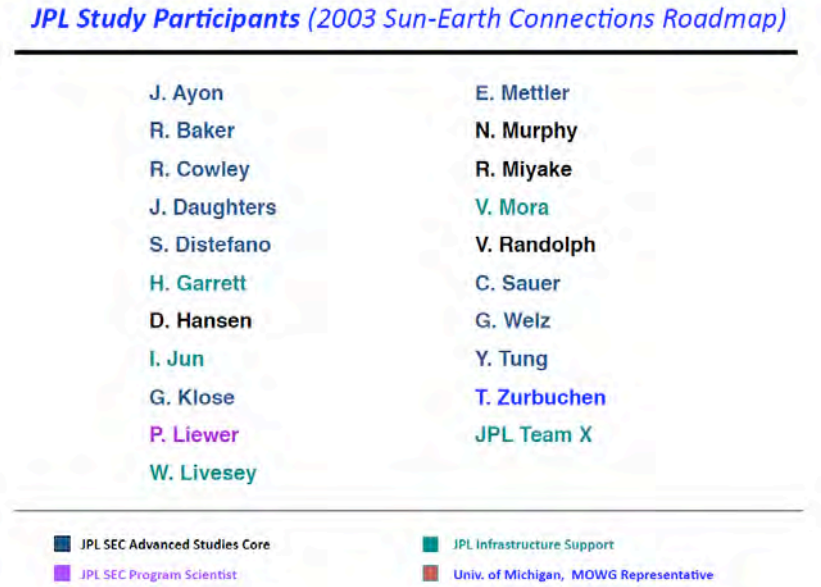
For the Telemachus mission the resources available for the payload were tightly constrained (33 kg, 38 Watts), implying that efforts must be made to minimize instrument mass and power. Fortunately, there has been substantial progress in the past few years in developing the technological infrastructure in support of miniaturized instrumentation. As an example, the ULEIS energetic ion composition instrument (currently operating on the ACE spacecraft launched in 1997), is an isotopic spectrometer built using 1990's technology. It weighs 18 kg and consumes 21 watts. A similarly performing instrument is presently in development for the Solar Orbiter mission, with two independent telescopes that weigh only ~3 kg and dissipate only ~3 watts. Another example of a modern miniaturized instrument is the PEPPSI energetic particle instrument on the New Horizons mission to Pluto. It has six look directions and is able to determine the composition of medium energy particles, yet weighs only 1.3 kg and draws <2 watts. These low resource instruments were enabled by the development of electronics dramatically advanced in terms of radiation tolerance, ultra-low power dissipation, and high density packaging. Digital electronics that previously occupied several circuit boards can now be incorporated into a single field programmable gate array (FPGA), such as the radiation-hardened Actel RTAX with nearly 4 million system gates. In the analog circuitry regime, a number of institutions have developed Application Specific Integrated Circuits (ASICs) to perform signal conditioning and processing tasks. For example, a JHU/APL time-of-flight ASIC measures timing differences down to 50 ps and yet consumes only 5 milliwatts. Highly integrated signal processing ASICs are now available that can amplify and shape signals from an array of detectors while consuming power on the order of milliwatts. In addition, technological progess has proceeded apace in the area of detector development, such as radiation-hardened high density CCDs, that are required by Telemachus for survival during the Jupiter gravitational assist as well as high-intensity solar energetic particle events.

Summary: Telemachus is a technologically-ready low-to-medium cost dual-mode mission that addresses basic science questions of the dynamics of (1) the sub-surface, surface, and coronal dynamics of the polar regions of the Sun, and (2) the non-ecliptic solar wind and energetic particles that constitute the global energy and mass sources of the non-thermal plasma that dominates the heliosheath.

Telemachus therefore links, in a single mission, observation of the essential physics of the heliosphere – from below the surface of the Sun out through the solar wind to the termination shock and into the heliosheath. In other words, Telemachus links the heliosphere from its innermost to its outermost boundary with the interstellar medium.

References:

1. Members of the 2003 NASA Sun-Earth Connections Roadmap Team-X



2. Di Mauro, M. P., Helioseismology, *Astophys. Space Sci. Trans.*, **4**, 13-17, 2008.

3. Decker, R. B., S. M. Krimigis, E. C. Roelof, and M. E. Hill, Variations of low-energy ion distributions measured in the heliosheath, PICKUP IONS: THROUGHOUT THE HELIOSPHERE AND BEYOND, AIP Conference Procedings, *in press*, 2010.

4. Richardson, J. D., J. C. Kasper, C. Wang, J. W. Belcher, and A. J. Lazarus, Cool heliosheath plasma and deceleration of the upstream solar wind at the termination shock, *Nature*, **454**, 63-65, 2008.

5. Richardson, J. D., and E. C. Stone, The solar wind in the outer heliosphere, *Space Sci. Rev.*, **143**, 7-20, DOI 10.1007/s 11214-008-9443-z, 2009.

6. Krimigis, S. M., D. G. Mitchell, E. C. Roelof, and R. B. Decker, ENA (E>5 keV) images from Cassini and Voyager "ground truth", PICKUP IONS: THROUGHOUT THE HELIOSPHERE AND BEYOND, AIP Conference Procedings, *in press*, 2010.

7. McComas, D. J, and 35 coauthors, Global Observations of the Interstellar Interaction from the Interstellar Boundary Explorer (IBEX), Science, Volume **326**, Issue 5955, pp. 959-962, 2009.

8. Krimigis, S. M.; Mitchell, D. G.; Roelof, E. C.; Hsieh, K. C.; McComas, D. J., Imaging the Interaction of the Heliosphere with the Interstellar Medium from Saturn with Cassini, Science, Volume **326**, Issue 5955, pp. 971-3, 2009.

9. McComas, D. J., H. A. Elliott, J. T. Gosling, and R. M. Skoug, Ulysses observations of very different heliospheric structure during the declining phase of solar activity cycle 23, *Geophys. Res. Lett.*, **33**(9), CiteID L09102, 2006.

10. Fisk, L. A., and G. Gloeckler, Acceleration of Suprathermal Tails in the Solar Wind, *Astrophys. J.*, **686**(2), 1466-1473.

Determination of how charged interplanetary dust affects the flowing magnetized solar wind

Christopher T. Russell, Gian L. Delzanno, Mihaly Horanyi, Lan K. Jian, Janet G. Luhmann, Ingrid Mann, Robert J. Strangeway, Eberhard Gruen, Zoltan Sternovsky, Ralf Srama, Sascha Kempf, Scott Robertson, Tobin Munsat, Andrew Collette

Every collision of meteoroids and asteroids produces a wide range of sizes of interplanetary dust particles. These particles then follow different evolutionary paths depending on size. Figure 1 shows schematically the evolutionary path from comets, asteroids, Kuiper belt objects, and interstellar dust that lead to eventual removal from the solar system. The particles all become charged and must have some effects on the solar wind that passes by them, in addition to those due to ion and electron-dust collisions. For particles more than about 100 nm in radius, the effects of solar wind interactions on the dust dynamics is typically small. This is not true for the dust approaching a few nanometers in scale; the solar wind accelerates these particles to the solar wind speed (Mann et al. 2007, Czechowski and Mann 2010). The nano-dust was recently observed with plasma wave measurements onboard STEREO (Meyer-Vernet et al., 2009). The nano-dust velocities near 1 AU are far beyond those of the larger "beta-meteoroids" that are repelled from the Sun by

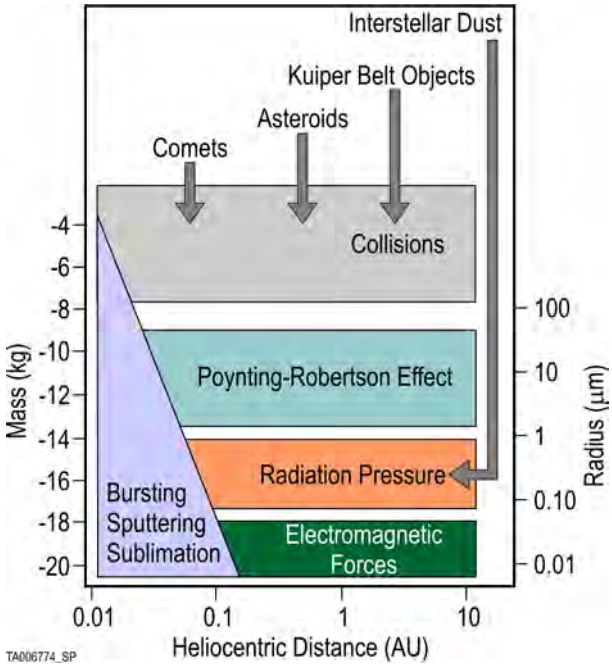


Figure 1. Schematic diagram of evolutionary path followed by asteroidal debris, comets, and dust, as the particles become smaller in size and closer to the Sun and eventually are swept out of the solar system (adapted from Mann et al. 2010).

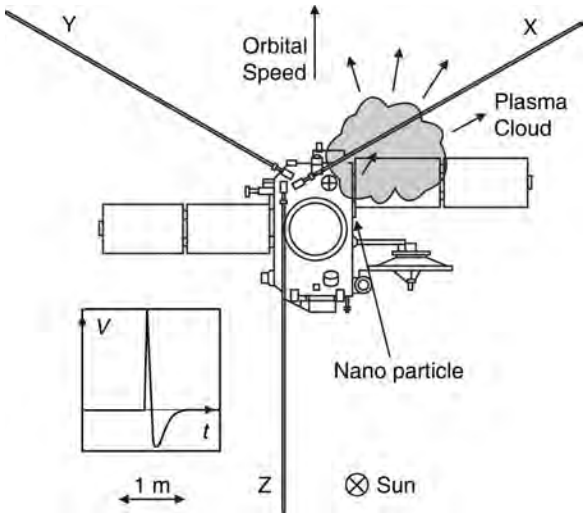


Figure 2a. Schematic diagram of the detection of beta-meteoroids traveling at the speed of the solar wind and creating voltage spikes on the STEREO plasma wave antennas (Meyer-Vernet et al. 2009).

radiation pressure (see Mann et al. 2010a). The nano-dust particles are detected on STEREO by the small puffs of plasma they create. These impacts are not rare. Rather, they are almost ubiquitous in the solar wind. Figure 2 illustrates the detection of the voltage pulses created by these particles and shows their distribution in ecliptic longitude in 2007 on the STEREO-A spacecraft (from Meyer-Vernet et al. 2009). The STEREO/WAVES instrument observes nano-dust as well as the larger "beta-meteoroids", but with distinctly different signals (Meyer-Vernet et al. 2010). Nano-dust particles are also observed with dust instruments initially in the vicinity of Jupiter (Grün et al. 1993). The nano-dust trajectories in the vicinity of Jupiter result from the combination of charging and deflection in the magnetosphere and in the solar wind (Horanyi et al. 1993, Hamilton and Burns 1993). Similarly streams of nano-dust have been recently detected near Saturn (Hsu et al. 2010). While the interaction of nano-dust with the solar wind is most impressive and even trapping of these particles can occur in the inner solar system (Czechowski and Mann 2010), also trajectories of the larger dust particles near the Sun are

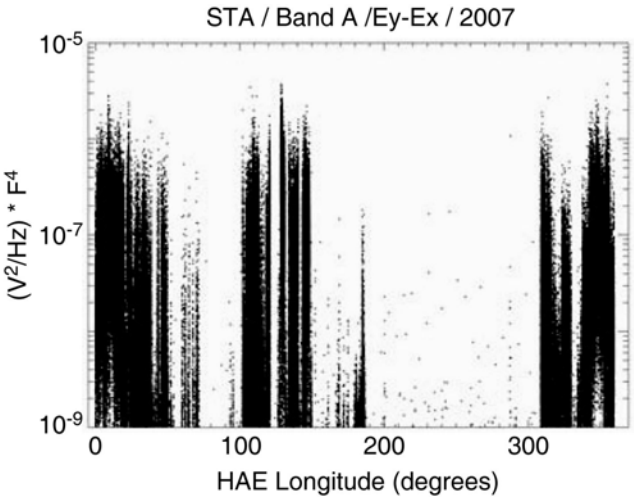


Figure 2b. Spectral power seen by STEREO A in 2007 as a function of ecliptic longitude (Meyer-Vernet et al. 2009).

considerably influenced and for instance modified during coronal mass ejections (Ragot and Kahler 2003).

More measurements by novel dedicated highly sensitive dust analyzers are needed to characterize these enigmatic nanometer-sized particles and measure their chemical composition. The measurements would enable us to identify the source of the dust by comparing their elemental composition with that of larger micrometeoroid particles of cometary and asteroid origin and will reveal interaction of nano-dust with the interplanetary medium by investigating the relation of the dust flux to solar wind and IMF properties. Such instruments are under development but need flights.

A second phenomenon has been attributed to charged dust. This one is a macroscale phenomenon and hence an unexpected interaction. Almost 3 decades ago a rare type of solar wind magnetic disturbance called an interplanetary field enhancement (IFE) was found to be much more frequent at Venus when the asteroid 2201 Oljato was passing between Venus and the Sun (Russell et al., 1983). Later Jones et al. (2003a) found such disturbances in association with comet De Vico, using the Ulysses spacecraft. Figure 3 shows four examples of such events in the solar wind.

The mechanism for the interaction with the solar wind was not understood initially and still needs further study for being fully understood. The Ulysses magnetometer detected ≈ 40 IFE*s* between 1.3 and 5.4 AU that are possibly generated by the solar wind crossing dust trails (Jones et al. 2003). A quantitative description of the formation of these IFEs based on typical dust trail parameters derived from observations was not presented (Mann et al. 2010b). In particular, it is important to determine whether the signals are due to the spacecraft passing through a stationary feature or due to a solar wind disturbance moving across the spacecraft.

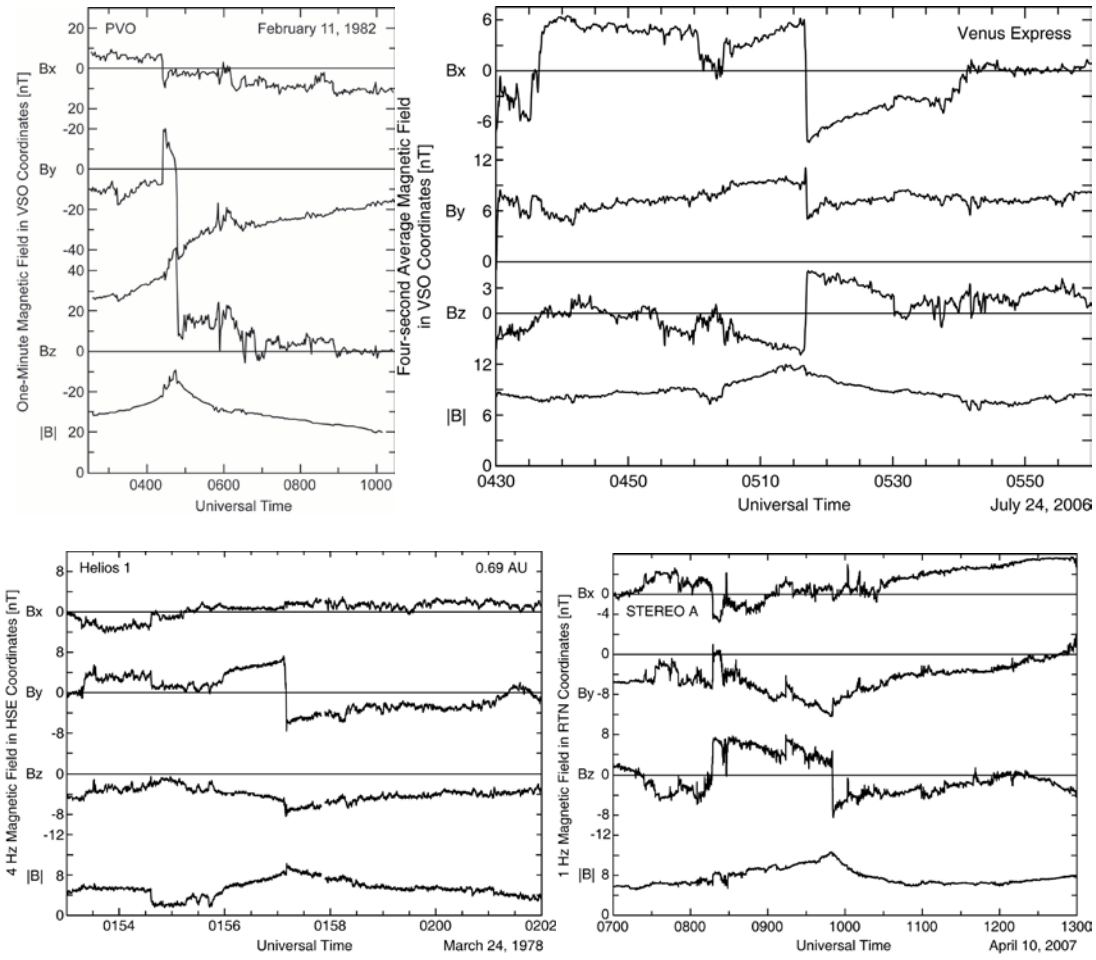


Figure 3. Examples of Interplanetary Field Enhancements (IFEs) seen by Pioneer Venus, Venus Express, Helios 1 and STEREO A. These events are identified by their sharp, cusp-shaped central peak in magnetic field strength, and a sharp change in direction of the magnetic field near the peak. Note the similarity in appearance over a very wide range of event durations, from minutes to many hours.

With the launch of STEREO, a new opportunity for studying IFEs arose when STEREO A, B, and Wind and ACE all saw the same IFE (Russell et al, 2010). Its motion could then be deduced. It was moving at the solar wind speed despite its large scale. The magnetic field pressure gradient associated with such a structure exerts a force on the plasma and on any charged dust. The shape of the structure suggests that the IFE is transferring momentum from the solar wind to the dust. Thus it appears that IFEs are temporal events and not standing structures. If this is true and if all IFEs are moving at the speed of the solar wind then the magnetic pressure force can be used to weigh the solar-wind-speed charged dust. In this scenario, we assume that

the IFE is created by the fine dust component released in a destructive cometary collision. The mass is about 10^7 kg or an object initially about 9m in radius, if all material was converted to nm-scale charged dust. The number of IFEs roughly matches the number of expected cometary collisions in this mass range. This rough estimate however needs to be supported with further studies.

STEREO has thus provided two measurements that show that the charged dust interacts strongly with the solar wind. This latter measurement shows that the dust does have an effect on the interplanetary magnetic field. We would expect this from Newton's third law, because if the dust is accelerated, there must be a force applied to the solar wind in return. Currently we have no theory or observations capable of determining what this feedback on the solar wind does to the plasma and the field. Observationally, we see that an IFE consists of a 10^5-10^6 km magnetic structure of complex nature with a strong cross-flow component of the flow, but we need theoretical and numerical simulation to determine how this structure arises.

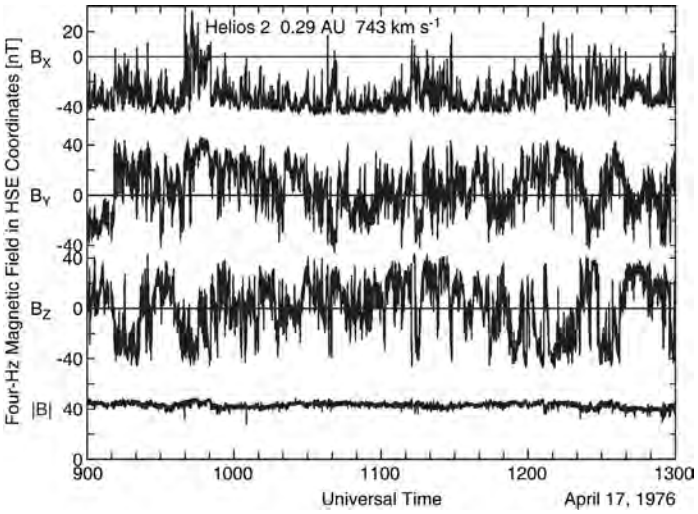


Figure 4. One-quarter second resolution measurements of the magnetic field observed at 0.29 AU on Helios 2 on April 17, 1976. Four hours of measurement are shown; the fluctuations consist of angular deviations of the field from a near radial orientation.

It is possible that a significant fraction of the turbulence in the solar wind comes from this dust-plasma interaction, and not from plasma-only processes. Charged dust particles are being created over a wide range of sizes. It is possible that phenomena such as that seen in Figure 4 are

associated with the smaller dust particle sizes, but a very large number of them. By not considering this ubiquitous source of charged particles in our studies of turbulence and solar wind properties, we may be setting ourselves up for failure in our turbulence studies.

We strongly recommend that the dust and plasma communities be given the opportunity to work jointly on the solar wind turbulence problem through joint funding opportunities and through mission opportunities such as multiscale, multisatellite measurements in the undisturbed solar wind with dust detectors capable of measuring the mass, speed, charge and elemental/chemical composition of the solar wind entrained grains, as well as sensitive plasma and field instruments. This is an extremely difficult problem to tackle because of the range of scales and the multiple non-linear plasma processes that might be involved, but it is critical for our understanding of the solar wind.

References

Czechowski, A., I. Mann (2010), Formation and acceleration of nano-dust in the inner heliosphere, *Astrophys. J.*, 714, 89-99.

Grün, E., H.A. Zook, M. Baguhl, A. Balogh, S. Bame, H. Fechtig, R. Forsyth, M.S. Hanner, M. Horanyi, J. Kissel, B.-A. Lindblad, D. Linkert, G. Linkert, I. Mann, J.A.M. McDonnell, G.E. Morfill, J.L. Phillips, C. Polanskey, G. Schwehm, N. Siddique, P. Staubach, J. Svestka, A. Taylor (1993), Discovery of Jovian dust streams and interstellar grains by the Ulysses spacecraft, *Nature*, 362, 428-430.

Hamilton, D.P., J.A. Burns (1993), Ejection of dust from Jupiter's gossamer ring, *Nature*, 364, 6439, 695-699.

Horanyi, M. G. Morfill, E. Grun (1993), Mechanism for the acceleration and ejection of dust grains from Jupiter's magnetosphere, *Nature*, 363, 6425, 144-146.

Hsu, H.-W., S. Kempf, C.M. Jackman (2010), Observation of Saturnian stream particles in the interplanetary space, *Icarus*, 206, 2, 653-661.

Jones, G.H., A. Balogh, C.T. Russell, and M.K. Dougherty (2003a), Possible distortion of the interplanetary magnetic field by the dust trail of comet 122P/De Vico, *Astrophys. J.*, 597, 61-64.

Jones, G .H., A. Balogh, D. J. McComas, and R. J. MacDowall (2003b), *Icarus*, 166, 297-310.

Mann, I., E. Murad, A. Czechowski (2007), Nanoparticles in the inner solar system, *Planet. Space Sci.*, 55, 1000-1009.

Mann, I., A. Czechowski, N. Meyer-Vernet, A. Zaslavsky, H. Lamy (2010a), Dust in the interplanetary medium, *Plasma Phys. Control*., Fusion 52, in print.

Mann, I., A. Czechowski, N. Meyer-Vernet (2010b), Dust in the interplanetary medium – interactions with the solar wind, Solar Wind 12, AIP CP 1216, eds. M. Maksimovic, K. Issautier, N. Meyer-Vernet, M. Moncuquet, F. Pantelliniet, 491-496.

Meyer-Vernet, N., M. Maksimovic, A. Czechowski, I. Mann, I. Zouganelis,, and K. Goetz, M.K.K., O. C. St. Cyr, J. L. Bougeret, and S. D. Bale (2009), Dust detection by the wave instrument on STEREO: Nanoparticles picked up by the solar wind?, *Sol. Phys.*, 256, 463-474.

Meyer-Vernet N., A. Czechowski, I. Mann, M. Maksimovic, A. Lecacheux, K. Goetz, M.L. Kaiser, O.C. St. Cyr, S.D. Bale, G. Le Chat (2010), Detection of fast nanoparticles in the solar wind, Solar Wind 12, AIP CP 1216, eds. M. Maksimovic, K. Issautier, N. Meyer-Vernet, M. Moncuquet, F. Pantelliniet, 502-505.

Ragot, B. R., S.W. Kahler (2003), Interactions of dust grains with coronal mass ejections and solar cycle variations of the F-coronal brightness, *Astrophys. J.*, 594, 2, 1049-1059.

Russell, C.T., R. Aroian, M. Arghavani, K. Nock (1984), Interplanetary magnetic field enhancements and their association with the asteroid 2201 Oljato, *Science*, 226, 43-45.

Russell, C.T., D.R. Weimer, N. Omidi, L.K. Jian, J.G. Luhmann, R.J. Strangeway (2010), Interplanetary field enhancements travel at the solar wind speed, *Geophys. Res. Lett.*, 37, L07204.

Ground Based Measurements of Galactic and Solar Cosmic Rays

A Heliospheric Physics White Paper for the NRC Decadal Survey

November 12, 2010

J.M. Ryan, J.W. Bieber, E. Cliver, G.A. de Nolfo, P. Evenson, J.R. Jokipii, M.A. Lee, C. Lopate, K. Munakata, I. Richardson, D. Ruffolo, M. Shea, D. Smart, C. Steigies

EXECUTIVE SUMMARY

The global network of ground based neutron monitors and muon telescopes constitutes a unique "instrument" in an armada of particle detection and measurement facilities now in space and to be deployed in the future. This network has a particle energy detection threshold of order 500 MeV, extending up to 50 GeV, with a full-sky field of view continuously making dozens of pointed measurements in key directions. It is sensitive to the highest energy solar cosmic rays and galactic cosmic rays, whose intensities and anisotropies are produced or heavily influenced by solar activity.

It measures the part of the solar energetic particle (SEP) spectrum that is most influenced by the particle acceleration processes taking place close to the Sun. Theses process are affected by coronal turbulence, shock speed and strength and radius, and by turbulence far upstream of the activity. Some particles may be accelerated in flares if the magnetic topology is such that they can escape into interplanetary space to be detected at Earth.

Monitoring of galactic cosmic rays provides information on interplanetary conditions not only at the radius of Earth, but also far out in the heliosphere through variations in the cosmic-ray intensities and anisotropies on time scales from minutes to decades.

Combining these measurements with lower energy measurements from space based instruments provides a comprehensive data set that can be used to investigate a variety of solar and heliospheric phenomena.

The network is now comprised of some new, but mostly aging, neutron monitors and directional muon detectors. Recently, air Čerenkov telescopes, primarily intended for TeV γ–ray astronomy, have been successfully performed differential measurements of the secondary muon component produced by solar and galactic cosmic rays. The network is not functioning as it once did or as envisioned during the International Geophysical Year, 1957-1958. Several key measurement sites have been closed or turned off, and the perceived lack of interest by the US makes foreign sites vulnerable to cuts and closings. There is little coordination between supporting organizations that are also fighting for their own existence New and inexpensive technology has had difficulty making its way into the field to modernize and greatly improve the performance of those stations now operating. Key recommendations to bring the network into a condition to support heliospheric physics include:

- A greater scientific priority and appreciation by agencies with a stake in heliospheric physics,
- increased funding (extremely modest by many standards),
- a national (or better still, international) coordination of the network,
- an organized effort to integrate complementary instruments and
- the training of young researchers.

INTRODUCTION

At energies upwards of 500 MeV, the intensity of solar and Galactic energetic particles becomes so low that very large detector masses are necessary to measure them. In this energy regime ground-based instrumentation provides a highly cost-effective method for making measurements with the required statistical accuracy. Generally these instruments acquire information on primary cosmic rays through measurement of the secondary particles created when a primary strikes a molecule in Earth's atmosphere.

Neutron monitors have been in operation for over fifty years, and provide a vital long-term perspective on solar variations with time scales such as the eleven-year Schwabe cycle or the twenty-two-year Hale cycle. They also make valuable measurements at much shorter time scales, providing

information on relativistic solar particles and on transient, solar-induced variations of Galactic cosmic rays such as Forbush decreases.

The scientific return from neutron monitors is enhanced when they are linked together in coordinated multi-national arrays. Indeed, in most modern applications the "instrument" is the array itself, and not any single detector in it. Analysis of intensities of particles arriving from different directions permits determination of the cosmic ray anisotropy, while combining data from detectors at different geomagnetic cutoffs provides information on the energy spectrum. Coordinated arrays now in operation include the 12-station Spaceship Earth network, which is optimized for measuring the angular distribution of relativistic solar energetic particles (Bieber et al. 2004), and the Neutron Monitor Database (NMDB) recently organized under the auspices of the European Union (Mavromichalaki 2010).

Neutron monitors provide a link between the sub-500 MeV energy regime routinely monitored by spacecraft and the high-energy regime probed by ground-based detectors such as muon detectors and air shower arrays, as well as the new large space-based detectors such as PAMELA and hopefully soon AMS. These higher-energy experiments often have phenomena outside the heliosphere as the primary research objectives, but owing to their extraordinarily high count rates, they also provide important new information on heliospheric phenomena. For example, the Global Muon Detector Network (GMDN) has shed light on the large-scale structure of interplanetary coronal mass ejections (Kuwabara et al. 2009), while *Milagro* and IceTop have yielded detailed information on the energy spectrum of solar energetic particles (Ryan et al. 2003; Abbasi et al. 2008).

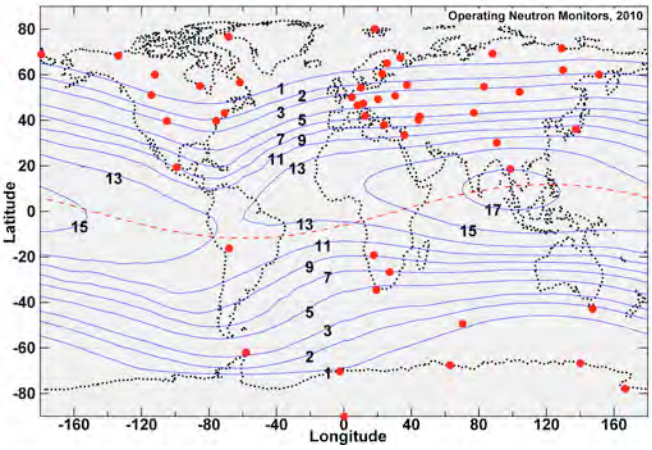


Fig. 1. Neutron monitor sites and geomagnetic cutoffs.

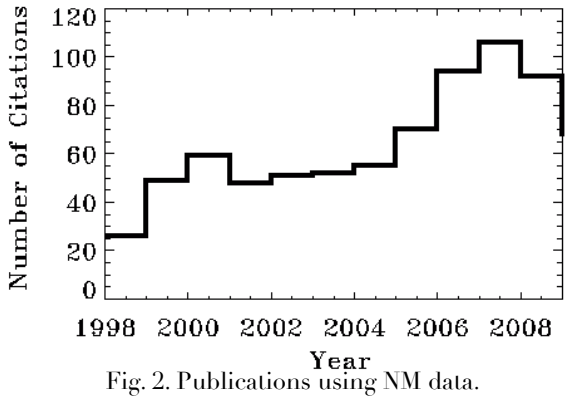


Fig. 2. Publications using NM data.

The status of neutron monitor operations in the past decade is summarized in Fig. 1. There were fourteen neutron monitors operated by U.S. institutions at the beginning of the decade (2001), all but one of which were supported by NSF. Since then, four have closed, including the two longest-operating monitors in the world (Climax and Mt. Washington). Of the ten monitors remaining, only two operate with NSF support. The remainder continue to operate, for the time being, under institutional support.

The present situation with U.S.-operated neutron monitors is not sustainable. The small amount of available federal funding makes it difficult to attract young researchers into the field, and the amount of funding for engineering and data analysis support is too small to sustain the long term health of the instrumentation and the integrity of the data. Further, the amount of federal support currently provided is not commensurate with the high scientific return generated by these instruments.

A cursory search of current literature making reference to neutron monitor measurements is illustrated in Fig. 2. Despite the dwindling support for the network, the importance of the supporting role of neutron measurements in other fields is clear and growing.

SCIENCE OBJECTIVES

The science that can be studied with ground-based instruments is diverse. Some of these objectives are described below.

Solar Energetic Particles

Ground-based instruments are ideal for studying the highest energy solar energetic particles in so-called Ground Level Enhancements (GLE). Neutron monitor observations are critical to resolving the debate as to whether the acceleration of high energy particles takes place at the Sun or in CME-driven shocks; see Klecker et al. (2007) and Cliver (2009) for recent reviews. While it is generally conceded that shocks dominate SEP acceleration in large SEP events at low (~10 MeV) energies, McCracken et al. (2008) and others have argued that at least some GLEs have an early flare component. Continued observations at neutron monitor energies are necessary to determine the peak energy and efficiency of the flare acceleration process. Regardless of the acceleration process, the spectral shape and intensity carries information about the nature of the acceleration environment. Physical solar parameters such as the volume and time available for acceleration, diffusion coefficients and pitch angle distributions manifest themselves in (1) a rolloff in the spectrum at ground-based energies, (2) duration or time evolution of the signal and (3) anisotropy of the high-energy particles. In principle the spectral measurements in this energy range of the rollover shed light on this issue. One other issue that these instruments can potentially address is shock obliquity, i.e., the quasi-parallel vs. the quasi-perpendicular issue. Quasi-perpendicular shocks have difficulty picking up seed particles. but can accelerate them rapidly near the Sun, as opposed to a slower acceleration process in quasi-parallel shocks. Recent impulsive GLEs, such as that of 2005 January 20, press shock acceleration theory, bringing to the fore the shock geometry issue.

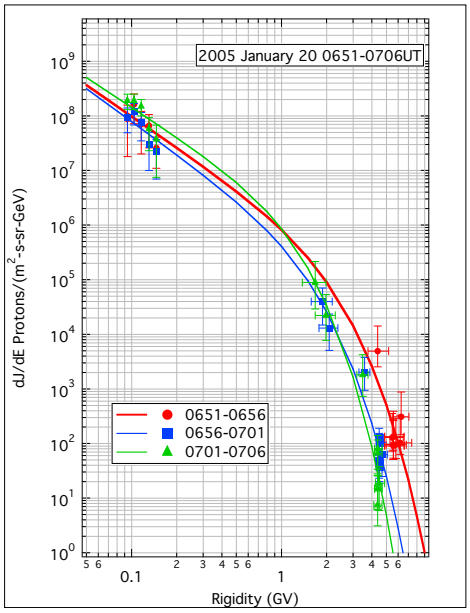


Fig. 3. GLE time-resolved spectra

The development of GLEs also tells us about the generation of turbulence that scatter the particles in interplanetary space or the presence of large magnetic particle-scattering structures. No one station or spacecraft instrument is capable of estimating all of these parameters or investigating these processes associated with GLEs. Shown in Fig. 3, for example, are the time-resolved spectra from the 2005 January 20 GLE measured with the *Milagro* instrument and several NMs, where the initial intensity spike is harder than the longer duration and isotropic component of the GLE.

Solar Flare Neutrons

Neutrons produced in solar flares by GeV protons and ions can reach the Earth and be detected with ground-based instruments. The measurement of the first solar neutron event was in 1982; the largest solar neutron event was measured in 1990. These relatively rare events, coupled with time-associated spacecraft and ground-based measurements of solar phenomena provide a unique opportunity for coordinated study of the solar flare/acceleration processes.

Short-term GCR modulation

Coronal Mass Ejections sweep away galactic cosmic rays (GCR) as they propagate outward from the Sun. The GCR depletion within a CME can be sensed early in the form of a local deficit in the pitch angle distribution of GeV GCRs, providing an indication of an advancing CME. The pitch angle distribution can be sensed in neutron monitors, but most graphically in the anisotropy of the muon intensity at ground level (see Munakata et al. 2000). So-called Halo CMEs are coronal mass ejections directed at the Earth. With no oblique perspective, they are the most difficult to detect optically. However, if the CME is large, the signal of its presence reaches us quickly in the form of the GCR anisotropy. The magnitude of the signal directly reflects the depth of GCR depletion within the CME and, thus, the potential impact of the CME on terrestrial phenomena or spacecraft near Earth. Once

the CME strikes the Earth's environment, all ground-based stations register the GCR depletion with a magnitude that depends on the local geomagnetic cutoff. These events are called Forbush Decreases (FD). The recovery from an FD is the backfilling of the GCR population in the Earth's environment. Its time scale depends on the conditions, not only at Earth, but at large heliospheric distances. When joined with measurements by cosmic-ray instruments in the outer heliosphere, the evolution of the FD recovery (measured, for example, as a function of particle rigidity) can be understood in terms of particle drifts and particle density gradients that are determined by, and hence provide information on GCR diffusion coefficients or the intensity of MHD turbulence. With the recovery measured as a function of rigidity (momentum per unit charge) this can constrain the wave spectrum in k-space at large distances from Earth.

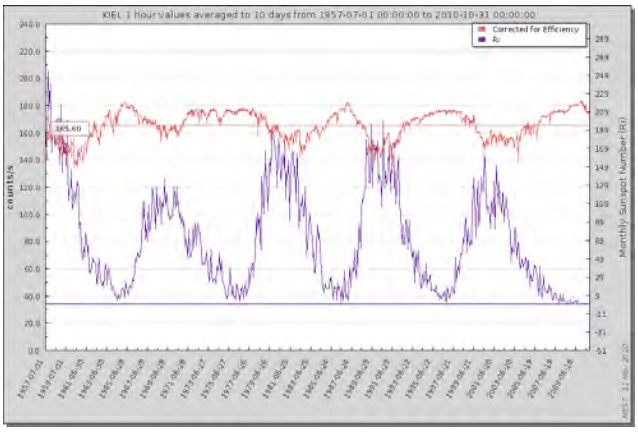


Fig. 4. NM count rate (Kiel) and sunspot number.

Long-term GCR modulation

Galactic cosmic rays at Earth have been monitored and measured over more than five 11-year solar activity cycles and their intensity is inversely related to solar activity. Solar activity is commonly characterized by sunspot number (Fig. 4). It is now understood that conditions in the outer heliosphere have a direct influence on the GCR intensity at Earth. Thus, NM observations provide an estimate of the accumulated activity beyond Earth's orbit. An early and reasonably accurate description of this was modeled by Gleeson and Axford (1968), who quantified the effect of diffusion in the outer heliosphere as a repulsive electrostatic potential. Models have been refined in the intervening years, but are still based on the assumption that conditions in the outer heliosphere determine the GCR intensity near the Earth. Furthermore, an exciting area now is the extent and nature of modulation in the heliosheath including the heliotail. Ground-level detectors can, in principle, provide information on this, including the associated hysteresis and solar-cycle dependence.

Fig. 4 shows alternating flat top vs. spike shaped GCR maxima in successive solar minima. The shape is determined by the 22-year cycle in the polarity of the solar global magnetic field, which affects the direction of GCR drifts in the heliosphere. During periods of high solar activity when the GCR intensity is depressed. Fig. 4 also shows peaks in the neutron monitor rate due to GLEs.

Another interesting feature illustrated in Fig. 4 is the fact that the greatest GCR rate in the last six cycles was observed both on the ground and by spacecraft during the recent deep and prolonged solar minimum (Mewaldt et al. 2010). The long term record in the galactic cosmic flux, combined with ongoing measurements will enable significant improvements in our understanding of the physical processes underlying the modulation parameters. The recent unusual sunspot minimum and its recovery will make possible important new tests of the modulation process in parameter ranges not seen before. A full recovery to more-normal conditions may take years, and proper measurements will enable a more-precise determination of the interplay of different parameters in modulation.

NM measurements of the GCR intensity are also influenced by secular variations in the geomagnetic field. These gradual changes in the local magnetic field are immediately reflected in the GCR count rate.

Historical reconstructions of heliospheric parameters

Long-term reconstruction of the solar wind magnetic field is a key application of the neutron monitor observations that highlights the need for a continuous long-term record of galactic cosmic rays. Neutron monitor observations cover a relatively short interval (1953-present). However, combined with contemporaneous interplanetary magnetic field data, they can be used to calibrate the much longer term (10^4-10^5 years) records of galactic cosmic ray intensity inferred from concentrations of cosmogenic nuclei obtained from tree rings (^{14}C) and ice cores (^{10}Be) (e.g., Steinhilber et al., 2010) which can then be

used in turn to infer the variation of the solar wind magnetic field over the millennia. Such information allows one to probe the behavior of the solar dynamo and to better understand sunspot droughts such as the Maunder Minima and the role of the Sun in climate change. However, the 50-year overlap of the neutron monitor and ^{10}Be records is too short to capture the complexity of the factors involved in transforming the measured concentrations of cosmogenic nuclei into reliable B measurements. Recently, geomagnetic measurements extending into the mid-19th century have been used for this purpose, but the magnetic field intensities obtained from geomagnetic and ice core measurements diverge for the last decades of the 19th century, underscoring the need for extending the period of neutron monitor observations.

INSTRUMENTS

Muon Telescopes

Cosmic rays at the top of the atmosphere produce muons as well as neutrons. Muons are registered with ground-based or underground hodoscopes. Two types of muon detectors are functioning, the scintillation detector (SD) and the proportional counter tube detector (PCTD). The SD consists of horizontal layers of plastic scintillators, viewed by photomultiplier tubes. By counting pulses of the coincidences between a pair of detectors in different layers, one has the rate of muons from the corresponding incident direction. A typical field of view (FOV) of the instrument is 30°. A PCTD instrument instead uses long proportional counter tubes with an overlying Pb layer to absorb the soft component radiation. Muon recording is triggered by the coincidence of pulses from different layers with the incident direction identified by the x-y locations of the interactions. A typical FOV in this detector is 7° half angle.

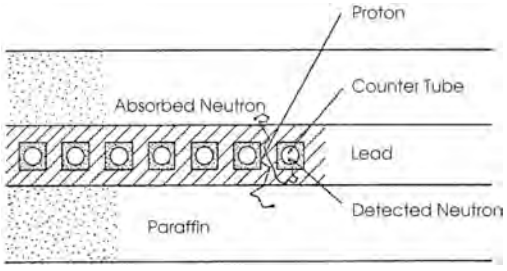


Fig. 5. Schematic of IGY neutron monitor.

Neutron Monitors

The neutron monitor was invented by University of Chicago Professor John Simpson in 1948 (Simpson 2000). There have been various types of neutron monitors, notably the "IGY-type" monitors (Fig. 5)deployed around the world during the 1957-1958 International Geopysical Year (IGY) and the much larger "NM64" monitors, "supermonitors." All monitors, however, employ the same measurement strategy: There is an outer superstructure of polyethylene and lead ingeniously designed to suppress local neutrons, while amplifying the signal from cosmic secondary neutrons. At the heart of a neutron monitor is a proportional counter. The early Simpson monitors detected neutrons through BF_3 gas enriched in the isotope ^{10}B, which produces a signal via the reaction ($n + {}^{10}B \rightarrow \alpha + {}^{7}Li$). Recent proportional counters use ^{3}He which produces a signal via the reaction ($n + {}^{3}He \rightarrow {}^{3}H + p$).

Solar Neutron Telescopes

Over the past ten years, an array of Solar Neutron Telescopes has been deployed near existing NMs to augment the observations performed with those NMs (Valdes-Galicia et al. 2008).

TeV Čerenkov Detectors

The IceTop air shower array is now nearing completion at the South Pole as the surface component of the IceCube neutrino telescope (Achterberg et al. 2006). IceTop will have approximately 500 m^2 of ice Čerenkov collecting area arranged in an array of 80 stations on a 125 m triangular grid with the primary objective of studying air showers from 1 PeV to 1 EeV. Each station consists of two, 2-m ⌀ tanks containing ice to a depth of 90 cm. Light from fast particles is measured with photomultiplier tubes.

Due to the high altitude (2835m) and the nearly zero geomagnetic cutoff at the South Pole, secondary particle spectra at the detector retain a significant amount of information on the spectra of the primary particles. The signal amplitude carries information about the composition and spectra of the incident particles, albeit integrated over broad regions of the spectrum. IceTop can measure the energy spectrum of these particles in the multi-GeV regime from a single detector with a well defined viewing (asymptotic) direction (Abbasi et al. 2008). IceTop has a counting rate approximately 10^4 times that of a neutron monitor. With its inherent energy resolution this permits study of Forbush Decreases,

Corotating Interaction Regions (CIR), and other transients in the cosmic-ray flux with unprecedented resolution (Kuwabara et al. 2007). Ice Top, with other instruments at the South Pole, will permit the first ground-based detection of a possible component of nuclei heavier than protons in solar particles in the multi-GeV regime.

The *Milagro* instrument, or more recently HAWC, operated in a similar way to that of Ice Top. These instruments detect(ed) the muon-Čerenkov light produced in a water pond or in tanks. *Milagro*, now de-commissioned, operated near Los Alamos, but HAWC (High Altitude Water Čerenkov detector), in a small version, has begun taking data on Sierra Negra in Mexico. At low-latitude, HAWC, once fully populated, will survey the sky for TeV γ-ray sources and, in the process be sensitive to secondary muons in GLEs. At low latitude, it will be sensitive to direct solar neutrons >500 MeV from solar flares. The utility of this Čerenkov technique was illustrated in Fig. 3, where *Milagro* data were used with NM data to produce a time-resolved spectrum of the 2005 January 20 GLE (Morgan 2010). The high count rates of Ice Top and HAWC permit studying fast temporal structure in GLEs.

International participation

By its very nature international participation in this system is critical. Faltering US support for the system is interpreted elsewhere as a low US priority, thereby encouraging cuts in support by foreign governments, with the entire system suffering the consequence. One bright point has been European support for an integrated database for the NM system. The European Commission has funded the creation of the real-time database for high resolution neutron monitor measurements (NMDB) during 2008 and 2009 as a scientific data repository in its 7th framework programme (FP7). Currently the Neutron Monitor database NMDB is accumulating data from more than twenty ground-based cosmic ray stations, most of which are operated by European institutions. The NMDB makes the data from all participating stations available in real-time in a common format. In addition to real-time data, NMDB also archives data from all participating stations.

The NMDB data coverage is currently limited to Europe and neighboring countries. Space Weather applications, like GLE and CME alerts, require a global coverage to achieve their full potential. Currently, the NMDB has no data coverage from the American sector, which reduces the reliability of the GLE alert and other applications. The inclusion of neutron monitor data from the Western Hemisphere will significantly improve the value and utility of the database. The implementation of this database is significant in that it demonstrates untapped willingness and enthusiasm for fully integrating a variety of international ground-based instruments into a single system. Until recently, this has been entirely cooperative, but informal arrangement with minimal coordination of data formats and access to facilitate research.

SUMMARY AND RECOMMENDATIONS

Measurements of the cosmic radiation have been made continuously for over half a century. During this time the community has made significant advances in understanding the nature of galactic cosmic radiation, solar particle events and processes and structure of the interplanetary medium. The advent of space measurements has augmented the ground-based measurements, leading not only to greater knowledge of our Earth's environment, but also motivating future research. It is not possible to address new questions that are raised without continuous measurements from a set of stable and robust monitors.

The research addressed with this ground-based network is extensive and diverse, affecting research in many fields, as compared to spacecraft experiments that are focused in their objectives by necessity. Compared to space experiments, the scientific return of the ground-based network is a bargain. For example, with modest money for supplies, the entire US network of neutron monitors could be maintained by little more than one full time engineer/technician. Investments for upgrading stations would reap benefits for decades. Although on one scale, the costs for these data are modest, they are large for individual researcher grants, from which one supports students. The de-coupling the network costs from the science-based grants allows researchers to provide both service and data to the community, yet conduct his or her own research, training the next generation of cosmic-ray scientists.

In addition to the many fundamental scientific achievements resulting from basic ground-based cosmic-ray detectors, there is a practical aspect in understanding Space Weather phenomena and its effects on national assets. We realize the important role that the neutron-monitor measurements play in

determining, understanding and predicting the Earth's space environment with the concomitant implications for climate, e.g., linkages between cosmic-ray intensity and cloud cover. The measurements from neutron monitors and the possibilities for long-term comparisons are a critical component in understanding our whole environment in a broad sense.

To continue this research the community must train young scientists to carry on the work of the cosmic-ray pioneers, many of whom have retired or left us. To achieve this, the present cadre of scientists needs sufficient funding to maintain the present cosmic-ray detectors and analyze the data. We recommend that there should be increased scientific priority on these ground-based measurements and an appreciation by agencies having a stake in various aspects of Earth's environment. Increased funding and interest from US National Agencies would have a strong and immediate effect on the scientific vitality of this research.

REFERENCES

Abbasi, R., et al., Solar energetic particle spectrum on 2006 December 13 determined by IceTop," Astrophys. J. (Lett.), 689, L65-L68, doi:10.1086/595679, 2008.

Achterberg, A. et al., "First year performance of the IceCube neutrino telescope", Astroparticle Physics, 26, 155-173, 10.1016/j.astropartphys.2006.06.007, 2006

Bieber, J. W. et al., Spaceship Earth observations of the Easter 2001 solar particle event, Astrophys. J. (Lett.), 601, L103-L106, 2004.

E.W. Cliver, in Universal Heliospherical Processes, ed. by N. Gopalswamy, D. F. Webb (IAU Symp. 257, Cambridge Univ. Press), pp. 401-412, 2009.

Gleeson, L. and W.I. Axford, 1968, "Solar Modulation of Galactic Cosmic Rays," Astrophys. J., vol. 154, p. 1011.

Kuwabara, T., J.W. Bieber, and R. Pyle, "Heliospheric Physics with IceTop", Proc. 30th Int. Cosmic Ray Conf. (Merida), paper 0729, 2007.

Kuwabara, T. et al., Determination of interplanetary coronal mass ejection geometry and orientation from ground-based observations of Galactic cosmic rays," J. Geophys. Res., 114, A05109, doi: 10.1029/2008JA013717, 2009.

Mavromichalaki, H., Worldwide integration of neutron monitors, Eos, 91, 305-306, 2010.

Selected Publications Using Neutron Monitor Data (includes PI institutions and external institutions)

K.G. McCracken, H. Moraal and P.H. Stoker, J. Geophys. Res., 113, CiteID A12101, doi: 10.1029/2007JA012829, 2008.

R.A. Mewaldt, A.J. Davis, K.A. Lave, R.A. Leske, E.C. Stone, M.E. Wiedenbeck, W.R. Binns, E.R. Christian, A.C. Cummings, G.A. deNolfo, M.H. Israel, A.W. Labrador, & T.T. vonRosenvinge, (2010) Record-Setting Cosmic-Ray Intensities in 2009 and 2010. Astrophys. J. Lett. 723, L1-L6.

Morgan, T., "Measurements of the 2001 April 15 and 2005 January 20 Ground-Level Enhancements by the Milagro water Čerenkov Detector," Ph.D. thesis, University of New Hampshire, 2010.

Munakata et al., 2000, "Precursors of geomagnetic storms observed by the muon detector network," Journal of Geophys. Res., vol. 105, pp. 27457-27468.

B. Klecker, E. Möbius and M.A. Popecki, Ionic Charge States of Solar Energetic Particles: A Clue to the Source, Space Sci. Rev., 130, 273-282, 2007.

Ryan et al., 2005, "Ground Level Events Measured with Milagro," Proc. ICRC, Pune, India, Volume 1, p. 245.

Simpson, J. A., The cosmic ray nucleonic component: The invention and scientific uses of the neutron monitor, Space Science Reviews, 93: 11 32. doi:10.1023/A:1026567706183, 2000.

Steinhilber, F. et al., Interplanetary magnetic field during the past 9300 years inferred from cosmogenic radionuclides, J. Geophys. Res., 115, A01104, doi:10.1029/2009JA014193, 2010.

Svalgaard, L. and Cliver, E.W., Heliospheric Magnetic Field 1835-2009, J. Geophys. Res., 114, CiteID A09111, doi:10.1029/2009JA015069, 2009.

Valdes-Galicia, J. et al., "The Solar Neutron Telescope Worldwide Network as a Tool to Study Particle Acceleration at the Sun," 37th COSPAR Scientific Assembly. Held 13-20 July 2008, in Montréal, Canada., p.3272, 2008.

The Abundance of Deuterium and He3 in the Solar Wind

Frank Scherb

Physics Dept., University of Wisconsin, Madison, WI 53706
scherb@physics.wisc.edu

Abstract

The relative abundance of deuterium (D) in the solar atmosphere is not known. D is not only destroyed in stars, it is also synthesized in the atmospheres of active stars (Prodanovic and Fields 2003). In several cases, production of D in the sun has been detected when solar flares occur, using both energetic particle measurements (Anglin 1975) and by detection of 2.223 MeV gamma rays emitted by D (Terekhov et al. 1996; Shih et al. 2009). We describe a project to measure the abundance of deuterium in the solar wind, and to monitor its evolution during a several-year period. The instrument consists of two grids, a tritium target, and semiconductor particle detectors. The grids, which are hemispherical and concentric, accelerate the incident solar wind ions using a potential difference on the order of ~80 to 100 kV and concentrate the ions on the tritium target. A fraction of the solar wind deuterons thus accelerated interact with the target to produce 3.6 MeV alpha particles, some of which are recorded by adjacent semiconductor detectors. A similar instrument was successfully tested in space in 1975 in order to observe positive auroral ions in a hydrogen aurora.

The Experiment

While in situ space measurements have shown the possibility of detecting high- energy (~a few MeV) deuterons linked to solar flares (Anglin 1975), the same is not true for less energetic solar wind deuterons (~< 10 keV). Therefore, we describe a project that uses a new method for detecting solar wind deuterium. This method is based on the nuclear reaction between D and tritium:

$D + T \rightarrow He4 + n + 17.6$ MeV.

The neutron and alpha particle that are produced have energies of 14.0 and 3.6 MeV respectively. **Figure 1** shows that the cross section of this reaction possesses a maximum for deuterons with energies around100 keV. Since the energy of deuterons in the solar wind is

below 10 keV, an instrument designed to use the reaction D + T must accelerate the deuterons to energies greater than 50 keV, in order to attain appreciable values for the cross section (Floyd and Scherb 1966). To accelerate the D particles of the solar wind, we generate an electric field between two hemispherical grids (Reynolds and Scherb 1967) by applying a potential difference on the order of 100 kV. The high- voltage accelerator was successfully tested in space on a high-altitude Javelin sounding rocket in 1975 in order to measure energetic auroral ions in the upper atmosphere of the Earth (Lynch et al. 1976; Lynch et al. 1977).

He3 ions in the solar wind will also react with the tritium target. The He3 + T reaction has three possible final states:

(a) He4 + D +14.3 MeV (43%)

(b) He4 + n + p +12.1 MeV (51%)

(c) He4 + p + n + 14.1 MeV (6%)

The He4 and D resulting from reaction (a) have unique energies of 4.75 and 9.55 MeV, respectively. Reactions (b) and (c) produce protons and alphas that do not have unique energies. Therefore, some of these particles will have energies near enough to 3.6 MeV to be confused with alphas from the D + T reaction. **Figure 2** , however, shows that reaction (a) can be used to resolve the ambiguity, while also yielding information on the abundance of He3 ions in the solar wind.

Measurements carried out with a crude, early laboratory model of the experiment indicated that for a D/H ratio in the solar wind of 10^(-5), and an instrument entrance aperture of 10^(3) cm^(2), the deuteron count rate would be about one count per sec in the slow solar wind (Floyd and Scherb 1966).

A campaign of observations spanning several years would not only produce a measure of the relative abundance of atmospheric D free of bias due to sporadic solar activity, but would also permit a better characterization of the production mechanisms of D in flares.

In **Figure 3**, a sketch of the instrument shows the solar wind particle accelerator, its position relative to the tritium target, and one of several alpha particle detectors. Other particle detectors, optimized to make direct measurements of solar wind composition, would be integrated into the instrument.

References

Anglin, J. D. 1975, ApJ, 198,733

Floyd, F. W., & Scherb, F. 1966, IEEE Trans., NS-13, 18

Lynch, J., Leach, R., Pulliam, D., & Scherb, F. 1977, JGR, 82, 1951

Lynch, J., Pulliam, D., Leach, R., & Scherb, F. 1976, JGR, 81, 1264

Prodanovic, T., & Fields, B. D. 2003, ApJ, 597, 48

Reynolds, R., & Scherb, F. 1967, Rev. Sci. Instr., 38, 348

Shih, A. Y., Lin, R. P., & Smith, D. M. 2009, ApJ, 698, L152

Terekhov, O. V., Sunyaev, R. A., Tkachenko, A.Yu., Denisenko, D.V., Kuznetsov, A.V., Barat, C., Dezalay, J.-P., & Talon, R. 1996, Astron. Lett., 22, 143

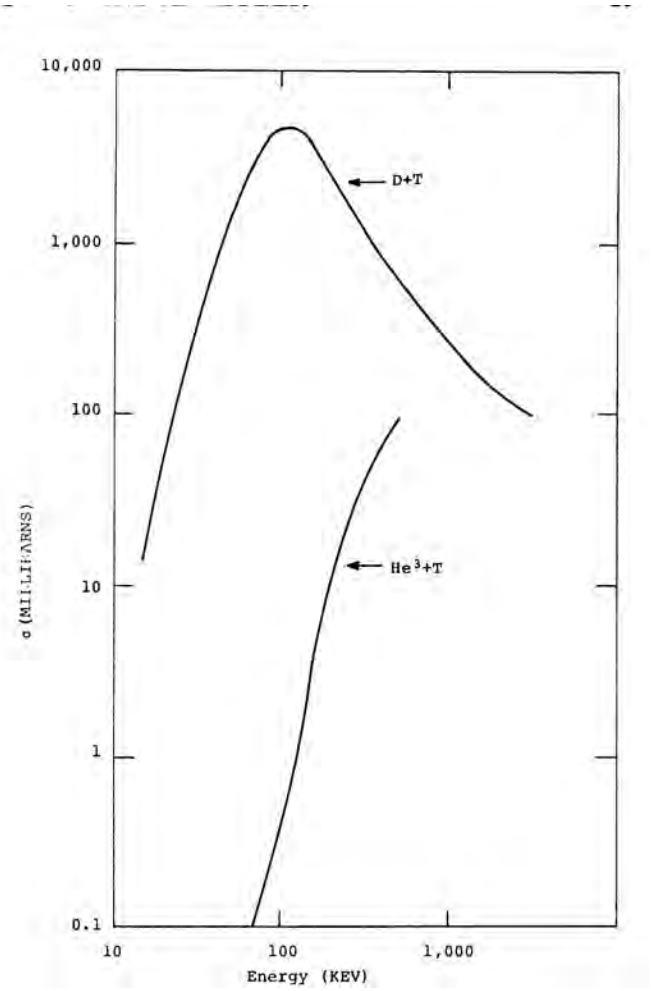


Fig. 1. Nuclear reaction cross sections

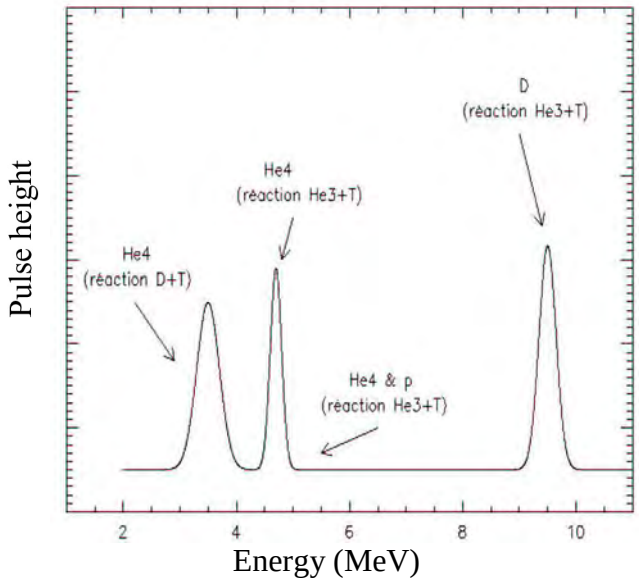


Fig. 2. Schematic pulse height spectrum of deuterium detector, showing effects of He3.

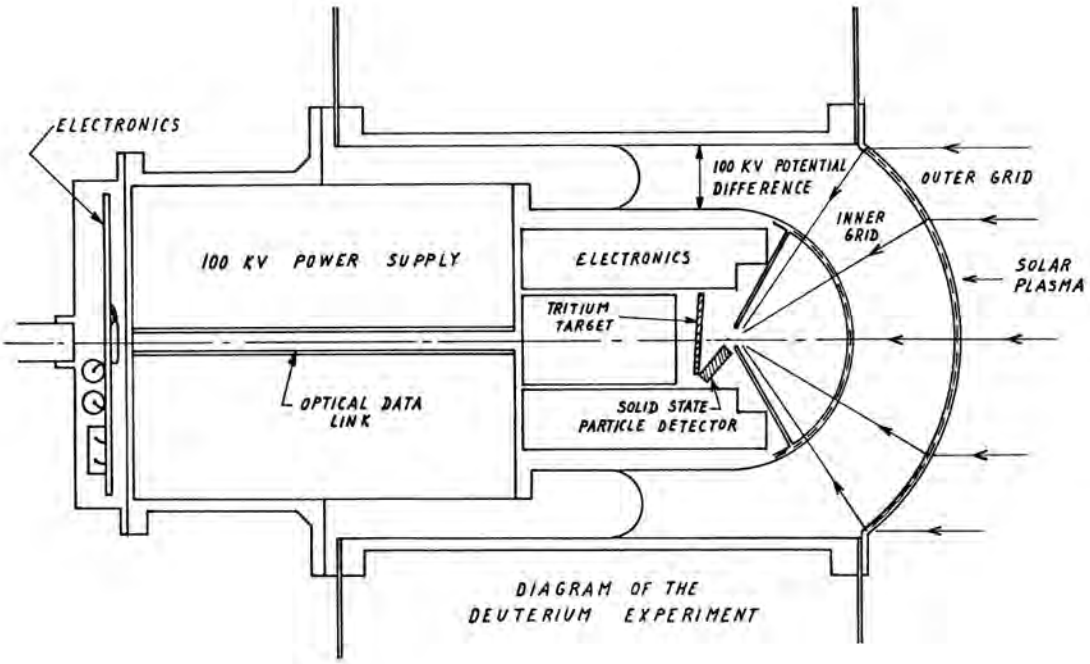


Fig. 3. Accelerator-detector instrument

A Science White Paper in response to the 2013-2022 Heliophysics Decadal Survey (theme of "solar and heliospheric physics")

The solar magnetic dynamo and its role in the formation and evolution of the Sun, in the habitability of its planets, and in space weather around Earth.

2010/11/11

Karel Schrijver (LMATC), Ken Carpenter (NASA-GSFC), Margarita Karovska (CfA), Tom Ayres (CASA), Gibor Basri (UC Berkeley), Benjamin Brown (U. Wisconsin), Joergen Christensen-Dalsgaard (Univ. Aarhus), Andrea Dupree (CfA), Ed Guinan (Villanova Univ.), Moira Jardine (Univ. St. Andrews), Mark Miesch (UCAR), Alexei Pevtsov (NSO), Matthias Rempel (UCAR), Phil Scherrer (Stanford Univ.), Sami Solanki (Max Planck Inst.), Klaus Strassmeier (AIP), and Fred Walter (SUNY).

For more Information, please contact:

Dr. Carolus J. Schrijver
Solar and Astrophysics Laboratory
Lockheed Martin Advanced Technology Center
3251 Hanover Street, Dept. ADBS / Bldg. 252
Palo Alto, CA 94304-1191, USA
phone: (650) 424 – 2907
fax: (650) 354 - 5445
email: schrijver@lmsal.com

The solar magnetic dynamo and its role in the formation and evolution of the Sun, in the habitability of its planets, and in space weather around Earth.

Summary

A magnetic dynamo will operate in the Sun throughout its lifetime. It regulated the formation of the Sun and our planetary system, it affects the habitability of the planets, and it determines the space weather around them. Dynamos also operate in objects such as planets, accretion disks, and active galactic nuclei. *At present, there is no comprehensive dynamo theory from which we can derive the strength, the patterns, or the temporal behavior of the Sun's magnetic field in past, present, or future.* Consequently, we cannot usefully forecast long-term space weather or reliably model the long-term effects of solar magnetic variability on the atmospheres of the planets, including Earth. *Understanding the complex of non-linear couplings in a stellar dynamo requires that we combine numerical studies and theory with observations of the evolving surface field of Sun and stars, of the average properties of magnetic fields in a large sample of stars, and of solar and stellar internal flows.*

Numerical modeling has taught us valuable lessons about dynamos, most importantly that they are highly non-linear processes that couple a vast range of scales within stellar convection zones. Our computer models are therefore of necessity simplified, and hence require observational guidance from stars with a variety of properties. Observations of light-curves and the application of (Zeeman) Doppler imaging continue to give us some of that information, but more is needed in the coming decade. We can expect exciting advances from high-fidelity numerical modeling, high-resolution helioseismic studies (with, e.g., SDO), comprehensive studies of the solar atmosphere (e.g., SDO, Hinode, STEREO, IRIS, and with ATST, FASR, and other ground-based facilities), and stellar global variability and internal structure (e.g., Kepler, COROT, ground-based observatories). We should also image surface activity and map interior flows for Sun-like stars. Together, these opportunities will rapidly advance our understanding of the impact of magnetism of the Sun on Earth and its inhabitants.

A strategy to understand the origins and effects of the solar dynamo is inherently interdisciplinary, and should include (1) continued solar observations of surface activity and interior flows, (2) utilization of long-term records that exist in, e.g., ice formations and other deposits on Earth, (3) high-performance numerical modeling using tens of thousands of processing cores, (4) observation and analysis of long-term variability in the magnetic activity of hundreds of stars, and (5) UV/optical interferometry (with, e.g., the Stellar Imager[20] or Luciola[21]) for surface activity mapping and resolved asteroseismology to measure internal differential rotation. Anticipating much input to the Decadal Survey on the solar aspects of this problem, we emphasize complementary stellar studies, realizing that solar constraints are critical.

1 The Sun and other cool main-sequence stars

We know that the mean solar activity level matches that of similar G2V stars, but we have yet to understand the observed dependence of stellar activity on rotation, and phenomena such as dynamo saturation in young, rapidly-rotating stars or the apparent enhancement of activity in tidally-interacting binary stars[17]. Several solar and heliospheric activity indices have been identified as somewhat useful to forecast subsequent magnetic cycles, but these correlations are as poor as they are poorly understood; the wide range of forecasts for the present sunspot cycle demonstrates the lack of predictive value of most (or all) of the methods[11]. A major problem is that *the validation of a successful cycle forecast would take decades if only the recent Sun were used as a benchmark: historical solar records and stellar constraints are essential!*

Stellar observations revealed that solar-like dynamos require convection immediately below the stellar surface, with rotation acting as a catalyst that strengthens the dynamo, at least up to a

point of saturation. The helioseismic determination of the solar rotation with depth and latitude guided theorists to models in which the rotational shear layer (the tachocline) below the convective envelope is crucial as a reservoir for field even as it is stretched and strengthened by latitudinal differential rotation. Observations reveal a pattern change in stellar activity across the boundary where stars become fully convective, but nonetheless activity continues to depend on rotation and decreasing with age, dynamos function in T Tauri stars and (tidally locked) rapidly rotating giants with very deep convective envelopes; recent models suggest that the buoyancy of the field can be countered by convective downdrafts[2]. Dynamos appear also to function in brown dwarfs and may even aid global mixing in asymptotic and red giant branch stars[5].

The magnetic activity of cool stars spans an enormous range: the most active stars have more than half of a hemisphere covered by dark starspots (Fig. 1), with at most 2% for the Sun. In inactive stars, the coverage by spots is too small to be measured by current instruments, but the emissions from the hot outer atmosphere can be used as a measure of activity. At the extreme lower end of activity, entire coronae of inactive giant stars are dimmer than solar coronal holes[17].

Numerical models are rapidly advancing with increasing computational capabilities. They provide insight into the magnetoconvective couplings near the surface[18], multi-scale compressible convective envelopes[9], and convection and dynamo action in fully convective stars, albeit thus far at low Reynolds numbers[4]. These models are far from capturing the full range of MHD and plasma interactions from solar tachocline to atmosphere. The nonlinear differential equations that couple turbulent convection and magnetism to yield dynamo action cannot be solved analytically. Nor can the dynamo of the Sun and other stars be simulated directly, as the solar convection zone is highly turbulent. In the high magnetic Reynolds number environment of the solar interior, the magnetic field is highly intermittent with length scales down to ~1 km. The volume of the solar convection zone would require ~10^{18} gridpoints, evolved for millions of timesteps to achieve a full cycle of solar activity. This is more than a factor of a billion beyond present computational resources, and thus both analytical and numerical treatments must make approximations that simplify much of the physics.

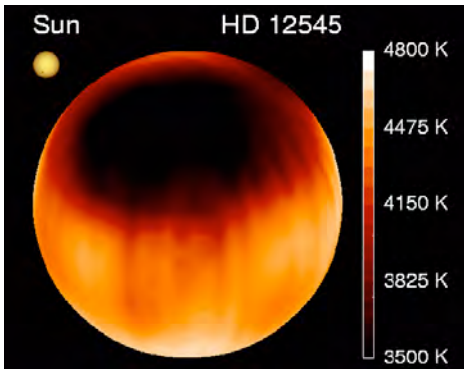


Fig. 1 Doppler-imaged temperature map of XX Tri (K0III). The 24-d tidally-locked binary component shows a large, persistent high-latitude starspot (or spot cluster) vastly larger than the entire Sun. Such high-latitude features are common among rapid rotators. [19]

Most dynamo concepts are kinematic in nature. A self-consistent model needs to include the interaction between the field and the flows. Some semi-analytical mean-field models have been tested that include this interaction (often captured by ad-hoc parameterizations), but comprehensive numerical experiments that include these couplings are in their infancy (Fig. 2). Thus observations are essential to validate the approximations made: comparisons of observed stellar dynamo patterns for different activity levels with those resulting from models is essential to guide dynamo research. Unfortunately, we know essentially nothing about these dynamo patterns, i.e., the stellar butterfly diagrams or - on smaller scales - the spectrum of emerging bipoles, their orientations, or their clustering properties for any star other than the Sun. For main-sequence stars with moderate to low rotation rates, activity tends to be cyclic, but no clear trend of cycle period with stellar parameters has been found, although there are hints of relationships between cycle period, rotation period, and the time scale for deep convection[14]. For truly active stars, various variability patterns exist, but generally no unambiguous activity cycle is seen. In this context, it is interesting to add that the moderately and very active stars never reach low

states of activity, in contrast to the Sun during its cycle minima. Perhaps the dynamos of these active stars differ substantially from the solar dynamo, with either no organized butterfly diagram or with overlapping starspot cycles that combine to reduce the overall modulation.

Doppler images of rapidly rotating stars revealed other surprises: starspot clusters extending to the poles (Fig. 1). Surface-flux dispersal models need very strong meridional advection to produce mixed-polarity regions at high latitudes even if flux emerges at mid-latitudes under the influence of the Coriolis force[8], yet state-of-the art models for convective envelopes suggest that these large-scale flows weaken rather than strengthen with increasing rotation rate[3,9].

Given the complexity and non-linearity of dynamos, improved observational guidance is sorely needed. For example, detailed magnetographic observations over three sunspot cycles have given us much more insight into the behavior of the Sun's magnetic field at and above the solar surface: we understand the roles of differential rotation, meridional flow, and supergranular dispersal quite well qualitatively, and even quantitatively on time scales up to several years. For longer time scales, understanding of the behavior of the large-scale solar field from cycle to cycle still eludes us: the global dipolar field, for example, does not show the anticipated hysteresis from one cycle to the next[15], and we do not know if this is related to the modulated meridional flow[7], to subsurface transport[1], to variations in tilt angles of active regions, or to some as yet unimagined effect. Why the solar dipole and heliospheric field in the recent minimum are only half as strong as during earlier cycle minima remains a puzzle[16]. We also still do not understand the role of helicity[12].

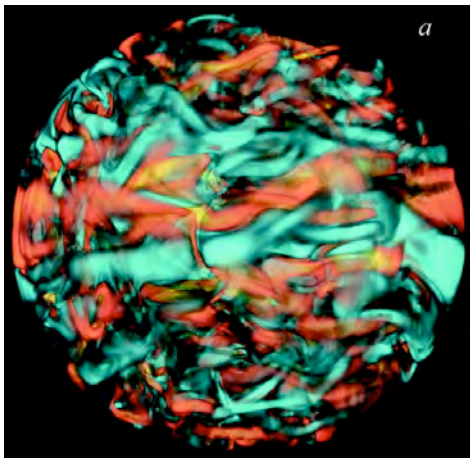


Fig. 2 Volume rendering of the toroidal magnetic field (red: eastward; blue: westward) from compressible MHD simulations of convection. This model lacks a tachocline and a realistic treatment of the near-surface layers. Strong, dynamic field is generated, but a cycling global dynamo is not excited in this simplified model. From Brun et al. (2004, ApJ 614, 1073).

To understand dynamos and the environments in which they operate, we must also study astrophysical dynamos that differ from the Sun. For example, close binaries should be studied because tidal interaction appears to strengthen dynamos. Fully convective dwarf stars, stars with shallow convective envelopes, and evolved stars also set valuable constraints.

2 Star formation, pre-main sequence stars, planetary systems and planetary atmospheres

In the initial phases of the contraction of the cloud out of which the Sun and the planets formed, magnetic fields were instrumental in transporting most of the angular momentum out of the core regions of the cloud. Without the expulsion of the bulk of that initial angular momentum, no star could have formed because centrifugal forces would have exceeded the pull of the gravitational field. Once a proto-star and a surrounding extended gaseous disk form, the star's strong magnetic field couples to the inner domains of the gaseous disk that surrounds a star for millions of years before the planets form. The star continues to accumulate mass from the disk for some time, but the disk cannot extend to the stellar surface: the magnetic field of the rapidly rotating star sweeps through the inner disk region, allowing matter to accrete only along the field. Accretion thus occurs in evolving columns connecting the disk to patches on the stellar surface. Where the material cannot reach the surface, the gas pressure aided by the magnetic field channels the material into jets of gas, shooting away from the poles of the star. More material is lost in a magnetically powered stellar wind. This balance in which matter moves toward and away from the star at different locations, affects the disk's properties. That, in turn, likely affects

how and where planets form, and what their masses and atmospheres are. The details of this scenario, and the precise role of the magnetic field in it, remain shrouded, because present-day telescopes cannot image the inner part of such disks around very young stars, and appropriate dynamo models are yet to be developed. The stellar magnetic field appears to be key to the rotational evolution of the system, however, as those stars with discs typically rotate more slowly than those without; moreover, the discs may modify coronal activity by truncating the stellar corona while disk accretion may be responsible for the large flares seen on these young stars[6].

Stellar activity also affects the evolution of planetary atmospheres, both by (X)UV irradiation and through the interaction with the stellar wind[10]. *The theories of stellar and planetary dynamos come together in the interaction of the magnetized stellar wind with the planetary magnetic field.* This determines the ablation of a planet's atmosphere by setting the distance to within which the magnetopause can come down to the planet, with 'auroral winds' flowing down the open field around the magnetotail, with (X)UV radiation and energetic particles energizing the highest layers of the planet's atmosphere, and charge-transfer collisions allowing neutral particles to stream away, eventually resulting in distinct atmospheres for planets as similar as Venus, Earth, and Mars.

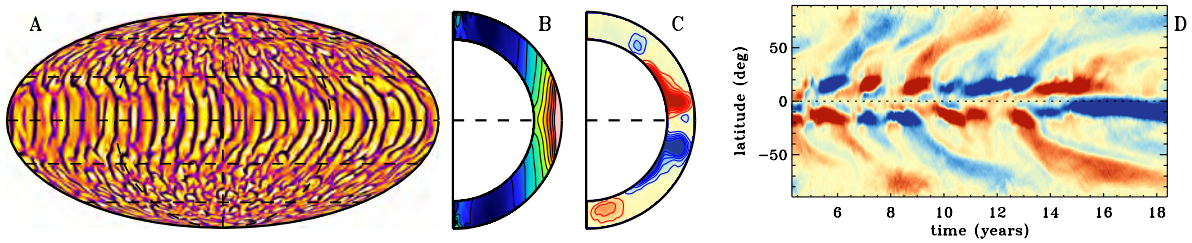


Fig. 3 Dynamo simulation of a Sun-like star rotating 5x faster than the Sun [3]. Rotationally-influenced convection (A: radial velocity near the surface: upflows bright, downflows dark) establishes a strong rotational shear (B: mean angular velocity; pink faster, blue slower) which generates coherent bands of toroidal field (C: red eastward, blue westward). These toroidal bands undergo quasi-periodic magnetic cycles with polarity reversals on a 4-year time scale (D: mean toroidal field at the base of the convection zone, of order 10kG). Such simulations have yet to achieve full realism, but already provide important insights.

3 Flows in the interior of stars

Our detailed knowledge of the stellar internal stratification stands in stark contrast to the simplified models of the large-scale flows. Solar data suggest that the magnetic field affects the large-scale flows, at least in the phenomenon of the torsional oscillation. But the observed variations in the surface meridional flow remain to be understood: are they associated with the magnetic field or are they a consequence of fluctuations in convection?

The situation is a little better for the differential rotation, because several techniques already help constrain it for some stars by using the fact that the periodicity of the disk-integrated signal depends on the latitude at which active regions preferentially emerge. Thus, as the sunspot cycle progresses, the rotation period changes with the mean active latitude. Stellar astronomers have assumed that the spread in observed rotation periods for a given star over the years is a measure of at least the magnitude of the surface differential rotation of a star, although with the caveat that there is no clear matching of the measured rotation period with the rotation period of the corresponding mean emergence latitude for the Sun. We infer that for Sun-like stars the differential rotation is comparable to that of the Sun, and are learning about the variation in differential rotation with spectral type along the main sequence[13]. Observations with, e.g., COROT and Kepler can be expected to provide much more information on stellar surface differential rotation. For fairly rapidly rotating stars for which a series of Doppler images is available, an even better constraint on differential rotation may be achieved (Sect. 4).

Numerical experiments show interesting trends (Fig. 3), but whether these are consistent with stellar behavior remains to be established. Recent models of much of the solar convective envelope at different rotation rates, for example, suggest a decrease in relative differential rotation with increasing rotation rate that is not supported by the limited available observations, associated with a marked decrease in meridional flow that is inconsistent with the strong increase suggested by flux transport modeling to explain polar spots[8]. Here, too, *observational guidance is essential to validate our approximating models of the flows of matter, energy, and field within the interiors of Sun and stars.* Estimates of internal rotation rates of stars should become available with the growth of stellar seismology, which can be compared with their surface rotation; in a few cases even the depth of the convective envelope can be constrained[0]. *A real breakthrough is anitcipated once stellar surfaces are resolved, and asteroseismic determinations of latitude and depth dependences of the flows become available.*

4 Observations: possibilities and limitations

Key to successfully developing a predictive dynamo theory is *a population study to compare the dynamo-driven activity in the Sun to that in a diverse sample of stars*.

For the Sun, we need (1) continued observations of surface activity and interior flows and (2) utilization of long-term records that exist in, e.g. ice formations and other deposits on Earth. In particular, frontiers of relevance to the dynamo appear to include (a) investigations into the structure and evolution of the solar differential rotation and meridional circulation, particularly at high latitudes, including how such flows couple to magnetism; (b) investigation of the distribution and evolution of magnetic flux on all scales in the photosphere and investigations of the characteristics of emerging flux from photospheric, chromospheric, coronal, and heliospheric observations, including the energy and flux budgets, the fibril nature of the field, the photospheric evolution of emerged flux (fragmentation, coalescence, emergence, submergence), pole-equator differences in fields and flows, helicity flux, multipolar components of the global poloidal field, dipolar tilt, etc.; and (c) study of comparative solar minima and maxima.

A breakthrough in our understanding of dynamos can also come from spatially-resolved imaging of the dynamo-driven emission patterns on a variety of stars. How these patterns depend on stellar properties (such as convection, rotation, meridional circulation, age, …), are crucial for dynamo theorists to explore the sensitive dependences on many poorly known parameters, to investigate bifurcations in a non-linear 3D dynamo, and to ultimately validate a model.

For other stars, magnetic fields can be mapped by

• *Rotational modulation.* Brightness variations caused by structures rotating across the visible hemisphere can yield longitude maps of activity if the intrinsic evolution of the source regions is slow compared to the stellar rotation, which is not the case for slow rotators like the Sun.

• *Eclipse mapping.* The crossing of one star in front of its companion allows some low-resolution recovery of atmospheric structure. But there are very few Sun-like eclipsing systems with orbits wide enough to ignore tidal coupling.

• *Doppler imaging.* Inhomogeneities on the surface of a rotating star cause line-profile distortions that can be used to map these inhomogeneities. Some latitude information can be recovered from the velocity amplitudes. The technique is limited to stars that rotate at least 5 times faster than the Sun, but not so fast that Doppler and thermal broadening are comparable. Thus, the method is not useful for stars of Sun-like activity. The use of Zeeman-sensitive lines enables access to slower rotators, but only for the largest-scale magnetic field, leaving active region scale structures well out of reach.

• *Interferometric imaging* is currently most frequently used at radio wavelengths. Only a few

optical interferometric imaging studies have been performed on stars. New ground-based instrumentation will boost the potential. However, UV/optical interferometry from space (as with the Stellar Imager) is essential for a breakthrough in dynamo modeling.

• *Direct imaging* is possible for very few stars. The star α Ori is one of the largest in the night sky, and UV/optical imaging that star with 30 elements across the photospheric diameter would require a telescope 10 times larger than HST. Nearby dwarf stars would require diameters of several hundred meters (see Sect. 5 and Fig. 65

For stars with rotation periods of less than about a week, useful information can be gathered by (combinations of) these methods. *But only large-baseline interferometers can provide latitude and longitude maps of the magnetic fields of stars that rotate as slowly as the Sun* (Fig. 4).

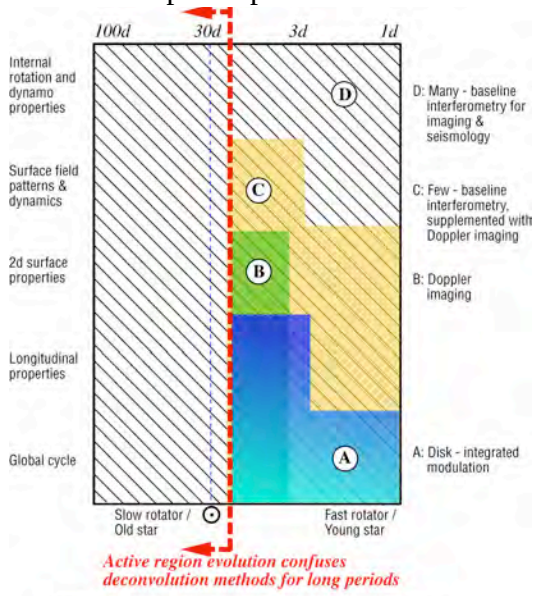


Fig. 4 Many questions regarding geometrical information on surface flows and surface magnetic fields on cool stars like the Sun require some form of imaging. Spectroscopic Doppler imaging can help in this area only by coarse mapping of stars that rotate neither too slowly nor too rapidly. Rotational modulation provides no latitudinal differentiation within the visible hemisphere. Space-based UV/optical interferometric imaging opens up the entire range of rotation periods to determination of patterns in flux emergence and migration.

The principal unknowns about stellar interiors that need to be established in order to understand stellar magnetic activity are the large-scale transport mechanisms for the magnetic field, i.e., the internal differential rotation and meridional flow. The Sun's internal differential rotation is measured using helioseismic means: the higher-order acoustic modes that resonate within the Sun are slightly altered in frequency because of the Doppler effects associated with the Sun's rotation. Modes with degrees ℓ = 0-3, measured in integrated starlight, provide the mean rotation period. Measurement of differential rotation within the convective envelopes necessarily requires spatial resolution to measure modes well above $\ell \sim 10$. Using the differential paths probed by modes of different order and degree, the latitude and depth dependence of the internal differential rotation can be determined. For stars, the acoustic signature of meridional circulation will remain out of reach, but the internal differential rotation can be determined by observing the *p*-modes in the intensity signal at stellar surfaces. This signal can be determined from measurements taken well within the period of the modes; for main-sequence stars, a 1-min spacing between measurements is required to measure *p* modes with 3-8 minute periods; a lower cadence suffices for evolved stars.

5 Feasibility of interferometric imaging

An interferometric imager should provide sub-milliarcsecond images for a significant sample of stars similar to the Sun, as well for other cool stars with very different characteristics. The Stellar Imager (SI[20]) is one such proposed UV/optical observatory. SI would have access to an exciting array of distinct stars and stellar systems (Fig. 5). For a resolution comparable to a medium-sized solar active region, a maximum baseline of ~500 m is needed when observing a dwarf star with an angular diameter of 2 milliarcseconds in the mid-UV. In addition to single F- through M-type dwarf stars, a variety of other target stars is accessible to such an observatory, including wide binaries, stars with planetary systems, and stars that probe the Sun's activity over time. Such an observatory will, for the first time, enable imaging of magnetic activity of a variety of Sun-like stars, and of stars with shallow convective envelopes, fully-convective cool, close

binary systems with dual active components magnetically coupled at a few stellar radii, compact RS-CvN-type binaries, mass-transferring Algol-type systems, and red (super-)giant stars.

6 Needs and opportunities for dynamo-related studies in the coming decade(s)

Development and validation of a predictive solar dynamo model requires advances in solar and stellar modeling and observations. Improved MHD modeling of the Sun combined with helioseismic measurements of subsurface flows and field-induced perturbations by will improve our understanding of the stellar internal dynamics. Disk-integrated observations of stars will provide low-resolution views of their internal structure and mean internal and surface rotation, and will constrain differential rotation and modulated starspot activity for many types of stars. Imaging of stars and helioseismic mapping of internal flows opens up entirely new sets of constraints to the development and testing of dynamo models for stars, including the Sun in the past, present, and future. And long-term solar activity variations can be studied in geological records, provided all transport processes involved in their deposition from Sun to Earth are understood. An interdisciplinary research approach is outlined in the Summary.

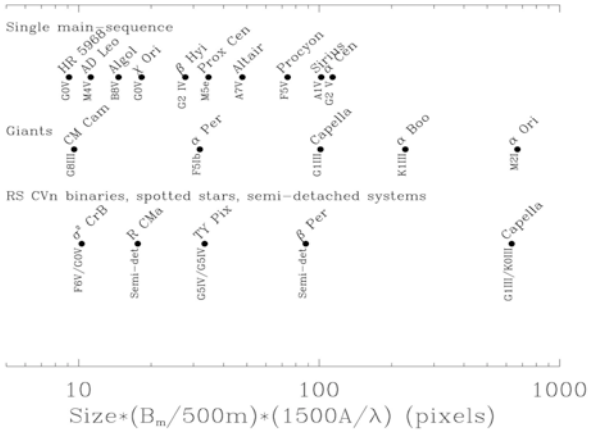


Fig. 5 Sample targets for the study of stellar dynamos. The horizontal axis shows the number of resolution elements across the equator for single stars or across the system at quadrature for binaries, assuming a baseline of 500m observing at 1,500Å.

[0] Balot, J., Turck-Chieze, S., Garcia, R., 2004, A&A 423, 1051.
[1] Baumann, I., Schmitt, D., Schuessler, M., 2006, A&A 446, 307.
[2] Brandenburg, A., 2009, eprint arXiv:0901.3789; Tobias et al., 2001, ApJ 549, 1183.
[3] Brown, B., Miesch, M. Browning, M., Brun, A., Toomre, J., 2010, ApJ, submitted.
[4] Browning, M., 2008, ApJ 676, 1262.
[5] Donati, J.-F., et al., 2008, MNRAS 390, 545; Busso et al., 2007, ApJ 671, 802.
[6] Getman, K., et al., 2008, ApJ 688, 437; Bouvier, J., et al., 2007, A&A 463, 1017.
[7] Lean, J., Wang, Y., Sheeley, N., 2002, GeoRL 29, 77.
[8] MacKay, D., Jardine, M., Cameron, A., Donati, J.-F., Hussain, G., 2004, MNRAS 354, 737.
[9] Miesch, M., Brun, A., DeRosa, M., Toomre, J., 2008, ApJ 673, 557.
[10] Moore, T., Horwitz, J., 2007, Rev. Geophys. 45, CiteID RG3002; and
Guinan, E., Engle, S., 2009, ArXiv/Astro-ph 0901.1860v1
[11] Pesnell, D., 2008, Solar Physics 252, 209.
[12] Pevtsov, A., Longcope, D., 2007, ASP Conf. Ser. 369, p. 99.
[13] Reiners, D., 2006, A&A 446, 267.
[14] Saar, S., Brandenburg, A., 1999, ApJ 524, 295.
[15] Schrijver, C., DeRosa, M., Title, A., 2002, ApJ 577, 1006.
[16] Schrijver, C., Liu, Y., 2008, Solar Physics 252, 19.
[17] Schrijver, C., Zwaan, C., 2000, 'Solar and Stellar Magnetic Activity', Cambr. Univ. Press.
[18] Stein, R., Nordlund, A, 2009, eprint arXiv:0811.0472.
[19] Strassmeier, K., 1999, A&A 347, 225.
[20] For the Stellar Imager, see the mission whitepaper, or: http://hires.gsfc.nasa.gov/si/
[21] Luciola: http://www.oamp.fr/lise/LuciolaProposalCorr-1.pdf

Solar Ion Acceleration and the Flaring Atmosphere

Heliophysics Decadal Survey White Paper

A. Y. Shih (NASA/GSFC; albert.y.shih@nasa.gov),
G. A. de Nolfo (NASA/GSFC), R. P. Lin (SSL/UCB), R. J. Murphy (NRL),
J. M. Ryan (UNH), D. M. Smith (UCSC)

1. INTRODUCTION

Solar flares are the most powerful explosions in the solar system, releasing up to 10^{32}–10^{33} ergs in 100–1,000 seconds in the release of stored magnetic energy, presumably through magnetic reconnection, and producing ions up to tens of GeV and electrons up to hundreds of MeV. Following such flares, ions and electrons up to comparable energies are often detected in situ near 1 AU in intense, gradual solar energetic particle (SEP) events (see Cane et al. 1986), but the delayed onsets and the mostly coronal/solar wind composition and charge states of the SEPs indicate that they are accelerated by shocks driven by associated fast CMEs at altitudes of ~2–40 solar radii (Kahler 1994; Tylka & Lee 2006), and not by flares. In large flares, accelerated >~20 keV electrons and >~few MeV ions often contain >~10%–50% of the total energy released (Lin & Hudson 1976; Lin et al. 2003; Emslie et al. 2004b, 2005). Thus, studying solar ion acceleration by observing its signatures is crucial to our understanding of the transient, efficient release of stored magnetic energy.

Not only is solar energy release and particle acceleration relevant to life on Earth due to space-weather considerations, but the Sun is an excellent laboratory to study similar processes occurring throughout the universe. We discuss the open science questions associated with solar ion acceleration, as well as the questions related to the flaring atmosphere that are linked to interactions with energetic ions (Section 2). We then outline the new measurements of remotely observable signatures that are required to address these science questions (Section 3). These signatures range from the typically studied gamma rays to the only recently observed energetic neutral atoms, and can be compared against in situ measurements of SEPs. Finally, we describe several instrument concepts to make these measurements (Section 4). The required technologies for most of these concepts either have already been proven or will be proven over the next few years.

2. SCIENCE TARGETS

What is the nature of solar ion acceleration?

How are ions accelerated?

While SEPs are thought to be accelerated by a first-order Fermi process at the CME-associated shock, the acceleration of ions by flares is a more open question. There are many candidates for flare acceleration, from acceleration in a DC electric field to stochastic acceleration by cascading turbulent waves. Even at the CME-associated shock, there are still questions about the shock geometries and how the SEPs diffuse across the shock for efficient acceleration.

These acceleration processes leave indicators: where the energetic particles are produced, their energy spectrum, and their composition. For example, the cascading waves in stochastic acceleration can produce a distinctive signature in the relative ion abundances depending on the resonances between the various modes. Changes in these indicators as a flare progresses (e.g., Rank et al. 2001) point to the evolution of energy release and particle acceleration over an event. Therefore, measuring these characteristics of the energetic particles is necessary for understanding the acceleration process.

What is the relationship between ion acceleration and electron acceleration?

While most if not all flares are thought to accelerate electrons, it remains an open question

whether all flares accelerate ions. The smallest flare (in terms of soft X-ray emission) observed to date with ion-associated emission was a C4.1 flare observed by ***CGRO***/COMPTEL (Young et al. 2001). Correlations such as the one shown in Figure 1 suggest that the flare acceleration of ions and electrons is roughly proportional, independent of flare size, and the lack of ion-associated signatures in smaller flares is simply a result of limited sensitivity (see Shih et al. 2009). However, intervals within flares can occasionally favor one population significantly (see Rieger et al. 1998).

Furthermore, the locations of ion-associated emission and electron-associated emission in some flares have shown surprising spatial separations, indicative of unexplained differences in the acceleration and/or transport processes (Hurford et al. 2003, 2006). One proposed solution attributed spatial separations to the fact that stochastic acceleration is more efficient for ions on longer loops (Emslie et al. 2004a), but the ***RHESSI*** observation for the 2003 Oct 28 flare shows that the footpoint separation is the same despite the spatial shift (Figure 2). Given the limitations of current observations, there are not any comparable comparisons for more and/or smaller flares to address this mystery.

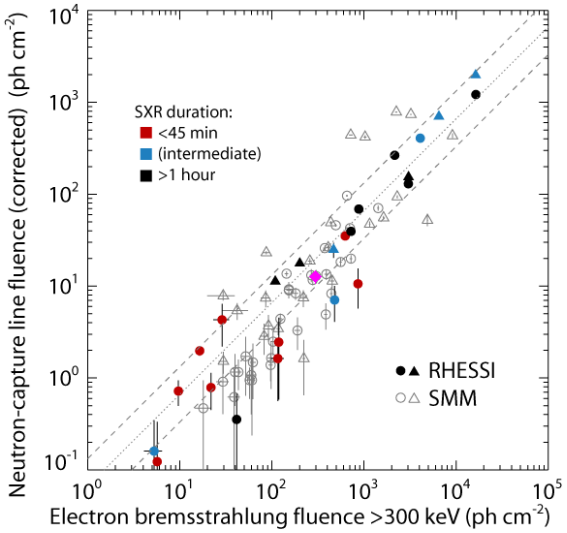


Figure 1 — Ion-associated emission is well correlated with electron-associated emission when flare-integrated, indicating that the two populations are roughly proportional over many orders of magnitude (Shih et al. 2009). The magenta point is the recent 2010 Jun 12 flare.

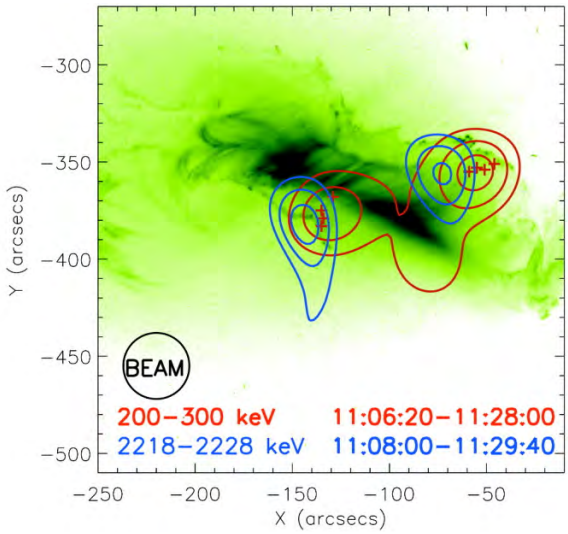


Figure 2 — Ion-associated sources (blue) are spatially shifted from electron-associated sources (red) in the 2003 Oct 28 flare, indicating an unexplained difference in acceleration and/or transport processes (Hurford et al. 2006). The coutours are from ***RHESSI***, and the background image is from ***TRACE*** (195Å).

What is the total energy content in accelerated ions?

Of the large fraction of released energy that goes into particles, there appears to be a rough equipartition in energy between >20 keV electrons and >1 MeV ions that interact at the Sun (Ramaty & Mandzhavidze 2000), although the flare-to-flare variation could still be quite large. Determinations of the total ion energy content, however, could be significant underestimates because the signatures associated with sub-MeV ions have yet to be observed in a solar flare, and thus the low-energy shape or cutoff of the ion spectrum is completely unknown. Obtaining more accurate estimates of the total ion energy content can put strong constraints on the overall flare budget and hence the possible acceleration processes.

The total ion energy content also has implications for so-called sunquakes, which are seismic disturbances associated with flares. Although the sunquakes for smaller flares may be caused by electron beams, the larger sunquakes may require the proton beams as well (Kosovichev 2006; Zharkova 2008).

What is the relationship between interacting particles and escaping SEPs?

SEPs are most important from a space-weather perspective, and their connection to particles that interact at the Sun is still unclear. While SEPs are generally thought to be accelerated by CME-associated shocks, the interacting particles accelerated by the flare may form a seed population out of which SEPs are accelerated. Furthermore, especially in impulsive SEP events without CMEs, the SEPs may in fact be the same population as the interacting particles. Thus, understanding the relationship between particles interacting in flares and escaping SEPs is an important part of any effort to forecast the SEP population before they are observed in situ.

This relationship can be studied by comparing the spectra, compositions, and timing of the two populations. Note that there may also be a component of the interacting particles that is directly associated with the SEPs: SEPs accelerated at the CME-associated shock may precipitate down over a large fraction of the solar disc to produce observable signatures.

What is nature of the flaring atmosphere?

Elemental abundances

Energetic ions are a useful tool to probe the flaring atmosphere, as the signatures of energetic ions depend not only on the ions themselves, but also on the characteristics of the atmosphere in which they interact. Foremost, observations of the strong narrow nuclear de-excitation lines (see Section 3) are a measure of the relative ambient abundances of six major elements in the ambient medium: iron, magnesium, silicon, neon, carbon, and oxygen. Note that energetic ions can thus probe the abundance of neon in the chromosphere, which is difficult to observe through other means. Energetic ions can also probe the helium-3 abundance in the solar atmosphere, which is of cosmological significance.

Some flares also show intriguing variations of the relative ambient abundances over the course of the flare. The ***CGRO***/OSSE observation of the 1991 Jun 4 flare showed evolving emission consistent with a relative enhancement of the elements with low first-ionization potentials (FIPs) (Murphy et al. 1997). Note that the corona typically shows an enhanced abundance of low-FIP elements compared to the chromosphere. Thus, FIP-related temporal variability could indicate that the interaction region in the chromosphere is being progressively filled with coronal material, or that the location of ion interaction is increasing in altitude, possibly due to changing magnetic mirroring. The ***RHESSI*** observation of the 2003 Nov 2 flare showed similar but quicker behavior, on the order of minutes, but the 2002 Jul 23 flare showed significant changes that could not be attributed to the FIP effect (Shih 2009).

Atmospheric conditions

For a number of flares, ***RHESSI*** observations of the positron-annihilation line have shown that the ion-associated emission is being produced in an ambient ionized medium at $>10^5$ K with densities $>10^{13}$ H/cm^3 (Share et al. 2003, 2004). This significant amount of material at transition-region temperatures is difficult to reconcile with models of the solar atmosphere. In addition, for the 2003 Oct 28 flare in particular, the ambient conditions appear to change significantly and quickly (on the order of minutes) to indicate a new ambient medium that is still ionized but at a much cooler temperature of $\sim10^4$ K. Such a rapid reduction in temperature or drastic change in the location of interaction is still unexplained.

Chromospheric mixing

The interacting particles will also produce radioactive elements in the chromosphere, some with long half-lives (Tatischeff et al. 2006). The most extreme anticipated example is cobalt-56, which has a half-life of 77.2 days. Since active regions can produce multiple significant gamma-ray flares, the long-lived radioactive material will accumulate over time and even survive over multiple solar rotations. However, if the material is convected to lower altitudes, it will no longer be observable. Thus, such long-lived radioactivity could function as a unique tracer of chromospheric mixing of the post-flare material.

3. MEASUREMENT REQUIREMENTS

There is a variety of ways for remotely observing both interacting particles and SEPs using neutral radiation, and these measurements can be compared against in situ observations of SEPs. Neutral radiation can propagate across interplanetary space without interacting with the magnetic field or other particles, but modeling may be required to recover the original characteristics of the energetic particles that produce these observable signatures.

Gamma rays

Gamma rays are produced through a large number of processes, generally beginning with nuclear interactions (see Murphy et al. 2007). Imaging of these lines with high resolution (~5 arcsec) can resolve ion-associated footpoint structure in many flares, for comparison against electron-associated footpoints, while coarser imaging (~1 arcmin) can at least distinguish between ions on closed field lines and those that precipitate from higher altitudes (e.g., the CME shock).

Ions of energies between ~1 MeV and ~50 MeV produce excited nuclei that produce gamma-ray lines when they relax. Energetic protons and alpha particles colliding with ambient carbon and heavier nuclei produce narrow prompt de-excitation lines (widths of a few keV to ~100 keV), while energetic heavy nuclei colliding with ambient hydrogen and helium produce much broader lines due to Doppler broadening (widths of a few hundred keV to an MeV). Spectral resolution of less than a few percent FWHM can unravel the information from these lines, including the Doppler shifts of the lines from ion directivity (Smith et al. 2003), and higher spectral resolutions (<~1% FWHM) can resolve the distortions to the line shape due to abundance effects. Very high spectral resolution (~0.25% FWHM) can increase the sensitivity to very narrow lines from long-lived radioactive nuclei.

Ions of energies above ~20 MeV produce secondary neutrons, which can thermalize in high densities and be captured to form an intrinsically narrow line at 2.223 MeV. This neutron-capture line is the only line imaged to date to show multiple spatially resolved gamma-ray line sources. A high-resolution spectrometer (<10 keV) maximizes the sensitivity for this strong line, and accurate measurements of the continuum produced by Compton scattering at the Sun of this line is another handle on the ion directivity and spectral shape (Vestrand 1990).

Sub-MeV ions can produce gamma-ray lines through radiative capture (MacKinnon 1989), but these lines have not been observed. Limits on observations by ***SMM*** (Share et al. 2001) suggests that a sensitivity of at least $<10^{-5}$ photons/cm^2/s will be needed to observe these lines.

Ions with energies greater than hundreds of MeV will produce pions, both charged and neutral. Neutral pions decay to produce a broad gamma-ray continuum peaked at ~70 MeV, while charged pions produce secondary electrons and positrons, which can consequently produce energetic bremsstrahlung emission superposed on the radiation from directly accelerated electrons. Thus, spectroscopy and imaging of the pion-decay component of the gamma-ray spectrum studies the extreme limits of ion acceleration and whether such ions are produced in the same accelerator as the rest of the energetic ions.

Positrons are also produced as a result of the nuclear interactions, and annihilate with the ambient medium to produce a characteristic line at 511 keV and a continuum at lower energies. The shape of the line depends on the temperature, density, and ionization fraction of the ambient medium (Murphy et al. 2005), and thus high-resolution spectroscopy (<2 keV) and imaging (~10 arcsec) can study the fast-changing atmospheric conditions and compare them against observations of other signatures (e.g., EUV).

Neutrons

In addition to producing the neutron-capture line, the neutrons produced in nuclear interactions can escape the Sun. Neutrons provide a separate handle on accelerated ions from MeV energies to hundreds of MeV, and can be critical for filling in the gap between the energies typically responsible for gamma-ray lines and those responsible for pion-decay emission. At the lowest energies, some of the neutron-producing alpha interactions are actually exothermic, and thus neutrons may be an effective way to probe low-energy alpha particles.

Neutrons above ~10 MeV can be measured in low-Earth orbit but neutrons below ~10 MeV will

decay significantly before reaching 1 AU. Thus, unlike the fluxes of other neutral radiation that scale inversely with the square of the distance from the Sun, the flux of low-energy neutrons will increase even faster as one observes from the inner heliosphere. Ion energies of ~50 to ~300 MeV, which may contain critical information about spectral rollovers representative of the acceleration process, are effectively probed only via 15–150 MeV neutrons. Compact, efficient instruments such as ***SONTRAC*** (Ryan et al. 1999) can perform low-background measurements of these particles with increased effective area. Neutrons above ~500 MeV are energetic enough to produce a signal in ground-level monitors, and are complementary to gamma-ray observations of the same population via pion-decay emission.

Energetic neutral atoms (ENAs)

The serendipitous discovery using ***STEREO*** of energetic neutral atoms (ENAs) at MeV energies associated with the 2006 December 4 eruptive event (Mewaldt et al. 2009, 2010) introduced a new probe of solar particle acceleration. These ENAs were likely produced by SEPs in the corona (~2 solar radii), but they may also have been produced by interacting particles in the chromosphere through a diffusive process of neutralization and reionization. As with other neutral radiation, ENAs ignore magnetic fields as they traverse the interplanetary medium, but they have the added benefit of essentially preserving the energy of the original energetic particles. Thus, ENA observations can provide a clean measurement of the acceleration spectrum, free from propagation effects, in addition to localizing where the SEPs are energized. For example, for ENAs produced by a CME-associated shock, imaging (<0.1°) can distinguish what shock conditions are responsible for efficient SEP acceleration (e.g., quasi-parallel versus quasi-perpendicular).

The ***STEREO*** ENA observations in the MeV range measure those SEPs that charge exchange with heavy ions (e.g., oxygen). Extending the ENA observations down to ~10 keV will include those low-energy SEPs that charge exchange with the small but non-negligible fraction of neutral hydrogen and hydrogen-like helium in the ambient medium. Thus, these ENA measurements can probe ion energies that are poorly or completely inaccessible through other signatures.

Extreme ultraviolet

Extreme ultraviolet (EUV) radiation provides a potential for observing sub-MeV interacting ions separate from gamma-ray measurements. Lyman-alpha 1216 Å line emission produced by ~300 keV protons that charge exchange in flight (Orrall & Zirker 1976) will be Doppler-shifted by the motion of the protons. Thus, for the presumably downward-directed flare ions, this emission will appear in the redshift wing of the primary line. Although not yet observed in the Sun, an increase in the red wing of Lyman-alpha was observed during the impulsive phase of a stellar flare (Woodgate et al. 1992). An analogous process can occur for alpha particles and the He II 304 Å line (Canfield & Chang 1985).

Therefore, high-resolution (<1 Å) EUV spectroscopy of these lines and their wings may measure the elusive sub-MeV ions. Next-generation instrumentation to study the thermal plasma during the impulsive phase of flares may achieve these requirements, so observations of this ion-associated signature would be inherently included. Also, note that high-cadence EUV imaging spectroscopy, free from saturation, can provide an independent measurement of the temperature of the flaring atmosphere and the rapid variations implied by the annihilation-line observations.

Terahertz radiation

Although much of the electromagnetic spectrum has been observed in flares, terahertz radiation is still essentially unexplored. Observations at hundreds of GHz by the Solar Submillimeter Telescope (SST) (e.g., Kaufmann et al. 2009) suggest the presence of strong THz sources, possibly localized to nonthermal sources. Possible sources of THz emission include:

- Thermal continuum emission with large emission measures
- Synchrotron emission of relativistic electrons separate from those producing microwave emission
- Synchrotron emission of secondary positrons produced in high-energy nuclear interactions

Clearly, THz measurements may provide significant diagnostic capabilities that complement other observable signatures.

4. INSTRUMENT CONCEPTS

The first three instrument concepts address the science targets that are associated with gamma-ray measurements, and could be standalone missions (~Explorer class) or parts of suites of instruments (e.g., Lin et al. 2010). These concepts benefit from ongoing technology development and the accrued experience from past gamma-ray instruments. However, we stress that gamma rays should not be the only avenue for future observations, and new instruments should also be developed for observing the other signatures detailed earlier. Importantly, continuing in situ measurements of SEPs, by spacecraft at the L1 Lagrange point or in the inner heliosphere, are vital for comparisons against remote observations of energetic particles. Also, inner-heliosphere spacecraft should include low-energy neutron detectors given the large science value. Finally, context from all other types of flare and CME observations are invaluable for studying solar ion acceleration.

Large-effective-area gamma-ray imager

RHESSI has demonstrated Fourier-transform gamma-ray imaging using bi-grid rotating modulation collimators (RMCs), but that instrument was not optimized specifically for gamma-ray observations. A design optimized for gamma rays can use thicker grids and a longer collimator to provide high-resolution gamma-ray imaging (~10 arcsec), and incorporate a spectrometer with high photopeak efficiency to be able to image even the pion-decay component. With this high sensitivity, it may also be able to measure never-before-seen gamma-ray signatures. Note that many of the science targets do not explicitly require gamma-ray measurements with high spectral resolution, as long as the combination of a large effective area and modeling of gamma-ray production processes is able to deduce the desired physical parameters.

This instrument concept would combine RMCs optimized for gamma-ray imaging with a large amount of detector volume, preferably $LaBr_3$ or some other scintillator. Thus, this design emphasizes a large effective area and high sensitivity, with a low level of complexity. However, as an RMC-based design requires multiple collimators to image at a range of spatial scales, a compact arrangement would likely make a full-Sun field of view prohibitive. Given a limited field of view (~3 arcmin), this imager must be pointed at a chosen active region of interest, and thus flares may be missed during particularly active times. Furthermore, the efficiency of imaging gamma rays in broad energy bands (or gamma-ray lines without high spectral resolution) will be deleteriously affected by scattering in the grid directly in the front of the detector.

High-resolution gamma-ray imager/polarimeter

Two newly developed technologies can be used in a next-generation gamma-ray instrument. First, 3D position-sensitive germanium detectors (3D-GeDs) combine the high spectral resolution of germanium with the capability of spatially resolving each energy deposition to within <0.1 mm^3. Using this information, it is possible to reconstruct the path of a gamma ray as it Compton scatters around the volume. This capability not only allows for rejecting background using Compton-imaging techniques, but also allows for determining precisely where the gamma ray ***first*** interacted with the detector.

This high spatial resolution leads into the second technology: the multi-pitch rotating modulator (MPRM). The modulating slats are arranged in a parallel with a quasi-continuous range of pitches, resulting in quasi-continuous coverage in the spatial-frequency domain (in contrast, ***RHESSI*** only sampled two spatial frequencies at gamma-ray energies). Each observed gamma ray is back-projected from its location of first interaction to the stripes on the Sun not blocked by the MPRM. There is no grid directly in front of the detector, which doubles throughput over a bi-grid design and also reduces the possibility of gamma rays scattering before reaching the detector. This technology is similar to imaging via coded apertures, but the MPRM is simpler to fabricate and has performance advantages given the resolution and field-of-view requirements.

These technologies are already being built and will be tested on the ***Gamma-Ray Imager/Polarimeter for Solar flares*** (***GRIPS***), a balloon-borne instrument with its test flight slated for spring of 2012. The MPRM will be mounted 8 meters away from the 3D-GeDs, which will provide an

imaging resolution of 12.5 arcsec. Long-duration Antarctic balloon flights in following years are anticipated to fully test the instrument and return science. This timing is well-suited for proving technological readiness for a mission designed to observe the maximum of solar cycle 25.

This instrument concept emphasizes high spectral resolution and high imaging resolution by using the technologies that will be proven by ***GRIPS***. A large number of 3D-GeDs, surrounded by a scintillator-based anti-coincidence shield, would make up the spectrometer. An MPRM would be mounted 20 meters away from the spectrometer on a boom, which will provide an imaging resolution of 5 arcsec. Unlike the previous instrument concept, this instrument would be able to measure the width of the positron-annihilation line, as well as produce images of the line to investigate changes in location. In addition, the MPRM design easily allows for a full-Sun field of view. However, depending on the capabilities of the anti-coincidence shield, this instrument may perform less well with respect to measuring and imaging higher-energy lines and pion-decay emission. This emphasizes the difficulty of comprehensive gamma-ray imaging and spectroscopy: i.e., it is difficult for one instrument to do everything.

Note that a spectrometer composed of 3D-GeDs is inherently also a very capable polarimeter at ~0.1–1 MeV. Since Compton-scatter tracks are reconstructed with high precision, azimuthal asymmetry in the scattering measures the amount and direction of polarization of the source. The science targets associated with the measuring X-ray and gamma-ray polarimetry are beyond the scope of this particular white paper (see McConnell et al. 2010), but would be inherently addressed with this instrument concept.

Focusing gamma-ray imager

The two aforementioned instrument concepts both use indirect techniques to image at gamma-ray energies. In contrast, there are two types of focusing lenses that can directly image gamma rays: Fresnel lenses and Laue lenses. Fresnel lenses take advantage of the small but non-negligible change in wavelength in a medium to focus gamma rays via refraction, with focal lengths on the order of millions of kilometers. Laue lenses, on the other hand, use Bragg diffraction in crystals with appropriately chosen spacing to focus gamma rays with focal lengths on the order of hundreds of meters. For both of these lens types, a given lens would only be optimized for a range of gamma-ray energies, and its focal length would vary with the gamma-ray energy, although more complex designs may be able to mitigate this effect (cf. addressing chromatic aberration in optical lenses). Laue lenses have the added disadvantage of a peculiar point-spread function.

This instrument concept would combine a gamma-ray focusing lens with a detector, which could be optimized for efficiency or for spectral resolution. Formation-flying spacecraft are likely required to achieve the needed separation between the lens and the detector. Using such lenses could dramatically improve imaging resolution, possibly even to the milliarcsec level. However, given the mean free paths of interacting particles, it is unclear whether gamma-ray signatures would have any sub-arcsec structure. Furthermore, while gamma-ray focusing increases the effective area of the instrument, the complexity associated with a lens of the size necessary to have a sufficient field of view (>3 arcmin) may far outweigh any benefit from not simply using a larger detector. Despite these obstacles, the required technologies are already being studied for instruments to observe astrophysical sources, and new developments and advancements may lead to useful adaptation to solar gamma-ray imaging.

ENA imaging spectrometer

The STE detectors on ***STEREO*** have demonstrated that low-threshold particle detectors will detect ENAs across the energy range of interest. The ***CINEMA*** Cubesats currently being built will soon demonstrate that a deflection electric field can be used to discriminate between ENAs and charged particles.

This instrument concept would build upon such developed instruments by using slit- or pinhole-imaging techniques in combination with a large area of ENA detectors to measure both their locations and spectra. Such an instrument could also be used to observe ENAs and charged particles from magnetospheric sources as well.

5. REFERENCES

Cane, H. V., McGuire, R. E., & von Rosenvinge, T. T. 1986, ApJ, 301, 448
Canfield, R. C., & Chang, C.-R. 1985, ApJ, 295, 275
Emslie, A. G., Miller, J. A., & Brown, J. C. 2004a, ApJ, 602, L69
Emslie, A. G., et al. 2004b, J. Geophys. Res., 109, 10104
Emslie, A. G., Dennis, B. R., Holman, G. D., & Hudson, H. S. 2005, J. Geophys. Res., 110, 11103
Hurford, G. J., Schwartz, R. A., Krucker, S., Lin, R. P., Smith, D. M., & Vilmer, N. 2003, ApJ, 595, L77
Hurford, G. J., Krucker, S., Lin, R. P., Schwartz, R. A., Share, G. H., & Smith, D. M. 2006, ApJ, 644, L93
Kahler, S. 1994, ApJ, 428, 837
Kaufmann, P., Trottet, G., Giménez de Castro, C. G., Raulin, J.-P., Krucker, S., Shih, A. Y., & Levato, H. 2009, Sol. Phys., 255, 131
Kosovichev, A. G. 2006, Sol. Phys., 238, 1
Lin, R. P., & Hudson, H. S. 1976, Sol. Phys., 50, 153
Lin, R. P., et al. 2003, ApJ, 595, L69
Lin, R. P., et al. 2010, "Solar Eruptive Events (SEE) Mission Concept", Decadal Survey white paper
MacKinnon, A. L. 1989, A&A, 226, 284
McConnell, M. L., et al. 2010, "Hard X-ray and Gamma-Ray Polarization Measurements of Solar Flares During the Next Solar Maximum", Decadal Survey white paper
Mewaldt, R. A., et al. 2009, ApJ, 693, L11
Mewaldt, R. A., et al. 2010, in AIP Conf. Proc. 1216, Twelfth International Solar Wind Conference, ed. M. Maksimovic, et al., 592
Murphy, R. J., Share, G. H., Grove, J. E., Johnson, W. N., Kinzer, R. L., Kurfess, J. D., Strickman, M. S., & Jung, G. V. 1997, ApJ, 490, 883
Murphy, R. J., Share, G. H., Skibo, J. G., & Kozlovsky, B. 2005, ApJS, 161, 495
Murphy, R. J., Kozlovsky, B., Share, G. H., Hua, X.-M., & Lingenfelter, R. E. 2007, ApJS, 168, 167
Orrall, F. Q., & Zirker, J. B. 1976, ApJ, 208, 618
Ramaty, R., & Mandzhavidze, N. 2000, in IAU Symposium, Vol. 195, Highly Energetic Physical Processes and Mechanisms for Emission from Astrophysical Plasmas, ed. P. C. H. Martens, S. Tsuruta, & M. A. Weber, 123
Rank, G., Ryan, J., Debrunner, H., McConnell, M., & Schönfelder, V. 2001, A&A, 378, 1046
Rieger, E., Gan, W. Q., & Marschhäuser, H. 1998, Sol. Phys., 183, 123
Ryan, J. M., Wenhui, L., Macri, J. R., McConnell, M. L., Castaneda, C. M., & Romero, J. L. 1999, Proc. SPIE, 3768, 496
Share, G. H., Murphy, R. J., & Newton, E. K. 2001, Sol. Phys., 201, 191
Share, G. H., et al. 2003, ApJ, 595, L85
Share, G. H., Murphy, R. J., Smith, D. M., Schwartz, R. A., & Lin, R. P. 2004, ApJ, 615, L169
Shih, A. Y., Lin, R. P., & Smith, D. M. 2009, ApJ, 698, L152
Shih, A. Y. 2009, Ph. D. thesis, University of California, Berkeley
Smith, D. M., Share, G. H., Murphy, R. J., Schwartz, R. A., Shih, A. Y., & Lin, R. P. 2003, ApJ, 595, L81
Tatischeff, V., Kozlovsky, B., Kiener, J., & Murphy, R. J. 2006, ApJS, 165, 606
Tylka, A. J., & Lee, M. A. 2006, ApJ, 646, 1319
Vestrand, W. T. 1990, ApJ, 352, 353
Woodgate, B. E., Robinson, R. D., Carpenter, K. G., Maran, S. P., & Shore, S. N. 1992, ApJ, 397, L95
Young, C. A., et al. 2001, in AIP Conf. Proc. 587, GAMMA 2001: Gamma-Ray Astrophysics 2001, ed. S. Ritz, N. Gehrels, & C. R. Shrader (Melville, NY: AIP), 613
Zharkova, V. V. 2008, Sol. Phys., 251, 641

The Case for Continued, Multi-Point Measurements in Space Science

A White Paper Submitted to the NRC 2010 Decadal Survey in Solar and Heliospheric Physics

Charles W. Smith and Nathan A. Schwadron

(University of New Hampshire)

Ruth M. Skoug

(Los Alamos National Laboratory)

Susan T. Lepri

(University of Michigan)

Christina M. S. Cohen

(California Institute of Technology)

James Weygand

(University of California at Los Angeles)

Gary P. Zank

(University of Alabama at Huntsville)

It is now 50 years since the start of the space age. In the beginning, every new launch brought new observations and new discoveries: the plasma environment above the neutral atmosphere, the extra-terrestrial radiation source(s), the magnetosphere, the radiation belts, the magnetopause, the bow shock, and the supersonic solar wind. Still, the discoveries continue. The discovery of new physics and subsequent scientific study over decades has raised far more questions than it has answered. With every new measurement and improved understanding multiple new questions are generated. While there remain fundamental discoveries yet to be made, the bulk of the space science discipline is now reaching a state of maturity where progress is made at a more detailed level. Taking nothing away from the latest missions and the newest observations, it is increasingly clear that to perform good science requires complimentary measurements from multiple points analyzed in combination with advanced numerical simulation and evolving theory.

The purpose of this paper is to argue three simple points:

1. Measurements from multiple, often widely spaced, locations are essential to provide the context needed to study the physics of the large, coupled space environment system.
2. Measurements over long time baselines that provide the most varied physical conditions provide the insight needed to see new or underlying physics that is obscured by observations that are too narrowly limited in scope.
3. Measurements from multiple, widely spaced, locations are required to support new missions making discovery measurements.

In addition to these points, we argue that new measurements are needed in old locations to make measurements at previously unresolved spatial scales or unmeasured energy ranges that will provide new insights to solve old problems.

Multi-point measurements are essential to science. Nowhere is this point made more clearly than in the magnetosphere where knowledge of the solar wind input is critical to understanding the magnetospheric response. Additionally, the largely subsonic flow conditions mean that all regions of the magnetosphere contribute to responses and dynamics at a single point. Much of magnetospheric physics is inherently non-local and widely separated multi-point measurements are essential to unravel the physics that defines this region.

Long baseline measurements are essential to science. The recent protracted solar minimum has provided unexpected new observations in solar wind and heliospheric magnetic field physics that challenge preconceived ideas of a solar minimum rest state, minimum interplanetary magnetic field flux levels, and baseline solar wind flux. These things were never predicted and they are seen in defiance of popular theories. Moreover, it is doubtful that NASA would have provided a mission for the specific purpose of studying solar minimum primarily because the recent new results were unanticipated and there was no warning of impending new science.

The extended period of low solar activity has had consequences for energetic particle studies as well. The galactic cosmic ray (GCR) intensities have reached levels higher than ever before measured in space. Small, isolated solar energetic particle (SEP) events have been observed simultaneously by spacecraft

separated in longitude well beyond the previously assumed longitudinal spread of such events. Measured anomalous cosmic ray (ACR) gradients during the extended period of high current-sheet tilt indicate their transport in the inner heliosphere is not due to rapid drift along the current sheet. Such unexpected observations over a wide energy range and resulting multi-spacecraft measurements can only be possible by spacecraft providing long baseline measurements spanning many years. These observations provide new insights and new questions for old missions with possible roles for new missions that were unanticipated previously.

Existing assets must be preserved for the simple reason that we can not adequately predict future new results in space science. Accepted paradigms are challenged when new conditions result in new measurements made by old assets. Only long baseline measurements spanning the most diverse conditions can reveal the true underlying physics governing any physical system.

New missions require multi-point support measurements. In an era where new small missions cost $300M and major missions cost $1B, there are limits to what can be flown. Pre-existing instruments and missions provide all of the contextual information required to do new science with old missions PLUS they support the new missions in the very same manner. No physical process (whether it is reconnection, particle acceleration, or the dynamics of global structure) can be pursued with local measurements alone and here the support of new missions falls on the shoulders of pre-existing hardware.

One of the most recent, notable examples of this is STEREO. Pre-existing assets at L1 are positioned to provide both a third point of solar longitude when studying transients or energetic particles as well "ground truth" for any predictions that STEREO may provide for Earth-approaching transients.

Understanding the longitudinal distribution of energetic particles, the structure and extent of interplanetary shocks and other magnetic/solar wind structures requires measurements at sufficiently separate locations. Relating these observations to space weather and their impact on the Earth requires measurements along the Sun-Earth line (e.g., at L1).

NASA's policy of making all data publicly available has resulted in enhanced science returns for all missions. For example, over 75% of all publications using ACE data do not involve an ACE/SWT member and this is the way it should be. Most missions have similar statistics. These "outside" authors extend the application of the data beyond the initial intents of the spacecraft team, diversify the science return, and build largely on multi-spacecraft, multi-point data analyses. NASA has spent decades and many millions of dollars building the in situ fleet we now enjoy and it must be maintained now or it will be lost for generations.

The Sun-Earth Environment

The physics of the Sun-Earth environment is far more complicated than that of isolated, non-interacting dynamical processes. Remote solar observations by SOHO, WIND, RHESSI, STEREO, and SDO not to mention numerous ground-based assets, reveal a high degree of magnetic structure at all scales in the photosphere and solar corona. Solar wind sources mimic the large-scale magnetic structure of the Sun

with the polar holes being major contributors and numerous low-latitude structures providing what is thought to be slow wind sources. Numerous solar models attempt to predict the 3D structure of the solar wind, but no attempts to date provide accurate and reproducible results. Multiple platforms such as ACE, WIND and STEREO are needed to test and refine these efforts. Erupting magnetic structures (CMEs) show a high degree of internal consistency in the solar wind, but they propagate through a structured solar wind showing massive large-scale structure as well as poorly studied small-scale structure. Energetic particle acceleration may reflect at least a comparable degree of structure, but inadequate spacecraft separation during times of high activity has made these studies virtually impossible. Solar wind magnetic structure on small scales may mimic the solar source, but many more studies are needed.

Diverse conditions drive new science. We have learned things about the solar wind during the recent, protracted solar minimum that were not predicted earlier: the solar wind density and flux is lower than previously seen and the IMF is weaker. Previous predictions for a minimum IMF intensity have been swept away. Both depletions exist at all latitudes of the heliosphere. Solar wind composition has changed. Comparing long baseline measurements and multi-platform observations provide the insights needed to form new theories as we resolve solar sources. At the same time the magnetosphere has entered a state not seen before. Without solar wind drivers the transient dynamics are different. The radiation belts have almost disappeared. This provides new insights and new questions for old missions. Without the older missions, we would be unaware of these unanticipated new results.

Old missions provide context for new missions. It is well established that magnetospheric science is today well-served by reliable L1 measurements characterizing the solar wind input. New missions (such as RBSP and MMS) require contextual observations to understand their own time-dependent observations. These contextual measurements include solar wind (ACE and WIND) as well as magnetospheric missions (GOES) and ground-based observations.

Interplanetary Solar Wind Studies

Every study of interplanetary dynamics has left us with new challenges and every answer has spawned new questions. The twin platforms of STEREO were launched to better understand transient dynamics from launch through interplanetary propagation, but the launch into the most protracted solar minimum of the space age has delayed many of those goals. Simulations teach us much on these topics, but they must be tempered by observations. On smaller scales, the pairing of ACE and WIND has provided new insights into medium-scale shock curvature and associated particle acceleration dynamics while additional pairings with Geotail, Cluster and Themis provide new insights into solar wind turbulence over scales from the correlation length down to dissipation scales. Single spacecraft measurements are incapable of providing these insights.

Missions deliberately designed and created as multi-platform investigations (Cluster and MMS) will serve many purposes in addition to their original goals. Cluster is now providing new insights into small-scale solar wind turbulence and MMS will eventually do the same at yet smaller scales. However, the bulk of the multi-platform studies come not from these missions, but from the fortuitous combination of

pre-existing assets at greater spatial separation. Whether the topic is turbulence, transients, or energetic particles, the Earth-orbiting and L1 assets together with more distant planetary and interplanetary missions provide the necessary associated measurements linking solar wind studies at all spatial and temporal scales. The subject could be as small as inertial range turbulence or as large as the formation of GMIRs and the related galactic cosmic ray solar cycle: the many separate assets of space physics form the largest and most versatile of the great observatories.

Planetary Studies

No one can deny the ground-breaking discoveries of the ongoing planetary missions. Undeniably, these are discovery missions in the truest sense. However, the purpose and application of these missions will also evolve with time. Once moons are mapped and their properties determined, once the radiation environments are better understood, once the planetary magnetospheric structure is determined, planetary science will turn to the type of studies now being performed in the Earth's magnetosphere – the mature study and understanding of the complex fields and currents that constitute the near-Earth (or near-planet) environment. Again, multi-point measurements will be required to achieve this understanding and while new multi-platform studies within the planetary magnetospheres will be desirable, these multi-point studies will begin with measurements of the solar wind and the resulting input into the planetary environment. This purpose will probably be served in the near-term by 1 AU missions.

Heliospheric Physics

The last decade and the most recent years have seen an unprecedented exploration of the heliosphere. Voyagers 1 & 2 have crossed the termination shock and have begun to penetrate the heliosheath. In the process they have raised new questions regarding the acceleration of anomalous cosmic rays and the propagation of galactic cosmic rays while also revising our understanding of the entire sheath region. IBEX has imaged the interstellar boundary region at all heliographic latitudes and longitudes and found signatures of previously unexpected interactions with the local interstellar medium. New Horizons is on its way to join the outer heliospheric fleet. None of these measurements could be properly understood without near-Sun baseline measurements of the solar wind and solar activity. The changing ram pressure of the solar wind, together with changing levels of transient activity, determines the location of the termination shock and a significant fraction of the heliosheath dynamics (the still-unmeasured interstellar medium determining the other fraction).

New Measurements

There is no denial that new measurements propel space science forward with the greatest impulse. No one is arguing against the development of new missions. However, new measurements without the context of simultaneous multi-point observations offer only limited value and are subject to greater misinterpretation or overly aggressive speculation. The good science, the solid science of the 21st century requires the greatest range of observations to provide the greatest scientific return on the investment.

Summary

Of necessity, every statement in this brief 7-page document is both incomplete and inadequate. We have attempted to remind the committee of the most undeniable fact that the pursuit of modern space science requires diverse, separated, long-term measurements. We have failed to do justice to any single mission. For instance, IBEX, which we describe as imaging the interstellar boundary region, has also imaged the Earth's magnetosphere, and lunar dynamics as well as interplanetary shocks. This gives IBEX a place in the study of the local Sun-Earth environment. This type of multi-tasking, multi-purpose application of resources is typical of the space physics assets today and similar statements can be made for virtually all missions.

These things are central to understanding the heliophysics system (Sun to Earth to Interstellar Space):

- Long baseline measurements and new conditions provide new relevance to older missions.
- Every observation benefits from the context provided by other instruments and other missions.
- New missions making original measurements require other spacecraft for context.

The astrophysics community responded to the need for new data by defining "Great Observatories" and these have made new and exciting measurements. For in situ heliospheric physics the single Great Observatory is the fleet of technically diverse and spatially separated missions that provide the measurements needed to understand the coupled nonlinear subsystems that together form the whole of the heliosphere.

The High-Energy Sun at High Sensitivity: a NuSTAR Solar Guest Investigation Program

A concept paper submitted to the Space Studies Board Heliophysics Decadal Survey
by David M. Smith, Säm Krucker, Gordon Hurford, Hugh Hudson, Stephen White,
Fiona Harrison (NuSTAR PI), and Daniel Stern (NuSTAR Project Scientist)

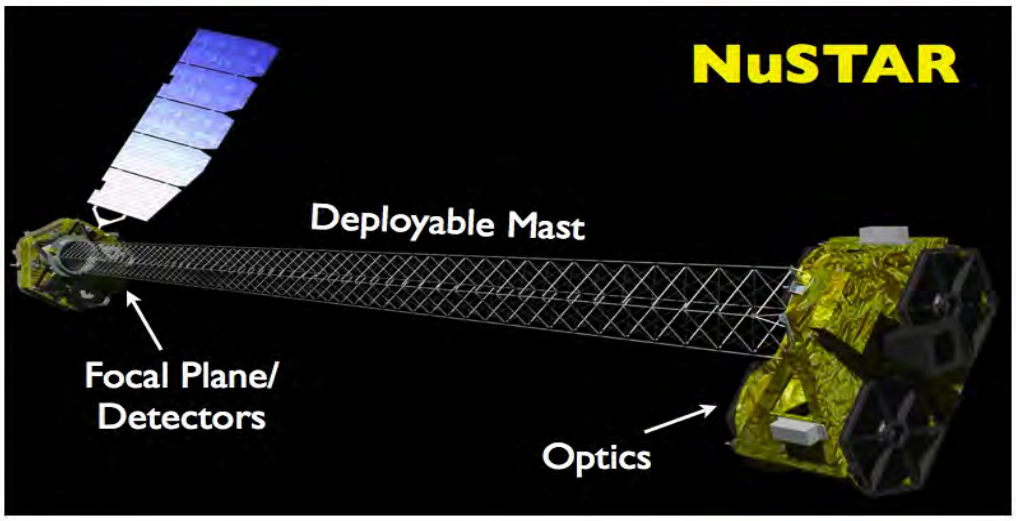


Figure 1: Artist's conception of the NuSTAR spacecraft

Introduction and Overview

Understanding the origin, propagation and fate of non-thermal electrons is an important topic in solar and space physics: it is these particles that present a danger for interplanetary probes and astronauts, and also these particles that carry diagnostic information that can teach us about acceleration processes elsewhere in the heliosphere and throughout the universe. Such non-thermal electrons can be detected via their X-ray or gamma-ray emission, their radio emission, or directly in situ in interplanetary space. To date, the state of the art in solar hard x-ray imaging has been the *Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI),* a NASA Small Explorer satellite launched in 2002, which uses an indirect imaging method (rotating modulation collimators) for hard x-ray imaging. This technique can provide high angular resolution, but it is limited in dynamic range and in sensitivity to small events, due to high background. High-energy imaging of the active Sun is technically easiest in the soft x-ray band (see, for example, the detailed and dynamic images returned by the soft x-ray imager on *Hinode,* as well as the corresponding telescopes on *Yohkoh* and *GOES),* but these x-rays are generally emitted by lower-energy thermal plasmas.

The purpose of this paper is to draw the attention of the solar community to a new mission that can address these topics with a sensitivity far superior to any previous detector. The *Nuclear Spectroscopic Telescope Array (NuSTAR;* Harrison et al. 2010*)* is another Small Explorer, currently being prepared for launch in February 2012, and aligned with NASA's Astrophysics Division. *NuSTAR* is a true hard x-ray telescope, using grazing-incidence optics with multilayer reflective coatings to make true images with 7.5" FWHM angular resolution over the 5-80 keV energy band. *NuSTAR*'s primary science goals concern supernovae, black holes, and pulsars, but, unlike virtually every other high-energy astrophysics mission to date, it is capable of being pointed at the Sun, a plan which has the support of the *NuSTAR* PI, Project Scientist, and science team. *NuSTAR*'s effective area at 10 keV is over ten times that of *RHESSI* (800 cm^2 versus 60 cm^2), but, even more importantly, its background is four orders of magnitude lower, thanks to the use of focusing optics. Overall, *NuSTAR* will be able to observe solar hard x-ray sources **200 times fainter** than our current capabilities allow. This sensitivity will open the door to brand new observations, such as the coronal hard x-rays related to type III radio bursts, and will allow comparison of the non-thermal characteristics of flares over eight orders of magnitude, from the brightest X flares seen by *RHESSI* to the faintest nanoflares visible to *NuSTAR* at a level about 200 times fainter than GOES class A1. Since flares increase in number as flux decreases (at least to the limit we can now observe), *NuSTAR's* high sensitivity also means that a large database of hundreds of (nano- and micro-) flares can be accumulated in days or weeks, as opposed to the years required to accumulate a library of C, M, and X flares.

During *NuSTAR's* primary science mission, in the years 2012 and 2013, approximately two weeks of the fully-allocated observing plan will be dedicated to solar observations. These will be planned by the *NuSTAR* solar science team (the authors of this concept paper, pending final approval by the *NuSTAR* Science Steering Committee). But we consider it important for *NuSTAR* solar science to be explicitly supported during the extended mission of the satellite, which we trust will be long and valuable (there are no spacecraft expendables). We propose here that the SSB recommend to NASA the establishment of a solar Guest Investigator (GI) program for *NuSTAR* as long as its extended mission is approved by the Astrophysics Senior Review. A formal GI program would:

1. Assure the participation of the wider solar community, with the inclusion of new ideas and science payoffs not anticipated by the *NuSTAR* solar science team,
2. Support dedicated observing campaigns at other wavelengths to exploit the full potential of the hard x-ray observations in a broader context, and
3. Make sure that solar GI proposals are not disadvantaged by competing against astrophysics proposals in panels composed of astrophysicists who may not be knowledgeable about the solar measurements.

Examples of *NuSTAR* Solar Science

1) Active region microflares

non-thermal emission from active-region microflares is most easily detected at photon energies around 10 keV (Benz & Grigis 2002, Krucker et al. 2002), an energy perfectly suited to NuSTAR. The smallest RHESSI microflares with statistics sufficient to analyze the photon spectrum (i.e. with about 1000 counts) contain about 10^{27} ergs in non-thermal electrons. Assuming that the spectral shape of smaller flares is similar, the enhanced sensitivity of NuSTAR will measure the energy spectra of microflares with energies down to $5x10^{24}$ ergs. This energy is about 8 orders of magnitude smaller than the largest flare seen, and lies in the so-called nanoflare range. NuSTAR will extend the current hard X-ray (HXR) microflare distribution to lower energies by at least two orders of magnitude, and provide new information on active-region microflares. Further, it will permit comparison of active-region nanooflares with the quiet-region network flares discussed in the next section.

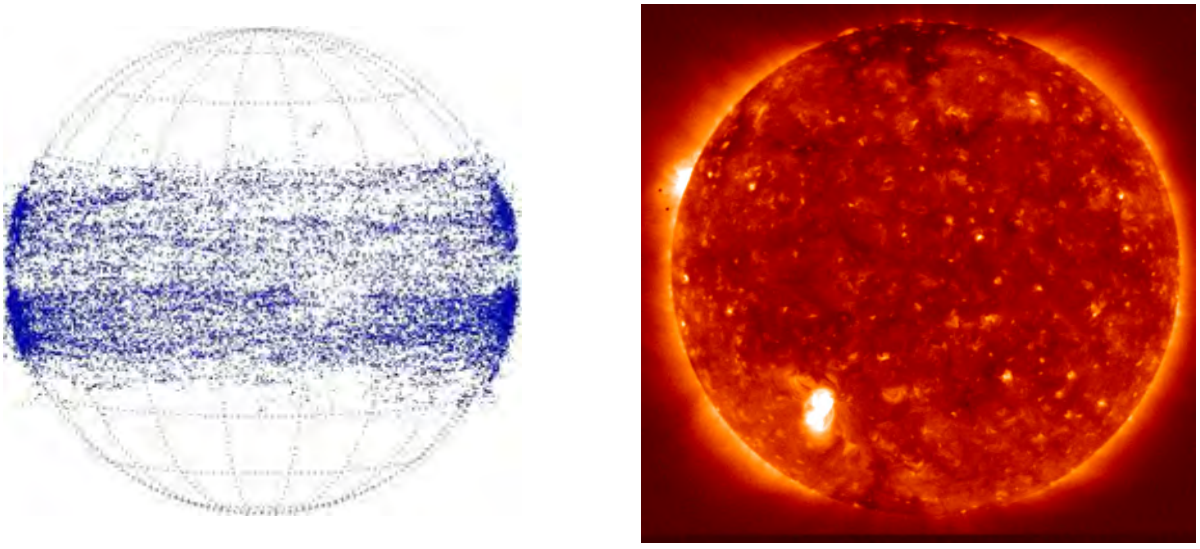


Figure 2. Left: map of hard x-ray microflares associated with active regions observed by RHESSI *(Christe et al. 2008). Right: soft x-ray bright spots in a* Hinode *image of the solar disk (NASA image).* NuSTAR *will have the sensitivity to see if the latter events have a hard x-ray component.*

2) Non-thermal electrons in the quiet corona

The high sensitivity of *NuSTAR* will allow a search for HXR counterparts of network flares seen in the quiet corona outside of active regions (Krucker et al. 1997,

Berghmans et al. 1998, Krucker & Benz 2000). From soft X-ray (SXR) and extreme ultraviolet (EUV) observations, the thermal energy content is estimated to be up to 10^{26} ergs (Krucker & Benz 1998, Parnell & Jupp 2000, Aschwanden et al. 2000). Radio observations have revealed the existence of non-thermal electrons in at least some network flares (Krucker et al. 1997), suggesting that network flares are heated in a fashion similar to regular flares occurring in active-regions. If network flares are indeed smaller versions of regular flares, the thermal heating seen in EUV and SXR is provided by non-thermal electrons. With the sensitivity of *NuSTAR* allowing us to observe non-thermal events with energies down to $5x10^{24}$ ergs, non-thermal bremsstrahlung of these electrons should be easily detected with *NuSTAR*.

EUV observations from EIT and TRACE predict about 2 to 250 network flares per minute above 10^{25} ergs in a field of view of 600×600 arcsec. While *NuSTAR* will not resolve the spatial structure of network flares, it will allow the separation of source locations, and simultaneous EUV observations taken by *SDO*/AIA, *STEREO*/EUVI and *Hinode* could help to resolve ambiguities. Spectral analysis will provide averaged values for the energy in non-thermal electrons in these events even when they overlap.

3) Tracing electron beams in the solar corona

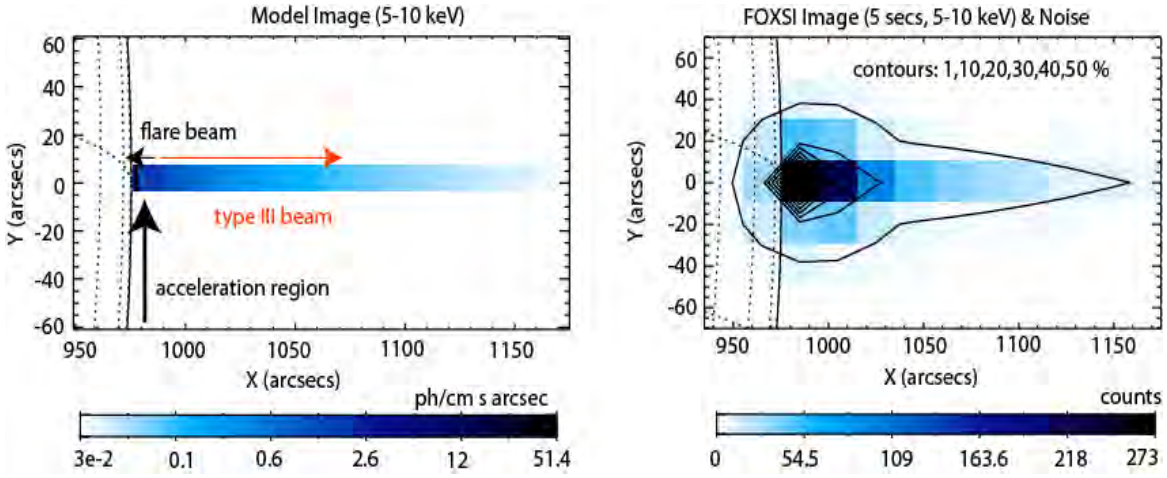


Figure 3: Simulated image of the electron beam associated with a type III burst, calculated for the FOXSI sounding rocket by Pascal Saint-Hilaire and Säm Krucker (adapted from Saint-Hilaire et al. 2009). NuSTAR will have over 5 times the effective area of the FOXSI rocket, and images to higher energies.

With the sensitivity of *NuSTAR*, it will be possible to detect HXR emission from electrons that produce coherent radio bursts seen at meter and centimeter wavelengths (e.g. Bastian, Benz, & Gary 1998). The best understood of these bursts are type III bursts, which are produced by electron beams in the corona. Although these beams

generally travel outwards towards low densities, they also radiate in X-rays albeit at a much lower rate than downward moving flare beams. The X-ray observation of a type III beam is therefore difficult for two reasons: the relatively small fluxes and the large dynamic range necessary since such bursts are often associated with flares. Assuming that the electron beam consists of 10^{33} electrons above 5 keV (consistent with in-situ electron measurements near 1 AU, Lin 1973) and propagate away from the Sun beginning from a density of 5×10^{9} cm^{-3} (corresponding to a plasma frequency of 600 MHz), *NuSTAR* will detected about 250 counts per second. Therefore *NuSTAR* will make the first HXR observations of a type III-producing beam. The hard X-ray emission will outline where the electrons are accelerated and along which path they escape from the Sun. More importantly, the *NuSTAR* spectra will permit determination of the energy distribution of the emitting electrons close to their source in the low corona. Radio measurements do not constrain the electron content of type III bursts because the emission is coherent. Joint radio and hard x-ray measurements of type III bursts will allow the electron content of these events to be calculated for the first time, and permit comparison of the low-coronal spectrum and intensity with the energy distributions that have been measured for decades at two sites: in situ at 1 AU, after what we believe to be extensive energy redistribution in the course of beam propagation, and at flare footpoints, by which time there may also be reacceleration, Coulomb losses, and other effects in the flare loops.

4) Non-thermal electrons associated with coronal mass ejections

For flares occurring 20 or more degrees behind the solar limb, not only the HXR footpoints but also the main thermal and non-thermal HXR emission from the corona are occulted, and flare-related emission from the high corona (>200 arcsec above the flare site) can be seen (Hudson 1978, Kane et al. 1992, Hudson et al. 2001, Krucker et al. 2007a). *RHESSI* provided the first high-resolution imaging spectroscopy of HXR emission from such events, revealing large sources (with sizes >200 arcsec) that expand and move outwards. The upward motion of the hard X-ray source is slower than the speed of the CME front, but similar to the speed of the filament seen behind the CME front. This suggests that energetic electrons are being advected outwards in magnetic flux tubes related to the CME. This is also the scenario for the expanding loops observed in microwave images by Hudson et al. (2001) in conjunction with an event far beyond the west limb. Multi-spacecraft observations (Kane et al. 1992) have already shown that such HXR emissions from the high corona could occur during the impulsive phase of the flare, simultaneously with the HXR footpoint emissions.

Krucker et al. (2007a) found evidence that all fast ($v > 1500$ km s^{-1}) farside CMEs that originate from flares occulted by 20° to 45° may show related HXR emissions from the high corona. With the sensitivity of *NuSTAR*, hard X-ray emission from the much more frequently occurring slower CMEs should be detectable. Hence, even with only a few

days of observations, several occulted CME should be seen. Since these sources are extended, *NuSTAR* imaging provides appropriate spatial resolution.

5) The quiet Sun

NuSTAR will make a giant improvement in our understanding of the hard X-ray spectrum of the quiet Sun. Several mechanisms outside the normal plasma physics of flares and solar magnetic activity could contribute; specifically we note cosmic-ray effects (both extrasolar and solar-system), as seen for example in the Fermi observations of inverse Compton scattering of sunlight from cosmic-ray electrons (Orlando et al. 2009). The current (RHESSI) limits on hard x-ray flux from the quiet Sun (Hannah et al., 2010) are somewhat improved over those derived earlier (Churazov et al. 2008), but *NuSTAR*'s sensitivity will suddenly drop these numbers into the domain of the albedo source from the diffuse cosmic x-ray background – the Sun will actually appear darker against the bright cosmic background, shining primarily with Compton back-scattered cosmic x-ray light. We also note the possibility of axion emission from the solar core, which could produce an X-ray signal via Primakoff conversion in the solar magnetic field (Carlson & Tseng 1996).

6) The earliest stages of flares

NuSTAR's high sensitivity will give an unprecedented look at particle acceleration in the precursor period to small, medium, and large flares when active regions are observed on the disk or at the limb. By comparing these observations with other wavelengths – such as EUV observations, which often saturate when flares reach their impulsive phase – it may be possible to identify the onset of particle acceleration and the rate and method of its expansion.

Scope of the proposed program

Based on the size of the qualified community, we might expect to receive 15-30 proposals per year for a *NuSTAR* GI program, once the capabilities of the instrument become well known thanks to the guaranteed solar observations during the primary mission. We expect to be limited more by the amount of observatory time available to be dedicated to solar observing than by the availability of good proposals. Let us say, then, that between 5 and 10 proposals should be funded each year. There may be only time for a few individual pointings on the order of a day or two, but each of these may observe a variety of phenomena, and solar GI awardees would be chosen and funded based on their proposed exploitation of the data, not just for novel observing strategies, although those may be considered as well. Consistent with the policy for all *NuSTAR* observations, no data would be proprietary, so multiple uses could be made of each

pointing, and multiple investigators funded. Individual awards should be large enough to fund the acquisition and analysis of simultaneous ground-based and other space-based observations, and/or directly relevant theory and modeling efforts, as well as the analysis of *NuSTAR* data. We suggest that approximately $500K per year should be available to fund the 5-10 selected proposals, of which 1-3 could be two-year proposals to be funded out of that total, if exceptional reasons warranted it. The management of the GI program should be negotiated among the Heliophysics Divsion, the Astrophysics Division, and the *NuSTAR* Science Steering Committee. It could be administered as a panel of solar specialists within the context of the regular *NuSTAR* GI competition, with the only difference being the source of funding for accepted solar proposals, or it could be administered separately by the Heliophysics division, with a different procedure and deadline, with *NuSTAR* team members brought in for technical support of the review. In the latter scenario, the total time available for *NuSTAR* solar observations would have to be negotiated among the parties ahead of time.

Conclusion

NuSTAR's 200-fold advance in hard x-ray sensitivity over current capabilities will open up entirely new parameter spaces for the study of particle acceleration at the Sun. Because the Sun produces ever more microflares at lower and lower luminosity in the range where we can now observe it, extrapolation leads us to expect that even short observations at these sensitivities will return a wealth of interesting events; and, if not, that will imply the discovery of a surpising and informative threshold effect for particle acceleration. The few weeks of *NuSTAR* solar pointings may therefore provide a science return approaching that of a dedicated mission. We believe the best way to assure this and harness the ingenuity of the solar community is the dedicated solar GI program outlined here.

REFERENCES

Aschwanden, M. J., Nightingale, R.W., Tarbell, T. D. andWolfson, C. J. (2000), 'Time Variability of the "Quiet" Sun Observed with TRACE. I. Instrumental Effects, Event Detection, and Discrimination of Extreme-Ultraviolet Microflares', ApJ 535, 1027–1046.

Bastian, T. S., Benz, A. O. and Gary, D. E. (1998), 'Radio Emission from Solar Flares', ARA&A 36, 131–188.

Benz, A. O. and Grigis, P. C. (2002), 'Microflares and hot component in solar active regions', Sol. Phys. 210, 431–444.

Benz, Arnold O.; Grigis, Paolo C.; Csillaghy, André; Saint-Hilaire, Pascal (2005), 'Survey on Solar X-ray Flares and Associated Coherent Radio Emissions', Solar Physics, 226, 121

Berghmans, D., Clette, F. and Moses, D. (1998), 'Quiet Sun EUV transient brightenings and turbulence. A panoramic view by EIT on board SOHO', A&A 336, 1039–1055.

Carlson, E. D. and Tseng, L. (1996), 'Pseudoscalar conversion and X-rays from the sun,' Phys. Lett. B. 365, 193

Christe, S., Hannah, I. G., Krucker, S., McTiernan, J. and Lin, R. P. (2008), 'RHESSI Microflare Statistics. I. Flare-Finding and Frequency Distributions', ApJ 677, 1385–1394.

Churazov, E., Sazonov, S., Sunyaev, R., and Revnivtsev, M. (2008), 'Earth X-ray albedo for cosmic X-ray background radiation in the 1-1000 keV band', MNRAS, 385, 719

Hannah, I. G., Hudson, H. S., Hurford, G. J., and Lin, R. P. (2010), 'Constraining the hard X-ray properties of the quiet Sun with new RHESSI observations', astro-ph 1009.2918

Harrison, F.A., et al. (2010): 'The Nuclear Spectroscopic Telescope Array (NuSTAR)', SPIE, 7732, 27

Hudson, H. S.: 'A purely coronal hard X-ray event', 1978, ApJ, 224, 235

Hudson, H. S.; Kosugi, T.; Nitta, N. V.; Shimojo, M.: 'Hard X-Radiation from a Fast Coronal Ejection', 2001, ApJL, 561, L211

Kane, S. R.; McTiernan, J.; Loran, J.; Fenimore, E. E.; Klebesadel, R. W.; Laros, J. G.: 'Stereoscopic observations of a solar are hard X-ray source in the high corona', 1992, ApJ, 390, 687

Krucker, S. and Benz, A. O. (1998), 'Energy Distribution of Heating Processes in the Quiet Solar Corona', ApJ 501, L213+.

Krucker, S. and Benz, A. O. (2000), 'Are Heating Events in the Quiet Solar Corona Small Flares? Multiwavelength Observations of Individual Events', Sol. Phys. 191, 341–358.

Krucker, S., Benz, A. O., Bastian, T. S. and Acton, L. W. (1997), 'X-Ray Network Flares of the Quiet Sun', ApJ 488, 499–+.

Krucker, S., Christe, S., Lin, R. P., Hurford, G. J. and Schwartz, R. A. (2002), 'Hard X-ray Microflares down to 3 keV', Sol. Phys. 210, 445–456.

Krucker, Säm; White, S. M.; Lin, R. P. (2007, 'Solar Flare Hard X-Ray Emission from the High Corona', ApJ, 669, L49

Lin, R. P., Evans, L. G. and Fainberg, J. (1973), 'Simultaneous Observations of Fast Solar Electrons and Type III Radio Burst Emission Near 1 AU', Astrophys. Lett. 14, 191

Lin, R. P., et al. (2002), 'The Reuven Ramaty High-Energy Solar Spectroscopic Imager (RHESSI)', Sol. Phys. 210, 3–32.

Orlando, E., and Giglioetto, N. for the Fermi Large Area Telescope Collaboration (2009), 'Fermi-LAT Observation of Quiet Solar Emission', arXiv 0912.3775

Parnell, C. E. and Jupp, P. E. (2000), 'Statistical Analysis of the Energy Distribution of Nanoflares in the Quiet Sun', ApJ 529, 554–569.

Saint-Hilaire, P., Krucker, S., Christe, S. and Lin, R. P. (2009), 'The X-ray Detectability of Electron Beams Escaping from the Sun', ApJ 696, 941–952.

4PI: A Global Understanding of the Solar Cycle

Keith Strong[1], Julia Saba[1], Dean Pesnell[2], Janet Luhman[3], Frank Hill[4], and Tom Duvall[2]
[1]SP Systems at NASA GSFC, [2]GSFC, [3]UCB, [4]NSO

4PI is a flagship-class mission (>$1B) that is baselined as a 4-satellite constellation designed to measure the magnetic field over the entire Sun and the solar structure and dynamics below the surface, at the surface, and above the surface, and track their evolution continuously. In our nominal configuration, three satellites would orbit the Sun in the ecliptic plane to facilitate local helioseismology studies. Their circular orbits would be less than 1 AU from the Sun so that they would periodically align with Earth instruments for cross-calibration; this would also allow them to track regions on the Sun with optimal coverage for longer periods than from Earth orbit. The fourth satellite would be in a high-inclination orbit to return never-before-observed data from the solar polar regions. All four satellites would return helioseismology, magnetic field, and coronal measurements. These data would be combined to produce full-Sun magnetic field maps, subsurface dynamics, and the ever evolving field configuration in the solar atmosphere. A bold modeling and analysis initiative is proposed in parallel with the mission development to enable mission planning, data assimilation, and the development of comprehensive physics-based solar cycle models.



Dopplergrams, magnetograms, and coronal imaging comprise a powerful combination of observations. Helioseismology measures solar oscillations and thus gives insight into the structure and dynamics of the solar interior. The photospheric magnetic fields show the evolution of both strong and weak fields that make it to the surface of the Sun. Coronal imaging shows the complex interconnections of the fields and eruptive processes that help dissipate those fields. 4PI will produce such data continuously, simultaneously covering all longitudes and the polar regions.

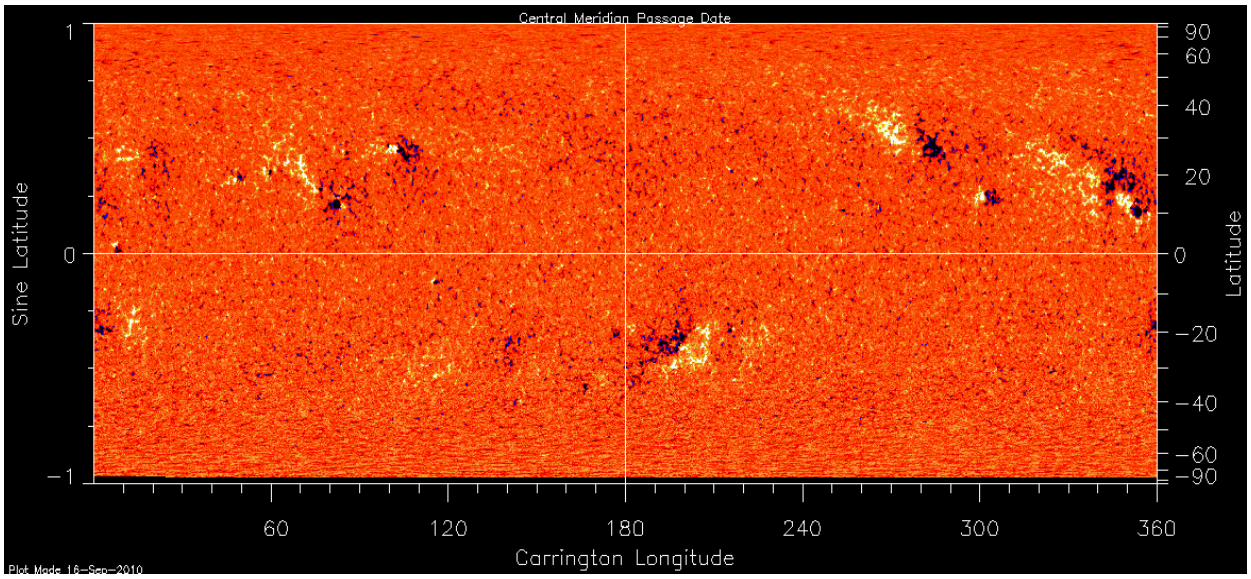
Introduction

For over 170 years, solar physicists have been trying to understand and predict the solar cycle. The surprising evolution of Cycle 23 demonstrated that we have not yet achieved a comprehensive understanding of how the Sun generates, transports, and dissipates magnetic field. Of over 80 published predictions using a plethora of data gathered during Cycle 23 and before, none forecast the extended, deep minimum**.** This inability to anticipate the behavior of

the Sun during the recent minimum shows that we continue to lack the necessary understanding of how the solar dynamo works.

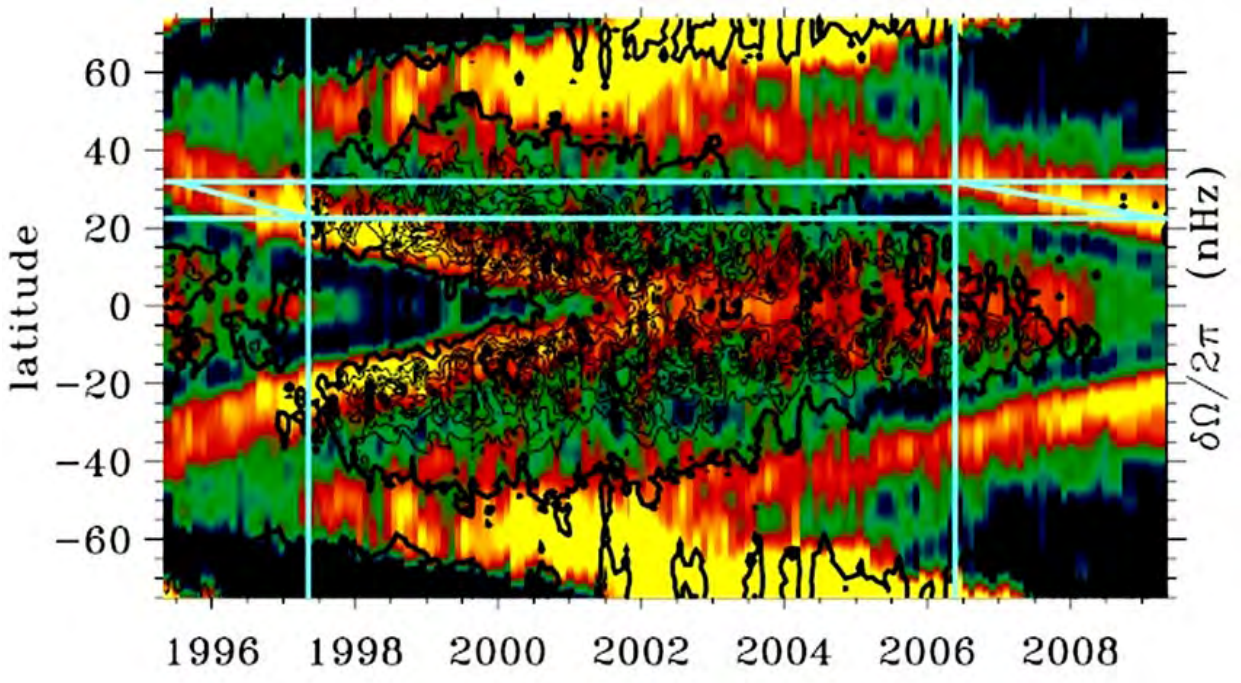
Many of the empirical forecasts of the solar cycle are based on historical data such as sunspot number and area. However, sunspot number data are of questionable quality and describe only a very limited part of the many different manifestations of the solar magnetic cycle. More quantitative measures of solar activity (e.g., F10.7, line-of-sight magnetograms, Ca K, green-line corona, soft X-ray flux, and total solar irradiance) go back only a few cycles and thus cannot yet provide a basis for accurate long-term forecasting. Besides, many of these indicators have relative calibration problems of their own. At the recent SHINE 2010 workshop, a group of solar cycle researchers agreed that to solve the solar dynamo problem we need accurate, long-term measurements of both the solar surface and coronal magnetic field and the solar interior dynamics combined with a robust theory and modeling initiative. We do not have these at present, and, while there are new solar facilities coming on-line, it is unclear that they have the capability to reveal the solar dynamo's workings definitively. We must start now to lay a firm observational and theoretical foundation to enable future solar physicists to solve this problem.

Given the great advances of the last decade from magnetography and helioseismology on SOHO, now augmented by Hinode and SDO, coupled with ground-based observations by SOLIS and GONG, where are the shortcomings in our understanding of the solar magnetic cycle? The primary issue is that we observe less than half the Sun's surface magnetic field at any given time and never get a clear simultaneous view of the polar fields. Thus, any comprehensive map of the global photospheric fields includes data that can be over two weeks old and has a potentially important area (the poles) missing or ill determined at best.



A map of the solar surface magnetic field configuration for Carrington Rotation 2100 (August 2010) from SOHO/MDI. It currently takes over 27 days to build up such a map so most of it is obsolete by the time the map is completed as solar features evolve on far more rapid timescales than a rotation period. Moreover, the polar regions above 60° latitude are not observed well, so the data are unreliable. 4PI will be able to produce such maps as quickly as the slowest cadence data (minutes); we will therefore be able to follow in detail the full evolution of every region as well as the global field.

Moreover, although these magnetic fields drive the observed solar surface and coronal variability, their origins lie below the surface. Only helioseismology can plumb those depths to look at the subsurface flows that help transport the magnetic flux generated in the dynamo and give possible access to and characterization of the dynamo region itself. However, currently helioseismology is hampered by our being able to view only a fraction of the solar surface. This introduces artifacts into the observed power spectrum of the oscillations, reducing the accuracy of the frequency measurements and possibly creating systematic errors in the inferred subsurface flows. The 4PI mission would eliminate these effects, provide measurements of the flows below the polar regions, and enable the direct correlation of the oscillations at diametrically opposite solar locations, which would allow direct probing of the deep solar interior.



The evolution of torsional oscillations inside the Sun at a depth of 1 Mm as a function of latitude. The black contours represent the surface magnetic field, appearing in close proximity to the position of the torsional oscillation. The reconstructions use spherically symmetric global modes. Note the first indications of Cycle 24 at high latitude in 1997 but the appearance of the first Cycle 24 sunspots over a decade after. [Figure from F. Hill et al., "Helioseismic Observations of Solar Convection Zone Dynamics," Astrophysical Dynamics: From Stars to Galaxies Proceedings IAU Symposium No. 271, 2010 (A.S. Brun & N. Brummel, eds.)]

While the new generation of helioseismic and full-disk vector magnetic field measurements from SDO will probe the solar interior in unprecedented detail, spatial resolution, and depth, the problems of the far side and polar behavior will remain. The planned Solar Orbiter is expected to obtain some better snapshots of the high-latitude solar magnetic field, but it will not provide the critical sustained and comprehensive look over a solar cycle.

Scientific Questions

4PI will address many aspects of the fundamental scientific questions that need to be solved to understand the riddle of the solar cycle:

- Why does the Sun's magnetic field vary from cycle to cycle?

- How and where does the Sun generate magnetic fields?
- How are those fields transported at various depths, to the surface, and at the surface?
- How do the fields evolve, e.g., how and where are they dissipated, and over what timescales?
- Why do the photospheric magnetic fields in the Sun's northern and northern hemispheres differ significantly in both timing and strength?
- Why is the solar cycle quasi-periodic at about 11 and 22 years?

Programmatic Flow-down Requirements:

To answer these questions it is clear that we need to measure the solar magnetic field at high temporal and spatial resolution over the entire surface of the Sun continuously over an extended period. To complement those data we need the Dopplergrams that provide the global and local subsurface dynamics needed to characterize and quantify the dynamo and transport mechnaisms. These combined data would reveal for the first time how the subsurface transport mechanisms are distributed globally and how they evolve as a function of latitude and longitude over a solar cycle. We also need to observe at least the morphology of the coronal magnetic fields in order to connect the magnetic evolution below, in, and above the photosphere.

Observational Requirements

The required mission would need to obtain multi-perspective views of the solar disk on a regular basis for a minimum of a solar cycle. The most important requirement is for optimal coverage of the entire solar surface. Incomplete coverage results in spatial aliasing, making it impossible to completely isolate individual modes of oscillation.

Current GONG and SOHO/MDI helioseismic measurements have been limited to a maximum extent of ±60° in latitude and longitude; beyond this, foreshortening caused by the spherical solar surface precludes accurate frequency measurements. Magnetic-field measurements are also subject to projection effects close to the solar limb and so are of limited precision. The ±60° limitation means that only 22% of the solar surface is usable for helioseismology and magnetic-field measurements. Thanks to its high-resolution cameras, SDO/HMI can reach ±80° in longitude and latitude, but even this represents less than 40% of the surface. For global helioseismology, the limited viewing surface introduces artifacts known as spatial side lobes into the oscillation spectrum, and modeling of these is a major challenge. For local helioseismology, the restricted viewing area means that we cannot infer the flows at the poles, where the meridional flows that are important aspects of the solar cycle must converge. Finally, 4PI would make possible the direct correlation of the oscillations at diametrically opposite solar locations, which would allow tomographic imaging of the solar tachocline, radiative envelope, and core.

The absence of reliable measurements of the polar surface magnetic field introduces large systematic errors in extrapolations of the field into the corona and beyond. These errors are propagated into models and predictions of the solar wind speed and direction, degrading space weather forecasts. Thus, accurate polar magnetic field measurements would greatly improve our ability to predict space weather, as well as provide an accurate global view of the magnetic field free of assumptions about its polar distribution and evolution.

It is quite difficult to observe structures deep in the solar convection zone, as one would have to average the data from many active regions to detect flows from rising flux tubes at a depth of only 30 Mm. To detect the return flow of meridional circulation near the tachocline of a

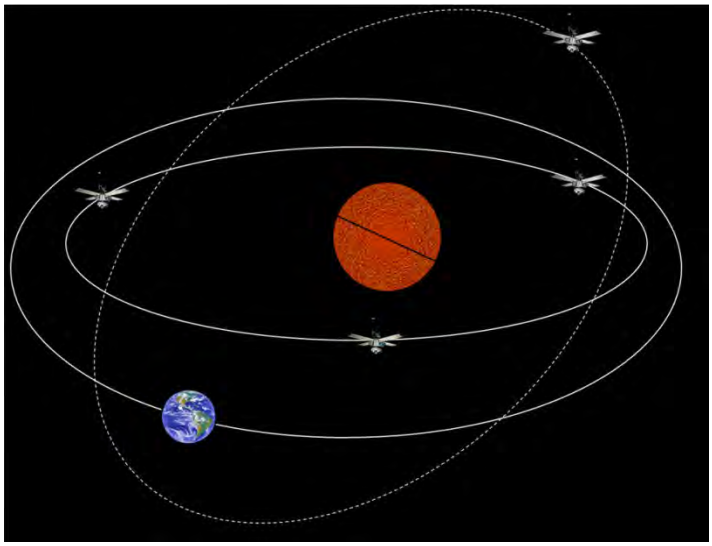
few meters per second would require averaging about a solar cycle's worth of data. In these measurements we are limited by solar realization noise, which is caused by randomly distributed wave sources over the solar near-photosphere. This noise can be reduced by averaging over a larger fraction of the sphere. Conceivably the emergence of three times as many active regions could be observed by having complete coverage in longitude. This would make the measurement of the flows quicker. Also, increasing the signal-to-noise ratio of the return flow of meridional circulation might make the difference between seeing time variation over a solar cycle and not.

The basic imaging instruments need to have a time cadence of no longer than 1 minute and a spatial resolution of 1 arcsec or better. Ideally, the Doppler velocity measurements should have a noise level of less than 15 m/s per pixel per minute, and a disk-average noise level less than 0.5 m/s per minute. The absolute accuracy of the velocity measurement should be better than 10 m/s per minute. The longitudinal magnetic-field noise-level goal is 10 G per pixel per minute with a zero-point accuracy of 0.1 G. For vector magnetic-field measurements, the polarization precision should be no worse than 0.05%. These measurement precision and accuracy goals are ambitious, representing an improvement over the SDO/HMI requirements of about a factor of 3 to 5.

The coronal imager should be similar to those on SDO and needs at least two channels – one to show the low T (<1MK) transition region and another for the active corona (>2MK). It should be possible, if sufficient funds were available, to fly other instruments on these same platforms. Obvious candidates would be particle and field detectors similar to those on ACE or Ulysses, irradiance monitors, and coronagraphs like those on STEREO; however, their inclusion must not compromise the primary mission goals.

A Notional 4PI Mission Profile

The baseline mission concept would be to have three of the spacecraft in the ecliptic but in an orbit somewhat inside 1 AU so that the spacecraft would occasionally align with the Earth-Sun line, allowing the instrument response to be calibrated against ground-based magnetographs and



There are many possible orbital configurations for the 4PI mission. The illustrated configuration is for a negative B angle so that the southern polar region is more visible from Earth while the out-of-the-ecliptic spacecraft, 180º out of phase with the Earth, observes the northern polar region. This arrangement means that at most times at least one spacecraft is observing any given area of the Sun and sometimes up to three spacecraft can observe the same limited areas, increasing the fidelity of 3D reconstruction. Other solar-observing assets, particularly ground-based instruments, would provide additional views as well as relative calibrations.

Dopplergraphs. The out-of-the-ecliptic spacecraft would have the same orbital period as the Earth so that it could be phased to observe the polar regions not visible from the Earth due to the

Sun's variable B angle. Regions where the fields of view of the various spacecraft overlap either would be points of cross calibration or could be used to get vector fields. These data would provide daily source surface maps and continual monitoring of newly emerging active regions.

The 4PI concept would be developed and operated under the NASA LWS program, but once its scientific merit is established it could eventually be turned over to NOAA to maintain, enhance, and perpetuate the instrumental constellation. The 4PI measurements could become the standard calibration source of all such instruments. Having 4π-steradian solar coverage would be useful for space weather forecasting when human explorers venture beyond the Earth-Moon system. Such a network should be the minimum requirement to ensure astronaut safety from solar events on round trips to Mars and other places beyond Earth orbit.

The 4PI Theory Initiative

There is little point in developing a bold, new observational concept if there is not an equivalently sweeping data analysis, modeling, and theory program to guide acquisition of and take advantage of the new data. When the meteorological community was faced with the challenge of understanding what was happening to the Earth's climate, they initiated a comprehensive climate modeling initiative at several centers around the country. Now they have dozens of different codes that model various aspects of our climate, which can be used to give appropriate scenarios to predict future directions of climate evolution. Improvements to those codes are still being developed 20 years after the initiation of the program, which demonstrates the need for a long-term commitment to such a program. A physical understanding of the solar activity cycle sufficient to lay the groundwork for more accurate prediction is at least a similarly complex problem and demands a similarly comprehensive approach.

We are proposing that a solar cycle modeling initiative be started at the onset of the 4PI program, initially to help guide decisions as to exactly what to observe and at what cadence and resolution to observe it. The fundamental goal of the program would be to develop comprehensive physics-based models of the solar cycle that can be tested with the 4PI observations and eventually be used to predict solar cycle evolution, possibly many years in advance. Theory and observation have a symbiotic relationship in the context of such an uncertain environment – each one forcing advances in the other. The development of such theories and modeling will almost certainly reveal new observational challenges. New observations will challenge modeling in the assimilation of large amounts of data in a dynamically changing situation such that developments can be predicted on a useful timescale.

Technology Challenges

There are several basic technical challenges to the 4PI concept:

- **Communications:** Transferring large amounts of data back from deep-space probes at varying distances. This problem is being worked by the STEREO team and the space exploration community. 4PI could be a proving ground for those technologies.
- **Computing:** The data handling and modeling will need a significant upgrade of our community's computer power and storage capability. There will be many valuable lessons to be learned from the SDO, STEREO, and Solar Orbiter missions as well as from the terrestrial global modeling community in this regard.
- **Calibration:** It will be vital to maintain detailed inter-calibration of the 4PI instruments so that the data can be realistically compared over an extended mission timeframe. A subgroup

dedicated to this problem is needed during the development, operations, and analysis stages of the program.

- **Orbital design:** In comparison to terrestrial observations, the tailoring of solar observing orbits is still immature. The initial phase of the 4PI mission will explore the tradeoffs between, for example, being close to the Sun (for better tracking of surface velocity patterns) and the difficulty of moving satellites into such interior orbits. The trade study of optimal coverage orbits with cost-effectiveness for different instrument configurations will be required at an early stage of the mission development.

Concluding Remarks

Many issues remain to be resolved concerning the details of the 4PI mission, such as optimum orbital configuration, payload details, mission costs, data analysis and computational requirements, schedule, and the nature and timing of the theory/modeling program, which will drive and derive the fruits of the mission. These and many other aspects of the mission need to be the subject of a detailed trade study involving technical experts and spacecraft engineers as well as instrumentalists, theorists, modelers, and data analysts to find the most cost effective and scientifically productive overall mission plan.

The primary benefits of 4PI are:

- 4PI provides a full-Sun view of the magnetic structure of the Sun including the polar regions that may be key

- 4PI provides full-coverage helioseismic measurements with unprecedented accuracy and precision

- The multi-directional coverage of coronal structures afforded by 4PI will allow improved 3-D modeling of coronal loops and better tracking of their dynamics and evolution

- 4PI can follow the development of active regions, coronal holes, filaments, and other solar features from birth to death, giving clearer insight into their roles in the solar activity cycle

- These unique 4PI observations would provide the basis for a revolution in the development of physics-based solar activity models

- The 4PI theory initiative gives an opportunity to develop operational codes that will significantly increase the chances of one day obtaining accurate predictions of the solar cycle

- 4PI would act as the pathfinder for a deep-space space-weather warning and communications system that should be a fundamental requirement for the Space Exploration Initiative.

Thus 4PI would provide the first-ever comprehensive view of the solar magnetic field and the structure and dynamics from beneath the solar surface into the corona and be able to track their development throughout at least one solar cycle.

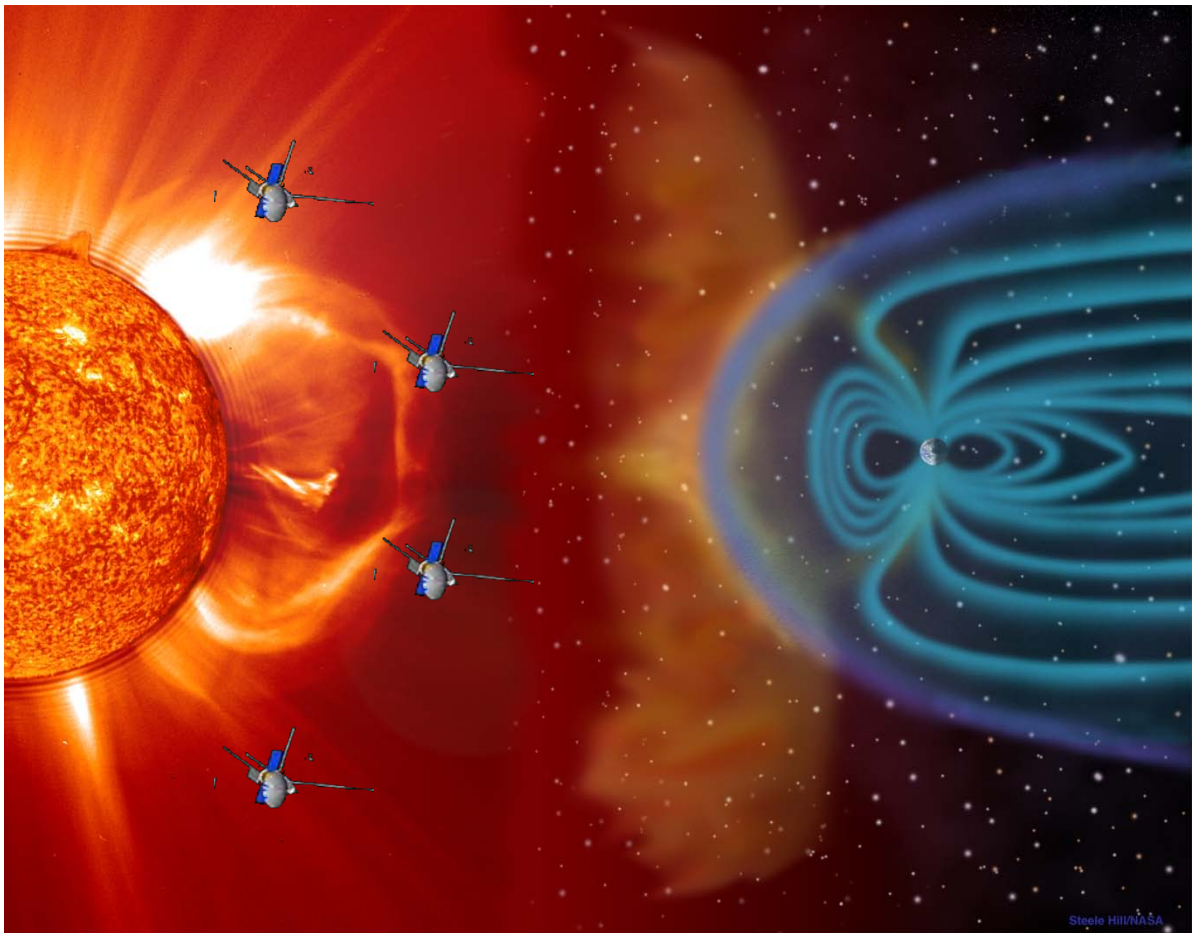
Energetic Particle Propagation and Coronal Mass Ejection Evolution in the Inner Heliosphere

Multi-point in-situ Solar Sentinels Observations

Adam Szabo
NASA Goddard Space Flight Center

Robert P. Lin
Space Science Laboratory
University of California, Berkeley

and
The Solar Sentinels STDT



Despite the significant advances in our knowledge and understanding of the Sun-heliosphere connection that have occurred over the past four decades, a number of fundamental questions remain to be answered. Our ability to answer those questions has been significantly limited by the fact that, except for the pioneering but relatively primitive observations made by Mariner 10 (1973), Helios 1 (1974-1986) and Helios 2 (1976-1980), the majority of the in-situ measurements of the solar wind, the interplanetary magnetic field (IMF), and solar energetic particles (SEPs) have been made at 1 AU or beyond. At these distances from the Sun, however, the solar wind plasma, fields, waves and SEPs have generally been so substantially modified by various poorly understood and often nonlinear processes as they propagate out from the Sun that the information they contain about their origins and evolution has been largely washed out. The upcoming NASA Solar Probe Plus and ESA/NASA Solar Orbiter missions will address a number of the outstanding questions, especially those relating to the heating and acceleration of the solar wind itself. However, questions concerning the sources and propagation of SEPs and the 3D structure and evolution of interplanetary coronal mass ejections (ICMEs) require multi-point in-situ measurements made at various radial and longitudinal positions in the inner heliosphere. The proposed Solar Sentinels mission will provide these required observations from four, identically instrumented ecliptic north spinner spacecraft placed in slightly different 0.25 x 0.76 AU near-solar equatorial orbits. Over the course of this mission, the Sentinels spacecraft will realize a number of scientifically advantageous configurations including radial or field aligned relative positions and longitudinal spreads of various extent. Thus, the radial evolution of SEP beams and solar wind transients, along with their 3D geometry will be possible to determine, providing closure to the pressing Heliophysics questions outlined below.

1. SCIENCE OBJECTIVES

When and where are energetic particles accelerated by the Sun?

Large SEP events are generally preceded both by a flare (as evidenced by electromagnetic emissions at various wavelengths) and by the launch of a fast CME and the formation of a shock in the corona and interplanetary space. Determining the relative roles of flares and CME-driven shocks in producing SEPs requires accurate measurements of the relative time profiles of the electromagnetic emissions associated with the flare, of the height-time profile of the CME, and of the timing of the energetic particle injection relative to the flare and CME. However, the information needed for timing studies of particle injections is washed out by the effects of scattering off magnetic irregularities as the particles are transported from the Sun to 1 AU. These scattering effects can be large and highly variable from event to event, with scattering mean free paths ranging from ~0.1 to ~1 AU. The detailed magnetic configuration of the inner heliosphere can also come into play [*Larson et al.*, 1997; *Bieber et al.*, 2002; *Ruffolo et al.*, 2006].

Solar impulsive electron events (see *Lin* [1985] for review), which are observed frequently throughout the solar cycle, are an important source of timing information. *Krucker et al.* [1999] and *Haggerty and Roelof* [2002] found that the inferred injection of near-relativistic (>~30 keV) electrons at the Sun was often delayed by ~10 min with respect to the type III radio bursts at the Sun, and concluded that the escaping near-relativistic populations were not directly related to those generating the prompt flare-related emissions. *Simnett, Roelof, and Haggerty* [2002] reported a correlation between near-relativistic electron events and SOHO CMEs and suggested that a CME-driven shock operating at a few solar radii above the solar surface is the primary accelerator for the near-relativistic electrons. This interpretation was challenged by *Cane* [2003], however, who argued that the near-relativistic particles and those responsible for the electromagnetic emissions belong to a single population and that substantial scattering during the electrons' transport from the Sun is the primary cause of the delay. Finally, recent detailed timing studies using Nançay Radioheliograph data together with EUV and white-light imaging of the corona suggest that the acceleration of near-relativistic electrons is extended in time and results from the restructuring of the corona after the passage of a CME [*Maia and Pick*, 2004; *Klein et al.*, 2005]. In a few solar impulsive electron events where the propagation is nearly scatter-free, *Wang et al.* [2006] found that there are two distinct injections, one at low energies, ~0.4-10 keV, early enough to be the source of

the type III radio bursts; and a second, about ~10 minutes later, of higher-energy (>15-100 keV) electrons.

The timing question can be resolved through simultaneous observations of low energy and near-relativistic electrons by the Sentinels spacecraft positioned at different radial distances from the Sun (e.g., at ~0.7, ~0.5, and ~0.3 AU). If the delay is caused by scattering, the average delay should be systematically smaller the closer to the Sun that the observations are made. On the other hand, if the delay reflects the actual order of things as the SEP-accelerator develops, the average delay will be the same at all radial distances. Also, multi-point measurements are necessary to distinguish between point sources for SEPs such as flare loops vs. broad-front sources such as shocks.

What conditions determine when CME-driven shocks accelerate energetic particles?

The properties of SEPs accelerated at CME-driven shocks are highly variable. For example, even in events in which CME speeds are the same, SEP intensities can range over ~4 orders of magnitude [*Reames*, 2000]. This variability is likely the result of the interplay of many factors, including the composition and distribution of the seed population, the properties of the CME and the shock, and the preconditioning of the inner heliosphere by earlier events. Measurements of SEP elemental and isotopic ratios and ionic charge states suggest that suprathermal ions rather than the bulk solar wind are the primary seed population for particles accelerated by CME-driven shocks. A number of different source populations have been proposed to explain these observations. ACE ^{3}He measurements [*Mason et al.,* 1999] have shown that remnant particles from impulsive events may be an important component of the suprathermal seed population for large, gradual SEP events. But at this time it is unclear if the suprathermals come primarily from the largest flares, whose particles are then dispersed by longitudinal and radial scattering. Or is the suprathermal population continually replenished by numerous small impulsive events, whose intensities cannot be resolved from background at 1 AU? Do multiple flares from the same active region play a key role in filling the inner heliosphere with suprathermals, thereby providing some uniformity to the compositional and spectral characteristics of the flare suprathermals over a wide longitudinal span? Alternatively shock-accelerated SEPs may themselves be another seed population for further acceleration by subsequent CME-driven shocks [*Kahler*, 2001]. Thus, another open question is: what is the relative contribution of "fresh" (as opposed to "remnant") flare suprathermals? Multi-point inner heliospheric in-situ observations at various radial and longitudinal positions, such as what will be provided by Sentinels, are necessary to resolve these questions.

Another challenge to understanding SEP acceleration is the complexity of the energizing shocks and the driving CMEs. CME width is apparently an important shock parameter, as CMEs with widths < 60° rarely produce strong interplanetary shocks [*Gopalswamy et al*., 2001] or gradual SEP events [*Kahler and Reames*, 2003], even when the CME speeds are high (V > 900 km/s). By measuring the CME-associated shock and associated turbulence at several points, a global picture will emerge that will link particle properties to the local variables such as shock strength and geometry (quasi-parallel vs. quasi-perpendicular), magnetic field turbulence and orientation, and seed populations. The four Sentinels spacecraft will provide these crucial observations.

How are energetic particles transported from their acceleration site and distributed in radius, longitude, and time?

The transport of solar energetic particles from their acceleration site through the inner heliosphere to 1 AU and beyond is a problem of critical importance for understanding and eventually predicting SEP events. SEP transport is a complex phenomenon involving a variety of processes: field-line wandering, pitch-angle scattering by turbulent magnetic fluctuations, magnetic focusing by the radially diverging heliospheric magnetic field, adiabatic cooling, and solar wind convection. In the case of shock-accelerated events, the propagation and evolution of the shock must also be taken into account. A particular source of uncertainty in our understanding of SEP propagation is a lack of knowledge of the particle scattering mean free path inside 1 AU. As concurrent Helios and IMP 8 measurements of SEP events in 1980 demonstrated (see Figure 1) multiple inner heliospheric SEP events can be either completely washed out

or be merged by 1 AU [*Wibberenz and Cane*, 2006]. Multi-point measurements at different radial distances will allow the tracking of the evolution of such SEP events.

How do coronal mass ejections evolve in the inner heliosphere?

In addition to their role in SEP acceleration, fast CMEs are the principal drivers of sever geomagnetic disturbances. However, the global topology and evolution in the inner heliosphere of these geo-effective structures is far from established. All current CME models predict that the topology of interplanetary CMEs (ICMEs) is that of a twisted flux rope as a result of the flare reconnection that occurs behind the ejection. This predicted flux rope topology is expected irrespective of the particular mechanism that initiates the ejection or of the initial configuration of the erupting field. Observations at 1 AU, however, find that less than half of all ICMEs, even those associated with strong flares, have a flux rope structure. Many ICMEs at 1 AU appear to have a complex magnetic structure with no clearly defined topology. This conflict between theory and observations is one of the outstanding problems in understanding CMEs; moreover, it is also a critical obstacle to predicting space weather. Only multi-point in-situ measurements provided by Sentinels at various radial distances from the Sun have the potential to resolve this dilemma.

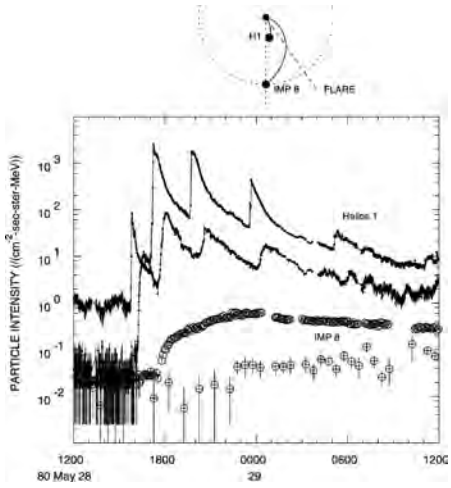


Figure 1. Electron and alpha particle time profiles recorded by Helios-1 at 0.3 AU and by IMP 8 at 1 AU during a series of impulsive particle events on 1980 May 28. Whereas Helios-1 observed multiple injections, no such structures can be resolved in the intensity-time profile at 1 AU. If both spacecraft were observing the same events, then this plot vividly illustrates both the effects of radial and longitudinal transport inside 1 AU and the need for multi-point observations [*Wibberennz and Cane*, 2006].

A related problem is the scarcity of prominence plasma detected in ICMEs. It is widely believed that CMEs are driven by the eruption of highly sheared magnetic field lying at the base of the corona – the so-called filament channel. These channels of sheared field generally (but not always) contain cool plasma observable as a prominence/filament. In fact, prominence material is frequently visible in coronagraph images of CMEs close to the Sun, but prominence plasma is very rarely observed in 1 AU in-situ measurements. One possible explanation is that the prominence material is very narrow and is likely to be missed by a single point observer. Multi-point in-situ observations of ICMEs, at slightly different longitudes, by Sentinels will provide the necessary measurements to determine the validity of this picture.

As noted above, a portion of ICMEs – so-called magnetic clouds (MCs) – retain flux rope like magnetic topologies as they reach 1 AU with their foot points presumably still encored on the Sun. Single spacecraft observations of MCs at 1 AU can be fitted reasonably well with cylindrically symmetric force free magnetic flux rope models [e.g., *Lepping et al*., 1990], though deviations from strict cylindrical symmetry have been noted [e.g., *Mulligan and Russell*, 2001; *Riley and Crooker*, 2004]. However, global MHD models of the evolution of ICMEs from the Sun to 1 AU predict significantly more distorted shapes (such as the "folded pancake" geometry predicted by *Odstrcil et al*. [2004]). The global shape of MCs is clearly the result of the interplay between the internal magnetic pressure of the cloud and the external impact of the ambient solar wind. However, the possible extent of MC deformation is not known. The distortion of cloud morphology during propagation through the inner heliosphere affects the surface geometry of the driven shocks on both the global and local scales, which would have significant impact on the particle acceleration process. Only multi-point in-situ observations at different radial and longitudinal positions of the same MC can provide the necessary measurements to determine MC deformation in the inner heliosphere.

2. SENTINELS OBSERVATIONAL STRATEGY

The primary objective of Solar Sentinels is to observe the inner heliospheric evolution of SEPs and ICMEs. Therefore, it is essential for the mission to cover as much radial distance and longitudinal spread between the Sun and 1 AU as possible. Moreover, all of the selected science objectives require multi-point in-situ observations, thus the individual spacecraft have to be as light as possible in order to maximize their number on a single launch vehicle. This dictates mission designs that do not require costly and massive heat shields. The Johns Hopkins University Applied Physics Laboratory (APL) has extensively studied this problem leading up to the 2006 *LWS Sentinels STDT Report*. It was concluded that for spinning spacecraft, with their spin axis perpendicular to the ecliptic, it is possible to get as close to the Sun as 0.25 AU while maintaining acceptable operational temperatures in the core of the spacecraft. Detailed trajectory computations also revealed that a 0.25 x 0.76 AU orbit is achievable with four Venus gravity assist maneuvers, (the 0.76 AU aphelion is dictated by the orbital distance of Venus). Furthermore, slightly varying the Venus encounter altitudes, the different spacecraft can be placed in slightly different orbits that result in desirable spacecraft configurations.

The minimum number of spacecraft required for adequate characterizations of either SEP events or transient solar wind structures such as ICMEs follows from the number of free parameters of the simplest phenomenological models. For SEP longitudinal intensities we assume that at least a Gaussian profile fit, with three free parameters, is needed. In addition, a fourth parameter, the radial dependence of the SEP intensities, expressed as a simple power law exponent, is also required. For ICMEs we assume a magnetic flux rope initially joined to the Sun at both ends. In that case we need measurements at two positions to give the ICME longitudinal extent, or equivalently, the location and radius of the major axis. An additional measurement is required to determine the latitudinal extent, which corresponds to the flux rope radius of the minor axis. The flux rope plane is generally tilted at some angle to the ecliptic, and the cross-section is distorted by a combination of solar wind expansion and solar wind stream interactions, so that at least a fourth spacecraft measurement is needed to determine the cross-sectional distortion. While the time series of magnetic field and plasma measurements at each of the Sentinels spacecraft yields information well beyond simple snap shot observations, the ICMEs are also subject to spatially dependent accelerations, shearing, and magnetic reconnections, all of which will require four spacecraft for the most basic characterizations.

The measurement requirements of the above stated science objectives are in-situ fields and particles observations at multiple inner heliospheric locations. The SEP objectives require the measurement of the solar energetic particles from their suprathermal seed populations to well above 100 Mev/nuc to cover the most dangerous particles to humans and satellites in space. In order to connect these particles to their sources, the local magnetic field, along with remote radio waves, X-ray and γ-ray measurements will have to be obtained. The ICME objectives will also require the measurement of the solar wind thermal particles with their detailed composition and full electron distribution functions to establish the internal ICME magnetic topologies. The detailed measurement requirements with the necessary accuracies are summarized in Table 1. It is also essential that the four Sentinels spacecraft be identically instrumented making intercalibration and comparison between various inner heliospheric locations possible.

3. MISSION DESIGN AND COST

The four spacecraft Solar Sentinels mission has undergone a detailed engineering study at APL during 2004-2006. The spacecraft are spin-stabilized, with the spin axis perpendicular to the ecliptic and a rotation rate of 20 rpm. This spin rate reduces the effective solar constant at the spacecraft from 16 to 5 Suns at perihelion. Optical solar reflector material and thermal louvers maintain a core spacecraft temperature of 0° to 25°C, while various passive techniques provide thermal control of exposed subsystems and instruments. Power is supplied by 16 solar panels mounted around the bottom of the octagonal spacecraft body and tilted at 45° to maintain the array operating temperature below the 180°C design limit (see Figure 2). The X-band telecommunications subsystem uses a gimbaled phased-array

Table 1. Solar Sentinels measurement specifications

Instrument	Parameter(s) or Quantity(ies) Measured	Sensitivity *Dynamic Range*	Energy Range *Resolution*	Angular Range *Resolution*	Time Integration *Cadence*
High Energy Ion Composition Analyzer	Composition and energy spectra of energetic nuclei $1 \le Z \le 28$ Energetic electrons	~0.1-10^6 protons/cm^2 s sr for >10 MeV	ions: ~3 - 500 MeV/nuc e^-: ~0.5 - 10 MeV *6 energy bins/decade*	As much of 4-π ster as possible *≤45°*	30 s for H; 10 s for e^-; 60 s for Z≥2
Low Energy Ion Composition Analyzer	Composition and energy spectra of energetic nuclei $1 \le Z \le 28$	~0.1-10^6 protons/cm^2 s sr for >0.1-10 MeV	ions: ~0.02 - 3 MeV/nuc *6 energy bins/decade*	As much of 4-π ster as possible *≤45°*	30 s for H; 10 s; 60 s for Z≥2
Energetic Electron and Proton Instrument	Energetic ions and electrons	~5 x 10^8 particles/cm^2 s sr MeV (protons and e^-)	Ions: ~0.02 - 10 MeV e^-: ~0.02 - 1 MeV *6 energy bins/decade*	As much of 4-π ster as possible *ions/e^-: 30°*	Ions/e^-: 10 s (e^- < 1 s for transients)
Suprathermal Electron Instrument	Distribution functions of suprathermal electrons	~ 1-10^8 e^-/cm^2 s sr for >2 keV	Suprathermals: ~2-100 keV *ΔE/E ~ 0.20*	As much of 4π sr as possible *Suprathermals: 20°*	Suprathermals: 10 s (<1 s for transients)
Solar Energetic Particle Charge State and Composition Analyzer	Ionization state of suprathermal through energetic particles $1 \le Z \le 28$	~1-10^5 ions/cm^2 s sr for >0.1 MeV *10^4*	10 keV/nuc – 1 MeV/nuc *6 energy bins/decade*	As much of 4π sr as possible	15 min
Solar Wind Ion Analyzer	3D distribution function of solar wind thermal protons and alphas	Density ~0.1 - 10^3/cm^3	~150 - 1500 (2000) km/s ~0.1 - 12 keV/e (protons) ~0.2 - 20 keV/e (alphas) (E/Q)/(ΔE/Q) ~ 20	±30° of ecliptic 360° in ecliptic (by s/c spin) 5° x 5 ° (or better) resolution	1 min (few sec in burst mode for transients)
Solar Wind Composition Analyzer	Composition and velocity distribution functions of major solar wind heavies including He, C, O, and Fe	10^7 over entire mission 10^5 at given location	100 eV/q - 100 keV/q *ΔE/E ~ 0.05*	±60° from ecliptic *10°*	1-5 minutes (burst mode to 10 s)
Solar Wind Electon Analyzer	Full electron distribution functions	Density to >1000 cm^{-3}	<1 - 3000 eV *ΔE/E ~ 0.10*	0° to 180° from spin axis ~3° x ~20°	~20 s (<3 s in burst mode for transients)
Dual Magnetometer	DC vector magnetic field	0 – 20 Hz *±64,000 nT*	n/a *0.1 nT accuracy for < 1000nT*	n/a	0.3 s (0.05 s in burst mode)
Search Coil Magnetometer	AC magnetic field	5 Hz – 20 kHz	< 10^{-3} nT/√Hz at 5 Hz, 10^{-5} nT/√Hz at 5 kHz	n/a	*> 20 kSPS in burst*
Radio and Plasma Waves	Radio and electromagnetic fields	~DC – 16 MHz	< 1 mV/m at ~DC ~10^{-19} W/m^2/Hz at 16 MHz	n/a	*0.5 s to 10^7 s (burst)* *4 s for* in-situ *spectra* *20 s for radio spectra*
Neutron Spectrometer	Time*-tagged neutron spectra (*at Sun)	0.001 neutrons/cm^2 s *10,000 (may be limited by x-ray dose)*	2 - 20 MeV *30% at 3 MeV*	Full Sun	Max 30 s in burst mode *Dynamic (burst mode and background mode)*
X-Ray Imager	Hard and soft X-ray spectra and imaging	Up to GOES X30 flares	1 - 150 keV *ΔE/E ~ 0.05*	150° in ecliptic; ±10° out of plane *~20 arcsec*	~20 s (<3 s in burst mode for transients)
Gamma Ray Spectrometer	Gamma spectra	5 x 10^{-3} gammas/cm^2 s at 2.2 MeV *10,000*	0.2 - 20 MeV *7% at 0.662 MeV*	Full Sun n/a	1 s in burst mode *Dynamic (burst mode and background mode)*

high-gain antenna for high-rate downlink and a medium-gain antenna for uplink and low-rate downlink. The two antennas are housed, together with a low-gain antenna, in a thermal-protective radome and mounted on a despun platform located on the top deck of the spacecraft. Data will be stored in the solid state recorders of the redundant command and data handling units and downlinked at a rate that ranges from 750 kbps to 23 kbps depending on the spacecraft-Earth range. In addition to science and housekeeping data, space weather data will be continuously downlinked from each spacecraft at a significantly lower telemetry rate. The complete cost of the mission, including an Atlas V 541 launch vehicle and 6 years of Phase E activities, was estimated by APL to be $750M in FY07 dollars. Since that time, the cost of the launch vehicle has increased significantly, and including inflation, the current cost estimate of this mission is closer to $1B.

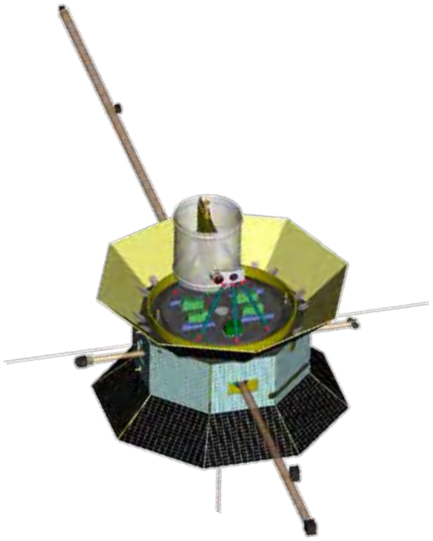


Figure 2. The Solar Sentinels spacecraft design will accommodate the payload outlined in Table 1 and meet the unique thermal and power challenges presented by the mission environment.

4. EVALUATION CRITERIA

Identification as a high priority or requirement in previous studies or roadmaps. The Solar Sentinels mission described here is explicitly identified by the 2003 Decadal Survey ("*The Sun to Earth – and Beyond: A Decadal Research Strategy in Solar and Space Physics*"), under the name Multispacecraft Heliospheric Mission, and by the 2009 Heliophysics Roadmap, under the designation "Solar Energetic Particle Acceleration and Transport, 2021", as high priority NASA Heliophysics objectives for the 2020-2025 time frame.

Makes a significant contribution to more than one of the panel themes. This mission contributes to multiple important scientific questions facing solar and heliospheric physics today. It will also enable the development and operational testing of multi-day space weather forecasting capabilities, contributing to the Research to Operations theme. Moreover, Sentinels observations will be essential for the continued development of global heliospheric and particle transport models.

Contributes to important scientific questions facing solar and space physics today. The Solar Sentinels science objectives focusing on the inner heliospheric evolution of ICMEs, and the transport of energetic particles are among the most important scientific questions facing solar and heliospheric physics today.

Contributes to applications and/or policy making. Solar Sentinels will enable the development and operational testing of multi-day space weather forecasting for Earth and for any other location in the heliosphere (including possible human missions to Mars or to an asteroid).

Complements other observational systems or programs available. The Solar Sentinels mission is highly complementary to the currently developed NASA Solar Probe Plus and ESA/NASA Solar Orbiter missions. While Solar Probe Plus will focus on collecting measurements significantly closer to the Sun and Solar Orbiter will add solar remote sensing observations, an overlapping Solar Sentinels mission will provide the necessary multi-point in-situ data from the inner heliosphere to answer the questions of ICME evolution and SEP acceleration and transport. Thus, while the Solar Sentinels orbit design is similar to that of Solar Probe Plus and Solar Orbiter (using Venus gravity assist maneuvers to access the inner heliosphere), there are significant differences in observational strategy between these missions. Solar Probe Plus and Solar Orbiter are essentially encounter missions concentrating their downloaded telemetry on the time periods of closest approach to the Sun below 0.3 AU. In contrast, Solar Sentinels will bring back continuous inner heliospheric data from the distance range of 0.25-0.76 AU and will add multi-point measurements. Therefore, Solar Sentinels is an essential mission to connect the exciting Solar Probe Plus and Solar Orbiter observations to 1 AU measurements.

Is affordable. The APL engineering team that developed the Solar Sentinels mission concept also provided detailed costing estimates. Thus, it is with confidence that Solar Sentinels can be designated as a Mid Size mission costing below $1B.

Technical readiness. The Solar Sentinels mission has been extensively studied by an APL engineering team and it relies on currently available technologies only. No technology development program is necessary to initiate this mission.

Fits with other national and international plans and activities. The Solar Sentinels mission also continues to naturally fit with the NASA Living With a Star (LWS) and the International Living With a Star (ILWS) programs.

REFERENCES

Bieber, J. W., et al., Energetic particle observations during the 2000 July 14 solar event, *Astrophys. J.*, **567**, 622, 2002.

Cane, H. V., Near-relativistic solar electrons and type III radio bursts, *Astrophys. J.*, **598**, 1403, 2003.

Gopalswamy, N., et al., Characteristics of coronal mass ejections associated with long-wavelength tyoe II radio bursts, *J. Geophys. Res*., **106**, 29,219, 2001.

Haggerty, D. K. and E. C. Roelof, Impulsive near-relativistic solar electron events: Delayed injection with respect to solar electromagnetic emission, *Astrophys. J.*, **579**, 841, 2002.

Kahler, S. W., The correlation between solar energetic particle peak intensities and the speeds of coronal mass ejections: Effects of ambient particle intensities and energy spectra, *J. Geophys. Res*., **106**, 20,947, 2001.

Kahler, S. W. and D. V. Reames, Solar energetic particle production by coronal mass ejection-driven shocks in solar fast-wind regions, *Astrophys. J.*, **584**, 1063, 2003.

Klein, K.-L., et al., Coronal phenomena at the release of solar energetic electron events, *Astron. Astrophys.*, **432**, 1047, 2005.

Krucker, S., et al., On the origin of impulsive electron events observed at 1 AU, *Astrophys. J.*, **519**, 864, 1999.

Larson, D. E., et al., Using energetic electrons to probe the topology of the October 18-20, 1995 magnetic cloud, *Adv. Space Res*., **20**, 655, 1997.

Lepping, R. P., J. A. Jones and L. F. Burlaga, Magnetic field structure of interplanetary magnetic clouds at 1 AU, *J. Geophys. Res*., **95**, 11,957, 1990.

Lin, R. P., Energetic solar electrons in the interplanetary medium, *Solar Phys*., **100**, 537, 1985.

Maia, D. J. F. and M. Pick, Revisiting the origin of impulsive electron events: Coronal magnetic restructuring, *Astrophys. J.*, **609**, 1082, 2004.

Mason, G. M, J. E. Mazur and J. R. Dwyer, ^{3}He enhancements in large solar energetic particle events, *Astrophys. J.*, **525**, L133, 1999.

Mulligan, T. and C. T. Russell, Multispacecraft modeling of the flux rope structure of interplanetary coronal mass ejections: Cylindrically symmetric versus nonsymmetric topologies, *J. Geophys. Res*., **106**, 10,581, 2001.

Odstrcil, D., P. C. Riley and X. P. Zhao, Numerical simulation of the 12 May 1997 interplanetary CME event, *J. Geophys. Res*., **109**, doi:10.1029/2003JA010135, 2004.

Reames, D. V., Particle acceleration by CME-driven shock waves, in *Invited, Rapporteur, and Highlight Papers of the 26th ICRC (Salt Lake City)*, eds. B. L. Dingus, D. B. Kieda and M. H. Salamon, AIP Conf. Proc. 516, p. 289, 2000.

Riley, P. and N. U. Crooker, Kinematic treatment of coronal mass ejection evolution in the solar wind, *Astrophys. J.*, **600**, 1035, 2004.

Ruffolo, D., et al., Relativistic solar protons on 1989 October 22: Injection and transport along both legs of a closed magnetic loop, *Astrophys. J.*, **639**, 1186, 2006.

Simnett, G. M., E. C. Roelof and D. K. Haggerty, The acceleration and release of near-relativistic electrons by coronal mass ejections, *Astrophys. J.*, **579**, L854, 2002.

Wang, Y. M., M. Pick and G. M. Mason, Coronal holes, jets, and the origin of ^{3}He-rich particle events, *Astrophys. J.*, **639**, 495, 2006.
Wibberenz, G. and H. V. Cane, Multi-spacecraft observations of solar flare particles in the inner heliosphere, *Astrophys. J.*, **650**, 1199, 2006.

COSMO

The COronal Solar Magnetism Observatory

A Response to the Request for Information from the
NRC Decadal Survey in Solar and Space Physics

Steven Tomczyk[1*], Joan Burkepile[1], Roberto Casini[1], Giuliana de Toma[1], Yuhong Fan[1], Philip G. Judge[1], Hao-sheng Lin[2], Boon Chye Low[1], Scott W. McIntosh[1], Peter G. Nelson[1], Stanley C. Solomon[1]

[1]The National Center for Atmospheric Research, High Altitude Observatory
[2]The University of Hawaii, Institute for Astronomy

[*]Corresponding author, tomczyk@ucar.edu, 303-497-1579
Project website: http://www.cosmo.ucar.edu/

November 12, 2010

Introduction

Measuring magnetic fields in the solar corona is crucial to understanding and predicting the Sun's generation of space weather that affects communications, GPS systems, space flight, and power transmission (Hanslmeier, 2003; Lambour, et al., 2003; Iucci et al., 2006). Most forms of solar activity, including high-energy electromagnetic radiation, solar energetic particles, flares, and coronal mass ejections (CMEs), derive their energy from coronal magnetic fields. The corona is also the source of the solar wind with its embedded magnetic field that engulfs the Earth. The ability to measure coronal magnetic fields will significantly advance the understanding of the underlying physical processes, and will lead to improved predictions of hazardous space weather effects on Earth.

The scientific community recognizes the critical need for accurate measurements of the coronal magnetic field. The last National Research Council decadal review of solar and space physics recognized that an important next step in solar physics is "...to develop the science of coronal mapping for measurement of the structure and strength of the coronal magnetic field." (Lanzerotti, et al., 2003).

Precise coronal magnetic field observations are essential to address some of the important questions that remain in solar physics. Yet until recently, magnetic fields in the corona have been extremely difficult to measure for three important reasons: 1) the magnetic fields in the corona are intrinsically weak; 2) the coronal spectroscopic lines are dimmer and broader than their photospheric counterparts; and 3) the optically thin corona requires interpretation of magnetic signatures integrated along extended path lengths. These challenges have been addressed recently using greatly improved infrared (IR) detectors, forward models, advanced diagnostic tools, and inversion codes. As a result, we now have the capability to directly measure the coronal magnetic field. Unfortunately, today's solar coronagraphs are too small to obtain the measurements at the required spatial and temporal resolution.



***Figure 1**. The solar corona during the total eclipse of 11 July 1991.*

The COronal Solar Magnetism Observatory (COSMO) is a proposed suite of complementary ground-based instruments designed to study magnetic fields and plasma conditions in the corona. The central instrument is a 1.5m-aperture coronagraph that will obtain daily measurements of the strength and direction of coronal magnetic fields over a large field-of-view at the spatial and temporal resolutions required to address the outstanding problems in coronal physics. Supporting instruments focus on chromospheric and prominence magnetometry, and observing the electron scattered K-corona. This suite of comprehensive observations will not be provided by any current or proposed observatory; the COSMO facility will enable breakthrough science and enhance the value of data collected by other observatories on the ground (e.g. ATST, FASR, SOLIS) and in space (e.g. Hinode, STEREO, SDO, TRACE, GOES, and SOHO).

COSMO will replace the current Mauna Loa Solar Observatory and be operated as an NCAR facility in collaboration with The University of Hawaii and The University of Michigan.

Key Science Drivers

The proposed COSMO design emerged from a broad base within the solar and heliospheric communities through the COSMO Science Advisory Panel which identified important questions that this facility can uniquely address. COSMO's unprecedented observations will contribute significantly to answering the following questions about fundamental coronal processes.

1. What determines the magnetic structure of the corona?

COSMO will investigate large-scale coronal structures in relation to surface field advection, differential rotation, and photospheric flux emergence. COSMO will provide new information on the evolution and interactions between magnetically closed and open regions that determine the changing structure of the heliospheric magnetic field.

2. What is the physics of the storage and explosive release of magnetic energy in the solar atmosphere that produce CMEs?

COSMO will detect the polarization signatures of non-potential fields to gain insight into the roles of energy storage, magnetic reconnection, helicity creation and transport, and flux emergence in CME formation. COSMO will also provide white light images of CMEs in the low corona that are needed to determine the basic properties of CMEs in the early stages of formation.

3. What is the nature of the changes in coronal magnetic structure that accompany the 11-year solar cycle?

COSMO will provide daily observations of the coronal magnetic field down to 1 Gauss over a 1° field-of-view and will obtain observations approximately 300 days per year. Long-term synoptic observations of the coronal magnetic field will directly address how the coronal field responds to the sunspot cycle and how the switch in polarity of the global field manifests itself in the heliosphere.

4. How are prominences formed, how do they evolve, and how are they related to CMEs?

COSMO will provide routine observations of prominence, filament, and chromospheric magnetic fields, and prominence flows. These observations will place constraints on prominence densities and determine how prominence and coronal magnetic fields interact, how and where magnetic energy is stored (e.g. flux and helicity transport) and how it is released (e.g. instabilities, reconnection, dissipative heating).

5. Where do CME-associated shocks form, and what is their role in accelerating solar energetic particles?

Particles accelerated by CME-driven shocks have the highest particle energies and pose the greatest space weather hazards. COSMO can detect compressions and distortions in the field due to the formation and passage of a CME and associated waves. Density compressions resulting from shock formation can be large enough to produce intensity enhancements in white light coronal images (Vourlidas et al. 2003). COSMO has the potential to simultaneously detect both the shock and CME locations.

6. What is the role of magnetic reconnection and flares in CME formation?

COSMO will provide the first routine, high time cadence measurements of coronal magnetic fields in and around flares and associated CMEs as seen over the solar limb. These high time cadence observations can help determine when, where and how magnetic energy is released by the CME and by the dissipative processes that result in flares.

Observational Requirements

These science questions have been translated into observation and telescope requirements by the COSMO Science Advisory Panel and the Co-Principal Investigators team. The COSMO science questions will be addressed through synoptic observation of coronal magnetic fields over a 1° field-of-view, with a spatial resolution of 5 arcseconds at a temporal cadence of 10 minutes. COSMO will be sensitive to magnetic field strengths of a few Gauss for the faintest coronal structures and less than a Gauss for the brightest. In order to do this, COSMO will need to observe features down to a few millionths of the brightness of the solar disk, which requires a large aperture telescope, 1.5m in diameter.

The large field-of-view is needed to observe the global properties of the corona and is required to address all of the science questions, especially to track the outward motion and other basic properties of CMEs and prominences as they are ejected from the corona. It is also necessary to observe as low in the corona as possible to understand how the corona continually reacts to changes occurring in the photosphere and chromosphere. High temporal cadence is needed to capture the relentless dynamic evolution of the coronal plasma structure and explosive disturbances (e.g., MHD waves, shocks, prominence eruptions, CMEs, and flares) that are the essence of science questions 2, 4, 5, and 6. The physical properties of CMEs and eruptive prominences are best determined from the polarization signal of broadband filtered white-light observations because the scattered light from the corona is partially polarized. The Science Advisory Panel and Co-PIs have determined that a spatial resolution of 5 arc seconds is sufficient to address many of the scientific questions pertaining to the global magnetic structure of the corona.

The COSMO Science Advisory Panel has also identified the specific need to directly measure the magnetic field in the chromosphere and in prominences (science question 4). Recent observations suggest that MHD waves in the upper chromosphere have sufficient energy to accelerate the solar wind outside of active regions (De Pontieu et al. 2007; Aschwanden, et al. 2007); COSMO will provide routine magnetic field measurements in this key region. These measurements will provide critical information on plasma conditions in the very low atmosphere that are needed to couple the coronal magnetic fields with those measured at photospheric heights.

Technical Overview

The science goals and observational requirements outlined in the previous section will be met with a facility consisting of two major elements: a large-aperture coronagraph with its own pointing system and dome, and a solar-pointed spar for supporting instrumentation in a smaller adjacent dome. A rendering of the large-aperture coronagraph and dome is shown in Figure 2.

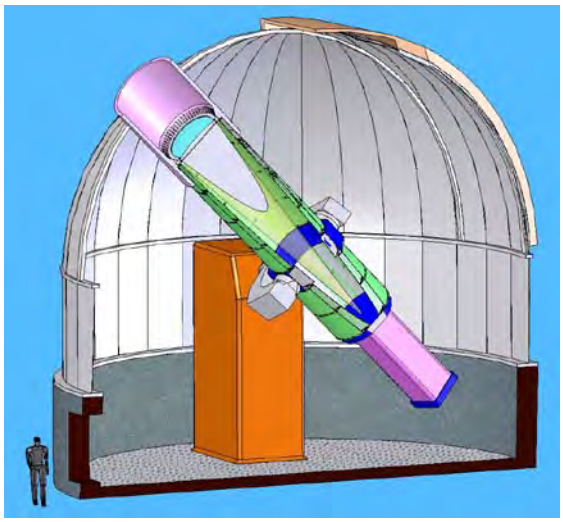


Figure 2. *Concept for the large aperture coronagraph and dome. A protective "snout" extending through the telescope slot when the dome is open allows the lens to be continually washed in HEPA-cleaned air. The sealed He-filled tube protects the back surface of the lens from contamination and improves internal seeing.*

The large-aperture coronagraph will have instrumentation to measure coronal magnetic fields. In order to exploit these measurements

to their full potential for advancing science, crucial information on prominence and chromospheric magnetic fields and the dynamical conditions of the embedded plasma will be collected by two additional instruments: a chromosphere and prominence magnetometer and a white-light K-coronagraph. These supporting instruments are designed to provide temporal and spatial resolutions that meet or exceed the temporal and spatial resolutions of the coronal magnetic field observations from the large-aperture coronagraph.

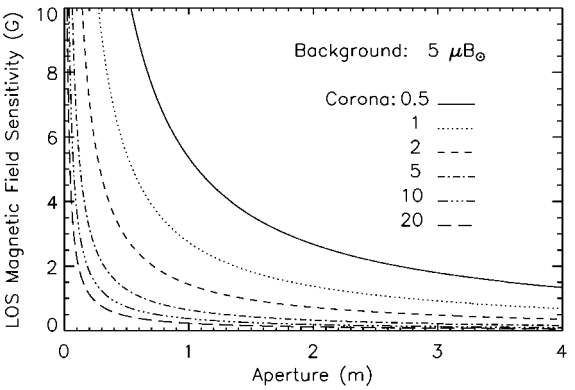


Figure 3. *Photon noise limited LOS magnetic field strength sensitivity using the FeXIII 1074.7 nm line vs. telescope aperture for an integration time of 10 minutes assuming a background sky brightness of 5 μB_{sun}. The lines plotted correspond to various intensities of the corona in units of millionths of the disk intensity. It is possible to achieve magnetic field sensitivity of 1 Gauss in 10 minutes for coronal structures with brightness of 2 μB_{sun} in the Fe XIII 1074.7 nm line with a 1.5-m aperture telescope. Structures as faint as 0.5 μBsun can be observed with a sensitivity of 3.5 Gauss in 10 minutes, or 2 Gauss in 30 minutes.*

Large-aperture coronagraph

The most important instrument in the COSMO suite is a large-aperture coronagraph with post-focus instrumentation to measure coronal magnetic fields. This solar telescope performs coronal magnetometry using the emission lines of Fe XIII at 1074.7 and 1079.8 nm and the He I chromospheric emission line at 1083.0 nm. An overview (Judge, et al. 2001) of methods for measuring coronal magnetic fields identified the most promising techniques. COSMO will use the Zeeman effect and Hanle effect in scattering polarization to study the widest possible range of magnetic fields in the corona and chromosphere. The line-of-sight (LOS) strength of the coronal magnetic field can be measured directly through the Zeeman effect observed in the circular polarization, and the plane-of-sky (POS) direction of the magnetic field is determined from the linear polarization of the scattered radiation. These magnetic field measurements will have a dynamic range from a fraction of a Gauss to several thousand Gauss. This is critical to provide information on both large-scale "quiet" coronal fields and active region fields.

To meet the science goals discussed above requires observation of the coronal magnetic field:

- with 1 Gauss sensitivity,
- and 5 arcseconds spatial resolution,
- for corona with 2 μB_{sun} brightness,
- in 10 minutes integration time,
- for a sky background level of 5 μB_{sun}.

A calculation of the available coronal flux (Figure 3) indicates that this demanding requirement can only be met with a coronagraph having an aperture of at least 1.5 m with an extremely low level of instrumentally scattered light and a 1º field-of-view. The aperture requirement is driven by the need to collect sufficient photons to achieve a photon noise limited magnetic field measurement at the required level.

Since the solar corona is a million times fainter than the nearby solar photosphere, observing it requires a coronagraph with extremely low levels of scattered light. A detailed engineering study of scattering from lenses and mirrors was conducted (Nelson 2006a; 2008) and showed that a lens objective scatters a factor of 10 times less light than a mirror with the same quality of figure, and a factor of four times less than a mirror with the same level of dust contamination.

The huge scattering advantage of a lens over a mirror led us to choose a refracting design for the COSMO large-aperture coro-

nagraph. To evaluate the feasibility of constructing a 1.5-m lens, we conducted a finite element analysis (Nelson, 2006b; 2008) which showed that gravitational flexure of this lens would have a negligible impact on the image quality. The FEA analysis also considered stress-induced birefringence, and this was found to be well below acceptable levels.

We note that blanks of fused silica of sufficient size and quality for the COSMO large aperture coronagraph are now readily available, due in large part to developments in support of optical fibers for telecommunications. A fused silica lens with a diameter of 1.6 m will be used as a camera lens in the Large Synoptic Survey Telescope (LSST; Oliver et al. 2008).

Post-focus instrumentation for the coronagraph includes a *narrow band tunable filter* to observe the entire field-of-view, and a *fiber-fed spectropolarimeter* capable of high spectral resolution line profile observations. Both of these instruments are based on prototypes developed over the past decade (see below).

Chromosphere and prominence magnetometer (ChroMag)

The second instrument in the suite is the ChroMag, an imaging polarimeter for magnetic and plasma diagnostics of chromosphere and prominences, using the spectral lines of He I (587.6 and 1083 nm) and Hα (656.3 nm), both on the solar disk and above the limb. This instrument will have the ability to observe magnetic fields and LOS Doppler shifts (Casini 2007). Measuring prominence magnetic fields embedded within the corona provides a more complete picture of magnetic field configurations that are needed to understand the role of prominences in CME formation and initiation. The ChroMag design (Elmore 2007b) will utilize a 20-cm aperture telescope and a tunable filter to measure the polarization across the line profile with a bandwidth of 0.025 nm in the visible region and 0.046 nm in the IR. It will have a full FOV of 2.5 solar radii and a spatial resolution of 2.3 arcseconds.

White-light K-coronagraph

The third element of the suite is a 20-cm aperture K-coronagraph that will measure white light coronal polarization brightness. This will provide a direct measurement of the column density of coronal electrons. It will reside on the solar-pointed spar in a small adjacent dome and provide observations of CME formation and early acceleration in the very low corona (down to 1.05 R_{sun}). It replaces the 1970s-vintage Mauna Loa Mk4 system, the world's only ground-based white-light coronameter. Modern detector technology will be employed to significantly improve spatial resolution, temporal resolution, and signal-to-noise out to 2.5 R_{sun} (Elmore 2007a). Given the absence of a white-light coronagraph on the SDO spacecraft, the K-coronagraph will be an essential complement to space- and ground-based coronal instruments.

Technical Readiness

Prototype instruments are producing results that, although limited by the small aperture of their telescopes, illustrate the potential and feasibility of the COSMO post-focus instruments.

One prototype resides on *Solar-C*, the University of Hawaii 46-cm aperture coronagraph at Haleakala (Kuhn, et al. 2003). It employs IR detector technology in an Optical Fiberbundle Imaging Spectropolarimeter (OFIS) to record the complete polarization (Stokes *I,Q,U,V*) of the forbidden line of Fe XIII at 1074.7 nm. Figure 4 shows a measurement of the LOS strength of the magnetic field (Lin, et al. 2004). Each fiber in the 16 x 8 fiber bundle subtended 20 arc seconds, and the observation required 70 minutes of integration time. Notably, the errors on the LOS component of the magnetic field are less than

1 Gauss. The precision on the magnetic field achieved with this instrument demonstrates the feasibility of the technique; however, the limited FOV and the coarse spatial and temporal samplings are not sufficient to address the science questions posed above.

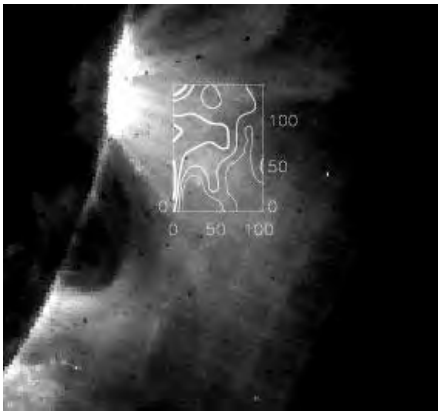


Figure 4. *Coronal magnetic field observations from Solar-C and OFIS shown as a contour plot of the LOS magnetic field strength determined from circular polarization, superposed on EIT observations for context.*

Another prototype application of coronal polarimetry with new IR detector technology is illustrated in Figure 5. This shows a map of the coronal field direction and strength measured using the NCAR/HAO Coronal Multichannel Polarimeter (CoMP) on the 20-cm aperture OneShot coronagraph at NSO's Sacramento Peak Observatory in New Mexico (Tomczyk et al. 2008). The CoMP instrument can observe the coronal magnetic field direction and strength, and plasma motion using the near infrared FeXIII emission lines with a full view of the low corona (~1.03 to 1.4 R_{sun}). The CoMP observations are photon noise limited, required an integration time of 2.42 hours, and were made at a spatial resolution of 9 arc seconds.

Unique high cadence observations made with CoMP in October 2005 have produced the detection of ubiquitously propagating Alfvén waves in the solar corona (Tomczyk et al. 2007). These waves are a possible source of mechanical energy needed to heat the quiet corona and act as a driver for the fast solar wind (DePontieu et al. 2007). Measurement of the speed of wave propagation can also be used to infer the strength and orientation of the magnetic field through the recently developed technique of coronal seismology (e.g. Tomczyk and McIntosh 2009). This provides an important additional constraint on the magnetic field complementary to the polarization measurements.

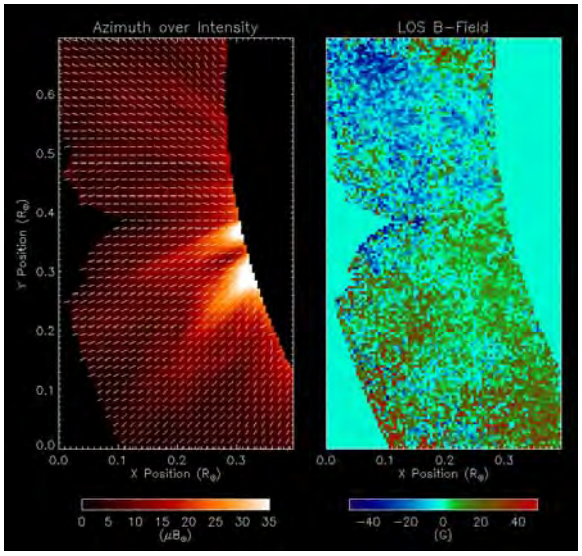


Figure 5. *Coronal magnetic field observations from CoMP. Left: the direction of the magnetic field shown as vectors on top of the FeXIII 1074.7 intensity. Right: the LOS component of the coronal magnetic field.*

The field strengths measured by OFIS and CoMP are in the range of 1-50 Gauss. These instrument developments demonstrate that technological advances, primarily in IR detector technology, now make it possible to measure the weak fields in the solar corona. However, these prototype results are severely limited by the modest apertures of the available coronagraphs. The science goals of COSMO require a coronagraph with an aperture of 1.5 m.

COSMO as a Complement to ATST

The Advanced Technology Solar Telescope (ATST) is a Major Research Equipment project to build a 4-m aperture solar telescope which has the ability to measure coronal magnetic fields above the limb using the same Zeeman and Hanle diagnostics as COSMO. Due to its large aperture, the ATST

is optimized for high spatial resolution observations over a maximum 5 arcminute field-of-view (Figure 6). ATST will be used to pursue a range of user defined observations on the solar disk and above the limb, but with a field-of-view that is small compared to large scale coronal structures. The capabilities of ATST are complementary to the large field-of-view, dedicated measurements provided by COSMO.

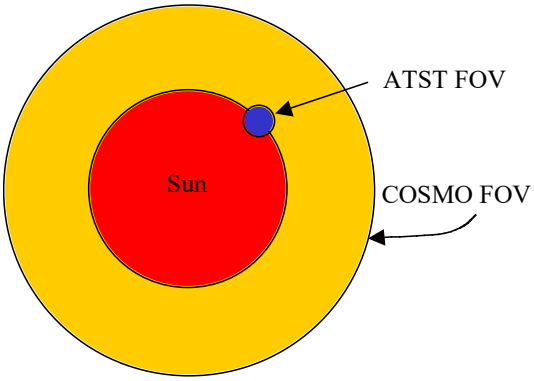


Figure 6. *Comparison of the FOV of the COSMO large coronagraph (orange, 1.0°) with the ATST FOV (blue, 0.08°). The size of the Sun (red, 0.5°) is shown for comparison.*

The light gathering ability of a telescope is given by the product of the area of the collecting aperture and the solid angle of the field-of-view. While ATST has a factor of 7 greater aperture area, that advantage is more than compensated by the factor of 144 larger solid angle of COSMO. *The COSMO coronagraph will have a light gathering power that exceeds that of the ATST by a factor of 20.* This, coupled with its dedicated nature, makes COSMO ideally suited for observation of the corona over a range of spatial and temporal scales that will not be available to the ATST. Such observations are crucial to advance understanding of the structure and dynamics of the solar corona.

COSMO COST ESTIMATE

We have completed a conceptual design of the COSMO facility and suite of instruments. From this, we estimate that the COSMO facility can be designed and constructed in five years at an approximate cost of $20M. This includes telescopes, instruments and buildings. Note that the design and construction of the COSMO K-coronagraph has been fully funded by NSF through NCAR at a cost of $2.2M and the design is currently underway. We expect to deploy the K-coronagraph to MLSO early in 2013. Following the completion of the COSMO facility, we will transfer the K-coronagraph to the COSMO spar.

The recurring operational cost of the COSMO facility will be borne mostly by the existing HAO operational effort. This currently consists of 3.0 FTE observing personnel at MLSO and 3.0 FTE data management personnel at HAO in Boulder. Once COSMO is complete, the existing MLSO facility will be closed and the HAO effort transferred to the new observatory, thus minimizing recurring operations costs. There will be an increase in data processing and reduction costs due to the greater volume and complexity of the COSMO data. Therefore, the estimated annual incremental operational cost will be: *1.0 FTE data management personnel, or approximately $200K/year.*

This level of data management support will allow COSMO to provide all key observations to the user community as a variety of data products: quick-look real-time images, movies, and activity alerts; level 1 data products, activity catalogues, and databases; and level 2 fully processed data products designed for easy use with other datasets and community models for data mining and reanalysis. COSMO will have an open data policy.

The Mauna Loa Solar Observatory has been serving the solar and heliospheric community for over 40 years. Based on this, we expect the COSMO facility will provide unique and valuable observations to the community over its expected 40 year lifetime.

REFERENCES

Aschwanden, M. J., Winebarger, A., Tsiklauri, D., Peter, H. 2007, The Coronal Heating Paradox, *ApJ*, **659**, 1673.

Casini, R. 2007, Prominence and Filament Magnetometry Simulations, *COSMO Tech Note 12*, Boulder, Colorado, High Altitude Observatory.

DePontieu, B., McIntosh, S.W., Carlsson, M., Hansteen, V.H., Tarbell, T.D., Schrijver, C.J., Title, A.M., Shine, R.A., Tsuneta, S., Katsukawa, Y., Ichimoto, K., Suematsu, Y., Shimizu, T., Nagata, S. 2007, Chromospheric Alfvenic Waves Strong Enough to Power the Solar Wind, *Science*, **318**, 1574.

Elmore, D. 2007a, Baseline Design of a Coronagraph to Measure K-corona Polarization Brightness, *COSMO Tech Note 8*, Boulder, Colorado, High Altitude Observatory.

Elmore, D. 2007b, Baseline Design for a Prominence and Filament Magnetometer, *COSMO Tech Note 11*, Boulder, Colorado, High Altitude Observatory.

Hanslmeier, A. 2003, Space Weather - Effects on radiation on manned space missions. *Hvar Observatory Bulletin*, **27**, 159.

Iucci, N., Dorman, L. I., Levitin, A. E., Belov, A. V., Eroshenko, E. A., Ptitsyna, N.G., Villoresi, G., Chizhenkov, G. V., Gromova, L. I., Parisi, M., Tyasto, M. I. and Yanke, V. G. 2006, Spacecraft operational anomalies and space weather impact hazards, *Advances in Space Research*, **37**, 184.

Judge, P. G., Casini, R., Tomczyk, S., Edwards, D. P. and Francis, E. 2001, Coronal Magnetometry: A Feasibility Study, *NCAR Technical Report* **NCAR/TN-446-STR**.

Kuhn, J. R., Coulter, R., Lin, H. and Mickey, D. L. 2003, The SOLAR-C off-axis coronagraph, in Innovative Telescopes and Instrumentation for Solar Astrophysics, *Proc. SPIE*, **4853**, 318.

Lambour, R. L., Coster, A. J., Clouser, R., Thornton, L. E., Sharma, J., Cott, T. A. 2003, Operational impacts of space weather, *Geophysical Research Letters*, **30**, 36.

Lanzerotti, L. J., et al. 2003, The Sun to the Earth - and Beyond, A Decadal Research Strategy in Solar and Space Physics, *N. R. C. Space Studies Board*, The National Academies Press, Washington, D.C., p 54.

Lin, H., Kuhn, J. R. and Coulter, R. 2004, Coronal Magnetic Field Measurements, *ApJ*, **613**, L177.

Nelson, P. G. 2006a, An Analysis of Scattered Light in Reflecting and Refracting Primary Objectives for Coronagraphs, *COSMO Tech Note 4*. Boulder, CO, High Altitude Observatory.

Nelson, P. G. 2006b, A FEA of Meter-Class Refracting Primary Objectives for Coronal Polarimetry, *COSMO Tech Note 2*, Boulder, Colorado, High Altitude Observatory.

Nelson, P. G. 2008, The Feasibility of Large Refracting Telescopes for Solar Coronal Research, in Ground-based and Airborne Telescopes II, *Proc SPIE*, **7012**, 701231.

Olivier, S. S., Seppala, L. and Gilmore, K. 2008, Optical Design of the LSST Camera in Advanced Optical and Mechanical Technologies in Telescopes and Instrumentation, *Proc. SPIE*, **7018**, 70182G.

Tomczyk, S., Card, G. L., Darnell, T., Elmore, D. F., Lull, R., Nelson, P. G., Streander, K. V., Burkepile, J., Casini, R., and Judge, P. 2008, An Instrument to Measure Coronal Emission Line Polarization, *Solar Physics*, **247**, 411.

Tomczyk, S. and McIntosh, S.W. 2009, Time-distance Seismology of the Solar Corona with CoMP, *ApJ*, in press.

Tomczyk, S., McIntosh, S.W., Keil, S.L., Judge, P.G., Schad, T., Seeley, D.H. and Edmonson, J. 2007 Alfven Waves in the Solar Corona, *Science*, **317**, 1192.

Vourlidas, A., Wu, S. T., Wang, A. H.;,Subramanian, P., Howard, R. A. 2003, Direct Detection of a Coronal Mass Ejection-Associated Shock in Large Angle and Spectrometric Coronagraph Experiment White-Light Images, *ApJ*, **598**, 1392.

-

SAFARI: Solar Activity Far Side Investigation

M. Velli, A. Ruzmaikin, N. Murphy, V. Angelopoulos, P.C. Liewer (JPL); O. Panasenco (Helio Research); A. Kosovichev, P. Scherrer (U. Stanford); F. Hill, R. Komm, J. Harvey (NSO); D. Braun, C. Lindsey (CORA); P. Riley, Z. Mikic (PSI); S. Jeffries (U. Hawaii); M. Thompson, S. Tomcszyk (HAO); B. DePontieu (Lockheed); F. Berrilli, A. Bigazzi, P.F. Moretti (U. Roma); L. Gizon, S. Solanki (MPS).

1. Introduction

The Sun's magnetic field dominates the outer solar atmosphere, structuring the coronal plasma and driving its energetic manifestations on all scales. Solar magnetic activity follows the well known ~11 year cycle, thought to be due to a dynamo operating in the solar interior, from the convection zone down to just below the tachocline, the shear layer at the base of the convection zone. The strongest markers of the solar cycle on the surface of the Sun are Sunspots, whose total area follows the solar activity cycle. Sunspots typically occur within larger magnetic complexes called active regions. There is a tendency for active regions to recur at the same longitudinal position on the Sun during a solar cycle and even from one cycle to another resulting in clustering of active regions at these *active* longitudes. Above active regions, permeating the solar corona, one observes loops, arcades of loops and streamers; overlying neutral lines in the photosphere one observes dark filaments against the disk revealed to be prominences surrounded by dark cavities at the limb. Explosive reconfigurations of these magnetic structures give rise to flares and coronal mass ejections (CMEs), which are among the main drivers of space weather.

The Solar Activity Far Side Investigation, or SAFARI, is a small-mission concept, or an element of a larger mission, devoted to exploring the origins of solar magnetic activity by carrying out observations of the velocity and magnetic fields at the solar surface from a vantage point widely separated from Earth in longitude. SAFARI images the Sun from an orbit trailing the Earth at 1 AU (SAFARI-S), while at the same time the Sun is imaged from the Earth (SAFARI-E). SAFARI carries out these observations using a compact Doppler magnetograph based on a simple, robust design with magneto-optical filters. SAFARI's ground based component, SAFARI-E, uses a similar observational technique, allowing precise inter-calibration of magnetograms and providing an opportunity to implement the novel technique of stereoscopic helioseismology, probing flows and structural heterogeneities deep in the convection zone, reaching below the tachocline and opening a new observational window into the Sun.

The combined measurements of solar magnetic fields from Earth and spacecraft viewpoints extends the longitudinal coverage of the solar disk allowing extended simultaneous observations which permit the full study of active region development and decay which cannot be observed in its entirety from a single point due to solar rotation. Global helioseismology will also naturally benefit from this situation since the modes of oscillation will be easier to disentangle. With stereoscopic local helioseismology, new acoustic ray paths can be taken into account to probe deeper layers in the interior, including the bottom of the convection zone. The STEREO spacecraft have shown how observing the solar corona and heliosphere from multiple vantage points can lead to breakthrough science concerning the some of the most dynamic events of solar activity, Coronal Mass Ejections: SAFARI provides simultaneous magnetograms for coronal structures which are on the limb when observed from the Earth, allowing

magnetohydrodynamic modeling of the corona and of coronal mass ejections ab-initio from measured magnetic fields at the Sun all the way into the heliosphere.

SAFARI directly studies changes in the Sun's interior and how those changes fully develop on the Sun's surface: flows will be inferred using all available methods of local helioseismology, including time-distance helioseismology, ring-diagram analysis, helioseismic holography, and direct modeling. SAFARI also provides a view of the Sun's poles from inside the Sun—areas hard or impossible to study from Earth. Finally, SAFARI's widely separated viewing points allow the separation of magnetic activity changes over time versus changes over location.

2. Science Goals

SAFARI's science goal is to quantify where solar variability originates, how it emerges to the solar surface and develops there.

Top level scientific objectives are

- What are the sources of solar magnetic activity clusters and preferred longitudes?
- What are the structure of flows and solar rotation in the solar interior?
- How do active regions grow and evolve on the Sun and what determines their persistence?

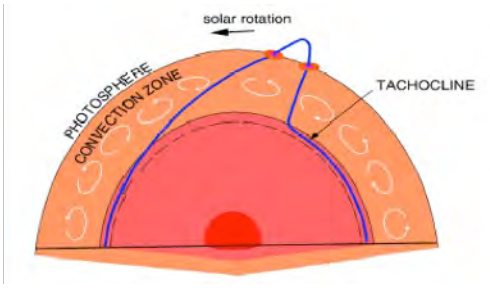


Figure 1 Equatorial cross-section of the Sun. The magnetic field (indicated by the thick line) generated in the tachocline emerges through the photosphere giving birth to active regions.

The objectives will be achieved by

- Detecting mass flows and thermal inhomogeneities deep inside the Sun
- Measuring the evolution of active regions over time scales of hours to months, and collecting data for reconstructing the magnetic field in the corona on time scales consistent with actual solar changes.

These prime science objectives are carried out by simultaneous Doppler and magnetic imaging from vantage points around the Sun with SAFARI-S, along with the Earth-side imaging from SAFARI-E components operating at Global Oscillation Network Group (GONG) sites, the Solar Dynamic Observatory (SDO), and potentially Solar Orbiter, through the use of helioseismic and magnetographic techniques.

What are the sources of solar magnetic activity clusters and preferred longitudes?

Theoretical studies indicate that solar variability originates from the interplay between magnetic fields and motions within the solar convection zone (Fig. 1). Modeling emphasizes a zone inside the Sun called the tachocline, a shear layer at $0.7R_S$. Fields, observed on the solar surface, change on time scales from minutes to the 11-year solar cycle and longer and conducting plasma moving in the solar convection zone is believed to be a generator (dynamo) for the solar magnetic field. Although the average internal solar rotation has been reconstructed with SOHO

and GONG data, basic plasma motions and the magnetic fields that the motions generate remain poorly determined.

Magnetic fields generated by the dynamo emerge through the photosphere, diffuse, reconnect, and move with the solar wind into the heliosphere. Magnetic activity often forms a particularly simple and interesting pattern: it clusters at a few "preferred" solar longitudes for many solar rotations. Clustering is a source for the most geoeffective fast CMEs. The preferred longitudes manifest themselves in the interplanetary magnetic field on time scale exceeding the solar cycle. The cause of the clustering remains unknown; it may be a manifestation of non-axisymmetric modes of the dynamo or a consequence of the non-axisymmetric structure of large-scale motions.

By imaging fixed inner regions of the Sun, SAFARI detects structures associated with deep convection and measures wave propagation speed anisotropy, a clear indication of the presence of magnetic fields. Measuring the field directly remains a challenging task. A theoretically estimated toroidal field which originates near the tachocline and emerges to the solar surface forming active regions is about 300 kG. The travel time difference produced by the field of this strength is beyond the accuracy presently achievable. Potentially detectable field enhancements of the order of 1000 kG are, however, not excluded and can be measured by using repeating time series and multiple bounces of the wave packets traversing the magnetic field concentrations. SAFARI detects the strong field enhancements and provides better limits on averaged magnetic field strengths near tachocline than any current or planned instruments.

SAFARI's Doppler imaging from extended longitudes combined with the Earth-side imaging is designed for detection of dynamo flows and their asymmetry inside the Sun.

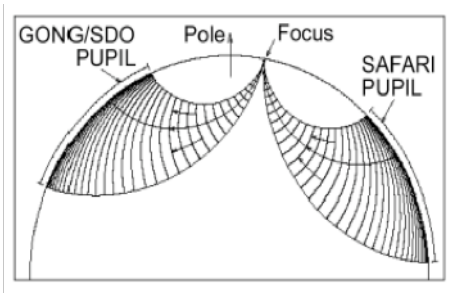


Figure 2. Wave configuration for probing a near-surface activity in the neighborhood of a solar pole. The "coherent acoustic egression" of the focal point is computed based on observations over a "pupil" visible from the Earth, while the "ingression," is computed over a conjugate pupil viewed by SAFARI, extrapolated by assuming reflection at the focus.

What are the structure of flows and solar rotation in the solar interior ?

The major systematic motions throughout the Sun are associated with rotation, and depend on latitude and depth, i.e., they are differential. Gradients of rotation, which stretch magnetic field lines into strong toroidal belts, drive the dynamo. Differential rotation is observed on the solar surface and reconstructed throughout the convective zone from GONG and SOHO data using the inverse methods of global helioseismology. The reconstruction, however, becomes inaccurate near the poles, near the tachocline, and deeper into the Sun's core.

SAFARI's three-dimensional deep probing provides improved gradients of rotation near the tachocline and determines the core rotation. SAFARI detects temporal variations of the rotation, such as apparent 1.3-year oscillations, revealed by GONG and SOHO MDI data with limited latitudinal and no longitudinal resolution. The Sun's polar regions have been poorly explored to date because Earth-based imagers are foreshortened at high solar latitudes. There are a number of unanswered questions concerning the structure and dynamics of these regions: the nature of the circumpolar jet streams inferred from SOHO data (related to meridional circulation), the origin of polar faculae, and the structure of solar convection near the poles.

SAFARI-S and SAFARI-E with SDO imaging open a unique opportunity to explore the near polar, internal solar regions using "acoustic lenses" focused near poles (Fig. 2) and wave packets excited on the Earth side of the Sun, bounced near poles and detected on the far side of the Sun.

Modeling indicates a critical importance of meridional motions and their location for the dynamo. Meridional circulation at ~10 m/s has been measured on the solar surface but the depth of its penetration into the Sun, in particularly a possibility to flow beyond the tachocline remains unknown. First attempts were made to trace it into the Sun using SOHO/MDI data but the accuracy was insufficient to determine the meridional flow near the base of the convection zone.

The longitudinal separation of simultaneous Earth-side and SAFARI Doppler images allows the detection of meridional flows near the base of the convection zone through a unique tracing of acoustic waves traveling under the solar pole.

Solar convection transports magnetic fields and heat within the Sun. A standard approach to studying this is parameterization of the convection in the form of turbulent diffusivity. This parameterization, however, completely loses the structure of convection, which is important for understanding the interaction between the motions and magnetic fields. A long-standing problem in solar physics is identification of the largest (called giant) convection cells. Giant cells must have lifetimes comparable to the Sun's rotation period; so rotation and convection must interact at a fundamental level. Motions at this scale must interact with global-scale magnetic fields driving the mean-field dynamo, which is responsible for the 11-year activity cycle. At present, only faint hints of an observational signal for this mode of convection have been detected.

By identifying and measuring convective structures, SAFARI establishes detailed dynamic conditions in the solar interior.

In the deepest parts of the Sun reside the low-frequency low-degree p modes (less sensitive to surface disturbances) and particularly interesting g modes (mostly confined to the radiative interior). These modes are keynotes for asteroseimology. Unfortunately, their detection and identification on the Sun are hindered by the presence of non-coherent convective motions, and by magnetic perturbations forming active regions. The lowest frequency structure reported and (tentatively) attributed to an acoustic p mode, is 0.535 mHz (from GOLF and MDI data). The validation of the nature of this structure and detection of the whole set of low-order p-modes, require a substantial improvement of signal to noise ratio.

Up to now, no unambiguous detection of a g-mode signature is reported. The upper limit for the velocity amplitude of g-modes on the solar surface is estimated to be 3 mm/s.

The use of SAFARI's stereoscopic helioseismology advances the search for the low-frequency modes. The realizations of velocity pattern observed by SAFARI-S are different from those observed by front side instruments (SAFARI-E, SDO), while the coherent signals coming from the modes are the same. The cross-correlation of SAFARI-S and SAFARI-E/SDO images greatly improves the signal-to-noise ratio needed for the detection of the low-frequency modes.

As the Sun formed out of a gas cloud, its mass, angular momentum and magnetic field have undergone a major re-distribution leading up to its current state. Although these processes stopped long ago, the core of the Sun, as the most deeply buried part of the solar system, may retain some clues of the early times, such as chemical inhomogenieties, remains of the initial angular momentum distribution, and of primordial magnetic field.

SAFARI's use of stereoscopic helioseismology to probe the Sun's core and its inhomogeneities, motions, and fields can bring some answers to these fundamental questions.

-

How do active regions grow and evolve on the Sun and what determines their persistence?

SAFARI is the first mission to observe the solar photospheric magnetic field from a location substantially far from the Earth-Sun line. Together with simultaneous Earth-based observations, it will revolutionize our understanding of the development of solar activity.

The magnetic field evolution for a specific area on the Sun cannot be observed for more than about 10 days because it rotates out of view from the Earth. Active regions systematically evolve through significant portions of their life on the unseen side of the Sun, so their evolution from one configuration to another is not studied. It is unclear what makes some active regions so energetic, while others are placid, why do some active regions last for a long time while others, seemingly similar, quickly vanish, and what is the magnetic flux budget history of an active region.

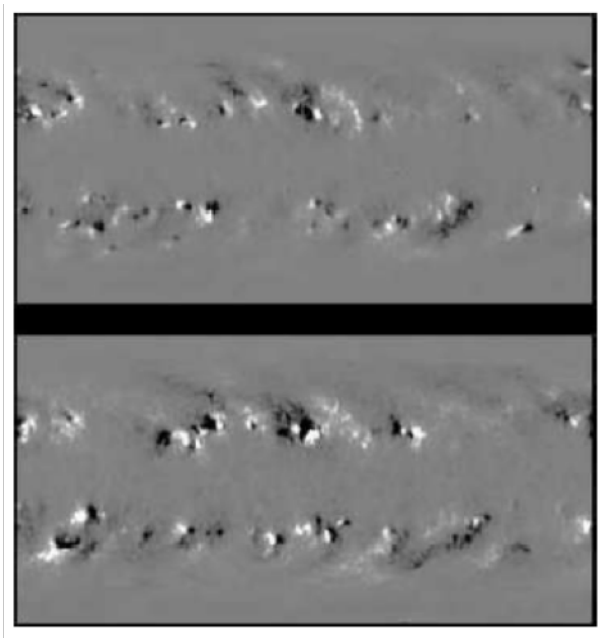


Figure 3. Solar magnetic features evolve over a solar rotation as seen in the image. The top magnetogram is from October 8, 1998, and the bottom from November 4, 1998 (K. Hanna of GONG-NSO).

In concert with Earth-side magnetic observations, SAFARI follows the detailed development of individual active regions and clusters of activity. This allows the Sun's large-scale magnetic fields to be studied without confusing spatial and temporal properties.

Clusters of activity are the sources of major fast coronal mass ejections (CMEs) and solar energetic particles and should be studied carefully.

SAFARI characterizes the magnetic conditions that lead to CMEs and related explosive events. It has long been known that flares associated with CMEs tend to occur in complex, dynamic, and sheared magnetic configurations. Recently, simultaneous, high-cadence magnetic field ground and space observations have consistently revealed the time scale and location of flare-related magnetic field changes. However, physical interpretation of these observations is ambiguous because they provide only the line-of-sight component of the vector magnetic field. It is not known if the changes are due to variations of field strength, direction, or location. Current vector field observations have lower sensitivity than required and thus have great difficulty in making and interpreting such observations. In concert with Earth-based observations, *SAFARI eliminates much of the ambiguity in interpretation by making robust, line-of-sight observations through two different views of the same event.* A range of Earth-Sun-spacecraft angles out to ~105° is satisfactory for this purpose. The related CMEs are observed from the Earth-side by the SDO coronal imagers.

SAFARI-S magnetic observations from the far side of the Sun combined with the SAFARI-E Earth-side magnetograms will bring us for the first time almost quasi full-Sun boundary

conditions. These boundary conditions will replace currently used synoptic maps and greatly advance MHD modeling of the corona, heliosphere, and solar interior. In particular, they will allow greatly improved reconstruction of the coronal magnetic field and the sources of the solar wind, allowing us to test our understanding of the formation of the corona and the sources of the solar wind with unprecedented accuracy.

3. Mission Concept

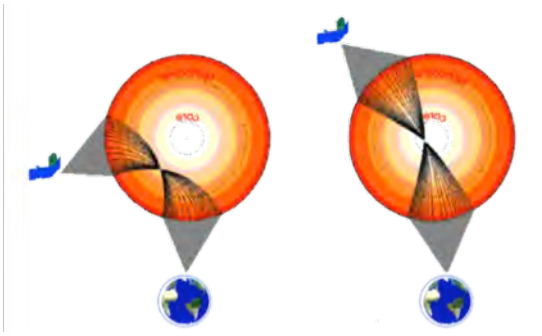


Fig. 4. Two views of SAFARI relative to the Earth along with multipoint helio-seismic observations penetrating the tachocline and the solar core

The SAFARI investigation concept involves sending a spacecraft into a trailing orbit around the Sun at 1AU using either direct launch or lunar gravity assist. The mission acquires a 45° – 60° drift from Earth per year and can accomplish its major science objectives within 2.5 years, with the possibility of an extended mission reaching Sun-Earth opposition in 4 years. SAFARI at a minimum carries only one instrument, a compact Doppler magnetograph. A simple realization sufficient to attain the scientific objectives is a 10 cm aperture telescope with a dual focal plane magneto-optical filter assembly allowing full Sun images for right and left hand circular polarizations in two wings of the potassium 770 line. Two 2048 x 2048 detectors receive 4 images with a cadence of 60 s. These are binned down to 512 x 512 and sent back every 60 s as Dopplergrams (with a resolution of 8 arcsecs) to reduce the overall data rate, while full resolution magnetograms (2 arcsecs) are obtained with a cadence of 10 minutes. To follow active region evolution and the helioseismic effects of strong magnetic field concentrations, the capability of higher resolution downloads of selected regions on the Sun will also be implemented.

4. Conclusions

SAFARI is a well-defined investigation concept that would lead to significant advances in our study of the magnetic activity of our star, the Sun. SAFARI addresses specific goals in the NASA Heliophysics roadmap and Strategic plan, and in particular the 2006 NASA Strategic Plan subgoal 3B, “understand the Sun and its effects on Earth and the solar system” and the 2009 Heliophysics Roadmap Priority Science Targets Heliospheric Magnetics (HMag): Understand the flow and dynamics of transient magnetic structures from the solar interior to Earth, topics F Open the Frontier to Space Environmental Prediction, sub-topic F4 Creation and variability of magnetic dynamos; topic H Understand the Nature of Our Home in Space, sub-topic H1 Causes and evolution of solar activity; and topic J Safeguard the Journey of Exploration sub-topic J2 Capability to predict the origin, onset, and level of solar activity.

-

Mission to the Sun-Earth L5 Lagrangian Point: An Optimal Platform for Heliophysics & Space Weather Research

Submitted By:

A. Vourlidas, R. Howard, Y.-K. Ko, D. Biesecker, S. Krucker, M. Thomson, N. Murphy, T. Bogdan, C. St Cyr, J. Davila, G. Doschek, N. Gopalswamy, C. Korendyke, M. Laming, P. Liewer, R. Lin, S. Plunkett, D. Socker, S. Tomczyk, D. Webb

1 SUMMARY

Our field of research, heliophysics, is evolving rapidly. Sun-Earth interactions, once treated as a area of pure basic research, has now evolved into the field of Space Weather with concrete and important societal benefits as well as a strong basic research interest. The once separate research fields of solar, geospace and space physics are integrating, leading to much more productive science, under the *heliophysics* label. These accomplishments are driven by a string of highly productive NASA missions (e.g., *Wind, SOHO, ACE, STEREO*) and are based on a very successful systems approach in the mission operations and analysis of their data. We have learned that some of the observations obtained from these missions, need to be available routinely to enable the science to address the next tier of questions.

In this paper, we argue that a *research-to-operations (R2O)* approach is the best strategy to foster the vibrancy of our field in the next decade. A mission to the L5 Lagrangian point should spearhead this strategy because it maintains the highly successful and productive systems approach to NASA's missions, augments the output and societal benefit of the LWS program, and enhances research both for the benefit of future space weather applications and for the basic understanding of the Sun-Earth system. NASA has implemented such a strategy before with extreme success. The *Nimbus* series of geostationary satellites drove atmospheric physics research for decades and simultaneously established dependable 3-day weather forecasting. In the end, NASA passed the baton over to NOAA with the creation of the operational series of GOES satellites. Heliophysics is at the same level of development now as atmospheric physics was in the mid-1960s and hence a NASA-led R2O program is the next logical step.

Our society's growing technological dependence on GPS, telecommunication, an interconnected power grid, global travel, and expected growth in space travel focuses attention to the conditions in Earth's near space and increases the demand for accurate space weather predictions. Research from NASA missions has uncovered the drivers of space weather and continues to provide new insights into the physics of solar eruptions. The realization that basic research from space can benefit societal needs led to the creation of a new mission line and research program—the Living With a Star (LWS) program which aims to improve our understanding of space weather through a series of research missions in geospace. However, the LWS missions were conceived and planned early in this decade before we had a chance to absorb the results from a full solar cycle of *SOHO* and *ACE* observations and to assimilate the breakthrough science that is currently coming from the *STEREO* mission. These missions are teaching us that even the venerable solar cycle may not be as predictable as we once thought, and that simply following the evolution of the solar structures on their journey to Earth can provide critical information for space weather studies. Having the benefit of hindsight, we argue that the current heliophysics mission plan fails to address the critical science of the Sun-to-Earth evolution of the solar wind and hence it may be inadequate to fully meet the LWS program goals in the near future.

To improve our predictive capabilities, we first have to understand how the solar structures evolve during their propagation to Earth. But to understand that evolution, we need to observe the propagation of a given structure and its interaction with the ambient solar wind continuously. Continuous coverage is necessary because of the complexity of the ambient solar wind (*e.g.* fast and slow streams, preceding

events) which can alter the structure of a coronal mass ejection (CME), or alter its path or both. Orbital dynamics (and costs) make it almost impossible to monitor the Sun-Earth space with a series of in-situ probes along the way like the NOAA National Data Buoy Center tsunami warning network of buoys dispersed in the Pacific. There is only one stable orbit at the L1 Lagrangian point, located about one hour upstream of Earth. The optimum solution is the L5 Lagrangian point from here we can survey the whole Sun-Earth space at once. The L5 point is located 60º east of the Sun-Earth line and is a gravitationally stable location.

The *Nimbus* program offers a proper analogy. The L5 orbit is the space weather equivalent of the geosynchronous orbit for weather satellites. Reliable 3-5 days predictions of terrestrial weather became viable only after the successful launch of the *Nimbus* satellite series in geosynchronous orbit in 1964-74. We argue that a heliospheric mission to L5 can accomplish a similar leap in space weather prediction and in our understanding of the inner heliosphere. The L5 point is a uniquely advantageous viewpoint because it provides global coverage of the inner heliosphere, allows imaging of solar activity and the in-situ sampling of solar wind structure, such as CIRs, four to five days before it arrives at L1 or geospace.

Table 1 Summary of L5 Mission Concept

Mission Profile	L5 orbit: 60º nominal relative to Sun-Earth line, orbit may vary from 40º – 90º relative to Sun-Earth line.		
Strawman Payload	EUV disk imager to 2 Rs. Line-of-sight Magnetograph HXR imager/spectrometer	White light coronagraph Heliospheric Imager Off-limb UV Spectrometer	On-disk spectrometer/imager Fields & Particles Solar wind composition.
Operational Capability	- 4-5 day warning for recurrent disturbances (e.g., CIRs, Irradiance variations). - 4-5 day warning of coronal activity (e.g., flux emergence, flaring history of active region. - 5hr – 4-day CME prediction capability using streamer swelling (**L5 unique capability**). - Few-hr eruption prediction by measuring height of Earth-aligned rising prominences (**L5 unique capability**). - Evaluation of CIR geoeffective potential by in-situ measurements of CIR shock and energetic particle production 4-5 days before they hit Earth. (**L5 unique capability**).		
Scientific Capability	- Solar wind measured at L1 originates at limb as seen from L5. - Sources/footpoints of Earth-connected field lines are occulted allowing HXR/SXR imaging of looptop sources, and measuring altitudes of erupting loops. - Sources/footpoints of field lines connecting to L5 s/c are close to central meridian (from Earth) allowing direct measurement of the photospheric magnetic field and coronal conditions. - Long-term studies of the heliosphere with correlation in properties between the solar structure and the solar wind at two vantage points. - Stereoscopic analysis of EUV sources, CMEs and CIRs using Earth-orbiting assets. - Two-point measurements of SEPs place stringent tests on models of SEP production. - Measurement of the longitudinal structure of the tachocline.		
Cost	$500M to $1B.		

The mission concept builds on our growing experience with observations off the Sun-Earth line from the *STEREO* mission[1]. The Behind *STEREO* spacecraft (STEREO-B) actually crossed the L5 point during 2008-09. It obtained many observations of Earth-directed events which provide the proof-of-concept for our proposed mission. For example, STEREO-B data are regularly used by NOAA for space weather assessment while active research from many groups is providing new insights in the evolution of CMEs. Hence, a future L5 mission will have a unique blend of operational and science objectives. It will also greatly enhance our understanding in solar and heliospheric physics by providing a second vantage point with Earth/L1.

[1] Webb, D. F. et al. 2010, Space Research Today, 178, 10

The mission science objectives, instruments and anticipated measurements are summarized in Table 1. In the following, we describe a set of key science questions that could be answered with this mission concept.

2 KEY SCIENCE QUESTIONS

2.1 L5 *is an Ideal Location for Heliophysics Research*

Observations from L5 can help answer key science questions over the full spectrum of solar and heliophysics studies starting from helioseismology and extending to CME initiation and the propagation of SEPs in the heliosphere. These questions are outlined below and the research benefits of an L5 mission are summarized in **Error! Reference source not found.**. The L5 location permits the direct viewing of structures propagating towards Earth, in contrast to the L1 location.

2.1.1 What Is The 3-D Structure Of The Corona Before An Eruption?

EUV and SXR observations of the low corona have shown that the most energetic CMEs originate in active region neutral lines. These regions are often characterized by complex loop topologies that include loops highly inclined from the local normal, serpentine filaments, semi-circular loops and loops with distant connections outside the active region. However, only a small subset of these structures participates in the formation of the CME and is subsequently ejected while the rest of the loop system remains relatively stable. Why is that? Why only part of the neutral line is activated during a CME? What determines the length of the post-eruptive flaring arcade? What is special about the erupting loops?

Table 2 Mission Summary

We cannot answer these questions because we do not yet have a good handle on the 3D magnetic topology of the corona, the sites of energy storage, or the distribution of currents within the coronal part of an active region. It is difficult to discern the evolution of erupting structures low in the corona from a single viewpoint. The spatial scales are small and the magnetic topology in the core of active regions is quite complex. The problem is more acute for observations from an Earth-based location because the structures that are ejected towards Earth are at large angles from the sky plane and hence almost invisible. Recent STEREO EUV work has shown that single viewpoint observations, even with high cadence, can lead to misidentification of the erupting structure[2]. Other STEREO observations have associated jets (and 3He rich events in the heliosphere) with small impulsive flares[3]. Are these flares the sources of the seed particles for SEP events?

STEREO results demonstrate the power of two viewpoint observations even at large angles from Earth. Unlike STEREO, our strawman payload includes both a magnetograph and an X-ray imager and will remain in orbit around L5. Thus, the mission will provide stereoscopic and quadrature observations of flares and erupting active regions on a routine basis with the combination of an Earth-based platform (e.g., from SDO).

2.1.2 What Are The Sources Of Seps And How Does Scattering Affect Their 1 AU Properties?

The uncertainty of the source of SEPs is a major hindrance in understanding their production and propagation. Both the flare reconnection region and the CME-driven shock are viable candidates for SEP production. A combination of the two, as flare-accelerated particles may be further processed by the CME shock, is another likely scenario. Composition is a way to trace the origin of SEPs but measurements at 1 AU are inconclusive because we have limited compositional information in the inner corona and lack sufficient *concurrent* off-limb coronal observations above the erupting site to understand the environment

[2] Patsourakos, S., Vourlidas, A., Kliem, B., 2010, A&A, in print
[3] Mason, G., 2007, Sp. Sci. Rev., 130, 231

through which SEPs accelerate and propagate, such as the seed particle population characteristics and shock geometry. How can we resolve the origin of SEPs? The answer to this question will have profound implications both for basic physics and for space weather predictions.
Any SEP events detected by an L5 spacecraft will tend to propagate along magnetic field lines from a flaring region located on the Earth-viewed disk thus allowing detailed measurements of the photospheric magnetic field and EUV structure. The L5 hard X-ray imager will frequently observe the looptop emission from these sources since many of the flaring regions will be occulted or close to the western limb as seen from L5. The associated CMEs and their shocks will be imaged clearly as limb events. The off-limb UV spectrograph will provide composition, density and temperature measurements of the corona above the flaring site *before and after* the event, the associated CME and its shock, as well as the post-CME current sheet. In-situ measurements at L1 and L5, such as the arrival time of SEPs at different energies and their composition will provide another piece of the puzzle. The relative timing between HXR, and EUV reconnection events, the CME shock parameters, the coronal conditions derived from the spectrograph data and the in-situ measurements, and the application of transport models will bring a new understanding not only on the origin of SEPs but also on their propagation in the heliosphere.

2.1.3 What Is The Connection Of Hard X-Ray Looptop Sources To The CME And Accelerated Particles?

RHESSI observations of looptop HXR sources have shown that these sources are visible whenever the bright emission from the flaring footpoints is occulted[4]. Some hard X-ray sources are located up to 100 arcsec or more above the surface and may contain a surprising number of non-thermal particles[5,6]. How are these particles accelerated and where do they come from? When combined with EUV observations, the HXR sources tend to lie close to the post-CME current sheet. So the HXR source and the post-CME current sheet may be closely related. But how are the HXR sources related to the CME formation? Are they a source of SEPs that may be later reaccelerated by the CME shock? Do the HXR footpoints mark the locations where the particles are accelerated instead?

To answer these questions we need frequent multiple view-angle observations of HXR looptop sources in association with their footpoints and with eruptive events. These are difficult to obtain from Earth as the scarcity of such events reveals. As discussed above, the footpoints of occulted flares, that are best for the detection of looptop sources from an HXR imager at L5, will be visible from Earth. While Earth-based SXR or EUV observations are frequently available, an HXR image at L5 will provide observations from a different view angle to study partially occulted events in combination of on-disk events.

2.1.4 What Are The Origins Of The Solar Wind Detected At L1?

The most obvious contribution of L5 observations is on solar wind studies. Currently, we cannot make a direct connection between the solar wind arriving at Earth and its origins. Upcoming missions such as Solar Orbiter and Solar Probe will focus on the problem of the solar wind origin but both are encounter missions providing measurements only for a few days/orbit through perihelia. An L5-orbiting spacecraft will allow long-term synoptic studies of the heliosphere. Tracking variations in the solar wind with corresponding variations in the solar magnetic and EUV structures at two vantage points will provide insights into the origin of the solar wind.

The solar wind intercepted by Earth including CMEs, CIRs and streamer blobs, is emanating from the western limb as seen from L5. It is therefore easily detectable, with minimal projection effects, by the

[4] Krucker, S., Lin, R. P., 2008, ApJ, 673, 1181
[5] Krucker, S., et al., 2007, ApJ Lett., 671, 193
[6] Krucker, S., et al., 2010, ApJ Lett., 714, 1108

white light coronagraphs, EUV imager and off-limb spectrometer. The evolution of the wind is also easily traced from L5 because the wind remains within the Thomson surface[7] during its propagation to 1 AU.

2.1.5 Where Do Active Regions Come From?

The large-scale solar dynamo is still not understood. Helioseismology with global modes has revealed the existence of the tachocline rotational shear layer at the base of the convection zone and has enabled its radial and latitudinal extent to be quantified[8]. The tachocline is now widely believed to be the seat of the solar dynamo and there is evidence of changes in wave-speed in the tachocline over the solar cycle[9], suggestive of a varying magnetic field there. Various theoretical works suggest that MHD instabilities in the tachocline should give rise to *longitudinal* variations in the region[10]: such longitudinal structures may be the origin of "active longitudes" where active regions preferentially emerge over multiple solar rotations and even over successive solar cycles. Time-distance helioseismology provides a way of mapping the solar interior in three dimensions. To date, however, their application has been restricted essentially to the upper convection zone, mapping bulk flows, magnetic structures, and thermal inhomogeneities there. The reason for the limitation is that to probe down to a given depth in the Sun with these techniques requires that both ends of acoustic rays that penetrate at least to that depth be observable simultaneously: observing the Sun from a single vantage point -- particularly given foreshortening as one observes further away from disk center -- severely limits which rays can be observed to those and means that only the upper convection zone is accessible. With identical magneto-optical filter (MOF) Doppler/magnetographs at both the Earth and the L5 point, time-distance helioseismic analyses using rays that penetrate to the tachocline and even deeper can be observed simultaneously, allowing longitudinal resolution of the region. As the Sun rotates, a full longitudinally-resolved synoptic view of the tachocline will be obtained. The technique can be used to reveal both bulk flows (using travel-time differences for waves traveling in opposite directions along the ray) and wave-speed anomalies (using mean travel times along rays). At typical 5-minute periods, waves corresponding to spherical harmonic degree 40 and smaller penetrate to the tachocline and below. This will provide a resolution in longitude at the tachocline of 10-15º using a 100x100 grid as a baseline for observations.

2.2 *L5 is an Ideal Location for Space Weather Studies*

A satellite located at L5 point will be trailing Earth by 60º and therefore will be able to image an additional 30º of the solar disk invisible from Earth. Since the solar rotation (from 1 AU) is about 13º/day, an L5 satellite will be capable of a 4-5 day warning regarding the emergence, complexity, and eruptive potential of active regions and irradiance variations. An L5 satellite will also provide up-to-date behind-east-limb magnetograms better suited for building background solar wind models that are essential for CME propagation models, and provide additional validation opportunity through the solar wind measurement at L5.

The progressive widening and brightening of a streamer, known as '*streamer swelling*', is a reliable short-term predictor of an eruption as it always precedes a many CMEs by a few hours to a few days (Figure 1). Synoptic coronagraph observations from L5 will provide our first reliable predictions for Earth-directed eruptions because Earth-aligned streamers will be located on the sky-plane as viewed from L5. The swelling rate is roughly inversely proportional to the final CME speed but we need more analysis to make these observations an operational tool. Additional information can come from low corona EUV observations of the rising prominence/cavity. Since these prominences become unstable after a certain height which is dependent on the separation of their footpoints, a simple height-time measurement during the early rise phase of the prominence can be used as a near-term forecasting tool.

---

[7] Vourlidas, A. & Howard, R. A., 2006, ApJ, 642, 1216
[8] Christensen-Dalsgaard, J. & Thompson M. J. 2007 in The Solar Tachocline, D. W. Hughes, R. Rosner & N. O. Weiss (eds), Cambridge Univ. Pr. Cambridge, UK, 2007, p.53
[9] Chou, D. -Y.& Serebryanskiy, A. 2002, ApJ, 578, L157
[10] Cally, P. S., Dikpati, M. & Gilman, P. A. 2003, ApJ, 582, 1190

-

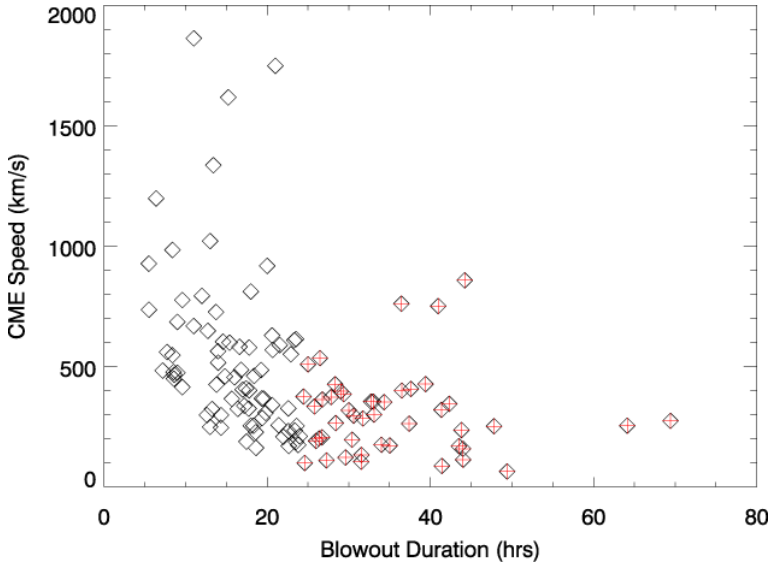


Figure 1 Relation between the duration of streamer swelling and the final speed of the associated CME[11].

After its ejection, the CME can be followed throughout its inner heliospheric path and its trajectory determined as it is currently done with STEREO[12,13,14,15,16]. Again, the L5 location is optimal for such studies because the CME will lie within the Thomson Surface of maximum scattering[7] throughout its travel to Earth. Even if there is no other coronagraph in operation (either at Earth or other vantage point) to provide stereoscopic information, the CME path can be estimated from a single viewpoint and the events tracked to large distance with the j-map technique[17]. The estimates can be adjusted continuously as observations further from the Sun become available and the Earth-arrival predictions can be updated.

In the heliosphere, the L5 satellite will use its in-situ instrumentation to detect recurrent disturbances such as CIRs and measure their geoeffective properties such as whether they are driving a shock, or whether they are accelerating energetic particles. These measurements will characterize the CIR geoeffective potential 3 days before crossing Earth thereby providing a high quality prediction which is unattainable from an L1 observatory. The benefits of an L5 mission to operational space weather are summarized in **Error! Reference source not found.**.

2.3 *Strawman Payload*

To meet the scientific and operational objectives outlined above, and the restrictions on spacecraft resources for a mission to L5, we have compiled a strawman payload in Table 2. We used the STEREO spacecraft as our starting point. It carries about 70 kg of science payload and operates within the X-band constraints (480 kbits/s) from L5. Technological progress will surely result in a more capable instrument payload with lower mass and power resources than STEREO. Already, CMOS imaging detectors developed for the Solar Orbiter and Solar Probe mission require 4x less power than the STEREO CCDs. Newer instrument designs enable us to cover the sun-earth space with 3 telescopes instead of the five telescopes on STEREO.

3 ESTIMATED COST

A mission to L5 is a deep space mission and will involve a large booster, moon and/or Earth flybys and DSN resources. Recent studies have shown that a stable orbit with minimum stationkeeping and fast downlinks (2-3 Mb/s) is achievable[18]. A proper mission study is required to derive a reliable cost. We

[11] Vourlidas, A. et al., 2002 in Solar-Terrestrial Magnetic Activity & Space Environment, Proc. of the COSPAR Coll., H. Wang & R. Xu. (eds), Boston: Pergamon, 2002, p.201
[12] Davis, C. J., et al 2009, GRL, 36, L08102
[13] Byrne, J. P., et al 2010, Nature Comm,, 1, 6, 74
[14] Liu, Y. et al 2010, ApJ, 722, 1762
[15] Lugaz, N. et al 2010, ApJ, 715, 493
[16] Lynch, B. J. et al 2010, JGR, 115, A07106
[17] Sheeley, N.R. et al, 2008, ApJ Lett., 674, L109
[18] Lo, M. Lianos, P., Hintz, G., AAS-AIAA Astrodynamist Specialist Conf, San Diego, 2/2010.

can, however, provide a rough estimate by using the STEREO mission as a basis. We adopt the same spacecraft, telemetry and mission profile. So our constraints are: (1) 620 kg total mass (2) 70 kg instrument mass, (3) 480 kbits/s telemetry over X-band, and (4) a relatively slow journey to L5 using lunar flybys. A slower drifting rate than STEREO's 22.5º/yr would be preferable because it allows more time for 2-point measurements with in-situ instruments hence increasing the science return of the mission. We estimate conservatively ~$240M for the instruments by assuming an average of $30M (FY10 dollars) per instrument (in-situ instruments may cost less but the spectrograph and magnetograph may cost more). Significant savings may be achieved if some imaging instruments (e.g., EUV imager, coronagraph, HI) are part of a suite as they were in STEREO. Since the total cost of the STEREO mission was ~$480M in FY06 dollars, we expect that a single spacecraft mission should cost less than $1B.

4 RELEVANCE TO THE DECADAL SURVEY

As we argued in the Summary, an L5 mission in the next decade is a logical next step for Heliophysics. The mission objectives extend across several of the National Capabilities Working Groups; namely, 'Research to Operations', 'Theory & Modeling', 'Explorers', and 'Innovations'. The mission combines both operational and research objectives and therefore has both societal benefits and advances the state of knowledge for the Sun-Earth system. It is not overlapping with the objectives of any other mission in operation around that time, yet it provides strong synergies with all of them. In this sense, an L5 mission is a cornerstone of the Great Observatory concept. The mission design and payload is of low risk. The STEREO mission has already demonstrated the feasibility and benefits of putting a payload of 620 kg at L5. Technological progress in hardware components (e.g., CMOS detectors, telemetry links) will result in the launch of a much more powerful platform than STEREO with reduced power and mass requirements. Most of the payload systems and components are at TRL 7 and above.

Table 2 Notional Payload for an L5 Mission

Telescope	FOV	Wavelength (Å)	Highlights	Mass (kg)	Power (W)	Telemetry (kbps)
EUV Imager	2 Rs (0° - 0.5°)	171, 195, 304, 284	STEREO Heritage	6	8	50
COR2 Coronagraph	2–15 Rs (0.5° – 4°)	600-700	STEREO Heritage	6	8	40
Heliospheric Imager	15–225 Rs (4° – 64°)	600-700	10x sensitivity over STEREO-HI	5	3	30
Off-limb Spectrograph	Two 3Rs-long slits at 2Rs and 3.5 Rs	500 – 1040	OVI dbl, CIII, Lyβ-γ, ArXII, SiXII, FeX, XIII, XVIII	4.4	4.4	50
Slit/Slot Imager	Full disk	170-210	Imaging spectroscopy 'on the cheap'	15	10	50
Magnetograph	Full disk	Na 5890 K 7700	Magneto-optical Filter	17	20	50
HXR Imager/Spectrometer	Full disk	64-150 KeV 0.7~1 keV resolution	Same as STIX on Solar Orbiter	14	24	0.5
Solar Wind Ions Solar Wind Electrons Energetic Particles Magnetometer	2π 4π 240º	5 º 22.5º	0.05 – 20 keV/q 0.005 – 6 keV/q 0.02 – 100 MeV/nuc	2 4 4 1.25	3.5 6 3.1 2	5

A Proposal to Lighten The Burden of International Traffic in Arms Regulations on Heliophysics Research

Submitted by:

A. Vourlidas (NRL), A. Rymer[†] (JHU/APL)

Background

It sounds illogical but it is true. Scientific satellites, the telescopes they carry, and the research needed to develop these telescopes, are classified as 'defense systems' by the U.S. government. This classification mandates adherence to the International Traffic in Arms Regulations (ITAR); a set of strict but wide-ranging controls on the export of technologies, hardware, and services that ordinarily applies only to items that are specifically designed or modified for military application. While we recognize that national security and foreign policy concerns necessitate such regulations, their blind application to basic space research brings only adverse effects. Specifically, the ITAR *stifle* international cooperation on which basic research thrives, *erode* the nation's leadership in space by limiting the participation of U.S. hardware groups in international projects, *create* a 'brain-drain' of highly educated and creative foreign researchers who flee the country; enhancing our competitors capabilities and diminishing our own, and *increase* the risk of mission failures due to lack of communication between the international partners. These problems affect every program with the NASA SMD, including Heliophysics.

As this is an on-going problem that affects everyone in the community, the Decadal Survey offers a suitable platform to raise this issue again, to demonstrate its burden on our daily workload and efficiency, and to propose specific actions that will maintain the scientific leadership of the U.S heliophysics program without raising security concerns.

History

Export control relating to scientific satellites was not always under ITAR jurisdiction. Between 1992 and 1996, export control was transferred to the Department of Commerce under the Export Administration Regulations (EAR). But in 1999, following a Congress-led investigation[1] of an ITAR breach by Space Systems/Loral, satellite systems were returned to ITAR. However, no effort was made to harmonize the very different EAR and ITAR regulations concerning space systems and associated research creating considerable confusion and conflicting regulations or interpretations. We note here that the hardware and launch activities of the joint ESA-NASA *Solar Heliospheric*

[†] Representing herself.

[1] U.S. House of Representatives, *U.S. National Security and Military/Commercial Concerns with the People's Republic of China*, Select Committee on U.S. National Security and Military/Commercial Concerns with the People's Republic of China, U.S. Government Printing Office, January 1999.

Observatory (*SOHO*) mission took place under EAR control. We have considerable doubt that the mission could have been as successful under an ITAR regulatory regime given the highly complex mixture of U.S. and European instruments, sensors, and overall hardware responsibilities.

Many of these problems were captured by a National Research Council Workshop and Report on the effects of ITAR on space science[2]. We used this report extensively in compiling this white paper since its findings are directly relevant to Heliophysics and to NASA-led missions.

Problems on Heliophysics Research

International Cooperation in Fundamental Research. The confusion surrounding ITAR and the serious repercussions of failing to comply with them are restricting the exchange of information between U.S and European or Japanese colleagues. For example, while U.S.-based hardware groups are the clear technology leaders in many observing regimes (e.g., EUV telescopes), they cannot easily influence future mission designs and payload arrangements from foreign space agencies since they cannot participate in detailed discussions during the mission concept phases. The absence of U.S. expert input could lead to poor mission planning, high mission risk and diminished scientific return. For example, the Chinese and European space agencies have been discussing for several years a solar mission to L1, called *Kuafu.* The payload contains instruments very similar to the NASA instruments flown on *SOHO*. For anyone familiar with space hardware development, it is clear that those instruments cannot be assembled, within acceptable risk, by third parties without U.S. know-how. However, much of that know-how resides in federal laboratories and institutions (e.g, NASA). Due to ITAR restrictions (perceived and real) U.S. scientists avoided participation in these discussions so far. We believe that, the lack of U.S. input into the *Kuafu* mission may lead to a wasted opportunity for new Heliophysics science. Amending ITAR to ease such international collaborations will certainly lead to better science and improved launch opportunities (e.g., missions of opportunity) for NASA.

Brain-Drain. Future hardware development relies on the continuous training of young scientists in the nuances of space hardware design and implementation. However, the pool of interested engineering and applied physics students is small. There are many bright foreign students and young scientists in U.S. academic institutions. Excluding them due to ITAR restrictions makes this pool even smaller and will have serious long term consequences for the field. Many students trained on instrument calibration and data analysis for space hardware, choose to return to their countries, taking with them valuable experience, rather than facing potentially career-restricting regulations.

University programs (hardware, education). Because ITAR regulations extend over every aspect of space hardware, including engineering instruction and laboratory experiments, the confusion has bred over-cautiousness and led many university

[2] Space Science and the International Traffic in Arms Regulations: Summary of a Workshop, Space Sciences Board, 2008

departments to lower the complexity (and hence quality) of their courses, diminishing significantly the pool of capable students.

Mission Risk. The NRC Workshop refers to the failure of a U.S. instrument[3] on the JAXA *Suzaku* mission as an instructive example of the increased risk associated with overly restricted flow of information with foreign partners. It is actually amazing that such failures have been so rare. However, risk is going to increase as missions become more complex. It is certainly true, that U.S. hardware providers face increased quality assurance and management costs and oversight during the hardware development for international missions. They are also inclined to discourage cooperation with foreign partners who may be more familiar with the idiosyncrasies of their respective space agencies test and integration procedures and hence miss an opportunity for risk reduction or mitigation.

Cost and Administrative Burden. ITAR compliance requires trained staff and legal support. This is usually not too demanding for larger organizations such as NASA or federal labs. However, it taxes severely smaller university groups and is a considerable hurdle for independent hardware groups .

Current Status

The problems on space-based research arising from blanket ITAR enforcement were realized soon after the switch from EAR to ITAR regulatory regimes[4]. Every stakeholder, i.e., universities, private industry, federal research centers and organizations, including Congress, has been issuing calls for a resolution of this issue. It is clear that this problem is understood by the government. Both the Bush Administration in 2008[5] and the Obama Administration in 2009[6] have announced reviews of ITAR policies, but the State Department is procrastinating behind the flag 'national security' considerations.

Suggestions for the Committee's Consideration

While the government are working to resolve ITAR related issues that are detrimental to the US continuing to lead and advance fundamental research it is unrealistic to expect a wholesale change in ITAR solely for the benefit of fundamental research. However, specific amendments can and are being implemented in the ITAR code all the time. This indicates, at least to us, that pressure from the community could have results. ITAR could be tailored with sufficient precision to satisfy both the needs of the research community and of national security.

---

[3] *NASA Can Improve Its Mitigation of Risks Associated with International Agreements with Japan for Science Projects*, Audit Report No. IG-06-020, Office of the Inspector General, NASA, Washington, D.C., Sept. 12, 2006

[4] http://www.aip.org/fyi/2001/110.html

[5] http://www.state.gov/r/pa/prs/ps/2008/jan/99562.htm

[6] http://nasawatch.com/archives/2009/08/itar-review-announced.html

In general we advocate that the Heliophysics decadal survey make specific reference to the detrimental effect that ITAR regulations have on fundamental space research, instrument development and education.
Specific changes/recommendations for ITAR amendments can provide a constructive way forward. The NRC Workshop report[2] makes several such recommendations. We submit here a list of similar set of recommendations that are applicable to Heliophysics.

Proposed Actions
Return research satellites under the EAR of the Department of Commerce. This has been proposed in legislation currently pending before the House (H.R 2410, Sec 826)[7]. We propose the community support this initiative.

Proposed ITAR Amendments
1. Extend the *fundamental research* exception (ITAR 123.16(10)) to include research laboratories that undertake basic research with universities and to explicitly include fundamental research on a remote sensing satellite (e.g., SOHO, STEREO, etc). This has been proposed by the Defense Trade Advisory Group, a federal advisory committee to the State Department on ITAR issues. We propose the community support that initiative.
2. Clarification of regulations concerning international participation in science missions. This will prevent unwarranted restrictions imposed by institutional administrators and removed confusion that can lead to inaction and the missing of opportunities for U.S. science. A manual of 'do's and don'ts' for international space science collaborations would be most welcome.
3. Clarify and Harmonize Definitions between EAR and ITAR. This initiative is underway as part of the current export reform effort.

[7] http://www.govtrack.us/congress/bill.xpd?bill=h111-2410&tab=summary

Heliophysics White Paper

Coronal Magnetic Fields

Stephen White, Carl Henney (*AFRL Space Vehicles Directorate*)

Tim Bastian (*NRAO*)

Marc DeRosa (*LMSAL*)

Dale Gary (*NJIT*)

Phil Judge, Scott McIntosh (*HAO*)

Haosheng Lin (*IfA, University of Hawaii*)

Summary. After a long period in which we have recognized the overriding importance of coronal magnetic fields for many phenomena in the atmosphere of the Sun but have been powerless to make suitable detailed measurements, recent developments are now making it possible to address this topic in a more programmatic fashion. This paper outlines the difficulties we face in making advances in this field, and describes the directions in and means by which we can expect to make progress in the period to be addressed by the Heliophysics decade report. A combination of algorithmic development in extrapolation techniques, improved observations of vector magnetic fields in the chromosphere and corona, radio measurements of absolute magnetic field strengths in the corona, and continued EUV and X–ray observations of field line topology are needed *together* for further progress.

Introduction

It is universally recognized that magnetic fields in general, and coronal magnetic fields in particular, are of crucial importance in understanding the major phenomena currently at the forefront of solar physics research. Magnetic fields

- are the only plausible source for the energy released during solar flares;
- are almost certainly a key component in the heating of the solar chromosphere and corona, a process that we still do not understand and that has broad implications in the rest of astrophysics;
- probably control the acceleration mechanism that results in the nonthermal energetic electrons that produce solar hard X–rays, as well as the energetic protons that can affect space assets;
- and are clearly involved in the eruption of coronal mass ejections.

The interaction between magnetic fields and the dynamic plasma in the solar atmosphere is complex and multi–faceted. The mechanism for conversion of the free (i.e., non–potential) magnetic energy to other forms has not been identified and remains a topic of intense study. Magnetic fields alone cannot do work on a particle since they only exert forces orthogonal to particle motion, so electric fields (which can be produced by changing magnetic fields) must also be involved. Classic ohmic resistivity is much too slow to explain the dissipation of energy in flares. *Reconnection* of magnetic fields is the main contender for energy conversion in the corona: oppositely–directed fields can be driven towards each other, reconnect and release energy in the form of heating, plasma flows out of the reconnection region, and turbulence. Dissipation of currents that are an intrinsic part of the non–potential fields can also release magnetic energy in the form of heat.

All conversion mechanisms share a common property: due to the low resistivity of the corona, the spatial scale in which energy conversion takes place must be small for non–ideal–MHD behavior to permit energy conversion. In the solar corona this scale is of order tens of meters, much too small for the spatial resolution of current instruments. Once magnetic energy has been converted to particle energy, other mechanisms can take over to produce the nonthermal energy distributions that we measure via X–ray, γ–ray and radio observations. In the solar paradigm, these nonthermal particles transport energy from the corona to the chromosphere where it heats plasma to soft X–ray temperatures. The resulting hot plasma is responsible for the X–ray flares that we observe on the Sun with the GOES satellites, and presumably those on active stars that we observe with satellites such as Chandra and XMM.

Detailed understanding of the role of the coronal magnetic field in these physical processes requires better observations than we presently have. The magnetic field distribution at the solar surface has been well studied, since we have long been able to measure line–of–sight magnetic field components at the photosphere using circular polarization induced in spectral lines by the Zeeman effect. We are now even able to sense the presence of magnetic fields below the photosphere using helioseismic techniques (e.g., González Hernández et al., 2007; Komm et al., 2009), and incorporate flow patterns in studying the evolution and emergence of magnetic flux at the surface (e.g., Wang et al., 2009; Arge et al., 2010).

However, we are only too well aware that the surface line–of–sight magnetic fields measured by conventional magnetographs are not sufficient to tell us what is happening in the Sun's atmosphere

above the surface where, e.g., energy release by conversion of free magnetic energy into particle energy must be taking place since particle acceleration to nonthermal energies cannot happen in the collision–dominated photosphere. To make further progress a quantitative understanding of coronal and chromospheric magnetic fields, and the ways in which the atmospheric fields are coupled to the dynamic and turbulent motions below the surface, is needed. However, it is crucial to recognize the fundamental difference between studying the two–dimensional magnetic field at the (relatively) well–defined photospheric surface, and the intrinsically three–dimensional field of the corona. The amount of information needed to describe a three–dimensional field is obviously much greater, and the complications imposed by having to take into account line–of–sight depths (i.e., the fact that material at any point along a given line of sight can contribute to an optically thin diagnostic, which is not an issue for photospheric measurements) pose a major headache in deconvolving raw data into magnetic field models. This simple consideration makes it clear that we cannot expect a quick resolution of this issue and need to be realistic in the problems that we tackle in the near term.

In this paper we describe the current state of studies of coronal magnetic fields, starting with the topic of extrapolation of surface field measurements that has been the primary source of coronal magnetic field models for research purposes, and then go on to describe examples of new instrumentation that will contribute important results to this field and can be constructed over the next decade.

Extrapolations of Photospheric Magnetic Field Measurements

Extrapolation of photospheric magnetic fields into the corona can be carried out as a boundary value problem, but in order to do more than a potential–field extrapolation (which has no free energy available for conversion, and therefore is not very useful for any of the physical problems listed earlier) we require, at least, a complete measurement of the vector magnetic field **B**. Recognition of the importance of coronal magnetic fields has led to major investments in two directions: (i) instruments capable of measuring **the full vector magnetic field at the photosphere**, and (ii) efforts to develop algorithms for solving the magnetic boundary–value problem in the solar atmosphere under the constraint of **nonlinear force–free** (**J** $\parallel$ **B**) conditions applicable to the quasi–static (i.e., slowly varying) corona. Determination of the full magnetic vector at the photosphere requires spectropolarimetric measurements and sophisticated modelling of the lower atmosphere, including motions, and is in itself a difficult problem even with suitable data. Once we have suitable magnetic data, the system of Maxwell's equations plus a boundary measurement set then constitutes a mixed elliptic-hyperbolic boundary value problem that has proven remarkably difficult to solve (Gary, 1989; McClymont & Mikić, 1994; Amari et al., 1999). There has been a concerted effort to develop robust nonlinear force–free–field extrapolation techniques, summarized initially by Schrijver et al. (2006) and more recently by Schrijver et al. (2008) and DeRosa et al. (2009). One of the problems with these extrapolations is that the photospheric magnetograms are not force–free and further the results are very sensitive to errors in the input magnetograms, and to imbalance in positive and negative flux. This has led to the need to "precondition" the magnetograms, i.e., to alter them in order to force them to match the conditions required for successful extrapolations (for a description of this process see Wiegelmann et al., 2006; DeRosa et al., 2009). Thus in parallel with the extrapolation efforts, a series of workshops is underway to compare vector magnetograms from different instruments and to reconcile differences between the measurements. In the long run global extrapolations are desirable,

but these are not really feasible unless we are able to measure the vector magnetic field over the entire photosphere, not just the hemisphere visible from Earth, and there is no prospect for this in the next decade.

A major issue for extrapolations is that the magnetic field in the photosphere and lower chromosphere carries less energy than the plasma and therefore is not force–free: the plasma can sustain forces and currents orthogonal to the magnetic field. This complication leads to the desire to use measurements of the vector field in the force–free upper chromosphere instead: this topic is addressed in more detail in a White Paper by Judge (2010). As Judge's discussion stresses, the surface in which any particular magnetically–sensitive line will form varies in time and space, and in particular the height will be different along different lines of sight. The practical use of the chromospheric magnetograms as boundary conditions for extrapolations is currently being explored. One advantage of the chromospheric measurements is that Chandrasekhar's virial theorem can in principle be applied, with no need for ill-posed extrapolations, to derive the total free magnetic energy in the overlying coronal volume.

Yet another major problem for developing reliable extrapolation algorithms is the lack of suitable measurements to compare with. Idealized models with exact solutions provide a test, but they cannot incorporate all the complications of the real solar atmosphere (Barnes et al., 2006). A standard approach is to compare field lines derived from extrapolations with field lines observed in EUV and/or X–ray observations of the corona. However, the mathematical rigor of this test appears never to have been established: the field lines visible in EUV images are a measure–zero subset of all the field lines in the volume (those that happen to carry excess density). Given the limited spatial resolution of the observations and the uncertainty in exact footpoint locations of field lines, the modeller is free to choose extrapolated field lines in the vicinity that best match those observed. Unfortunately, in the absence of better data, this is the only observational test that is routinely available. Nevertheless, more stringent and bias-free tests of extrapolations using actual measurements of magnetic fields in the corona have been demonstrated, for example, by dedicated programs using Very Large Array radio measurements of coronal magnetic field strengths at the base of the corona (e.g., Lee et al., 1998a) and using IR spectropolarimetric measurements of near-IR coronal emission line (e.g., Liu & Lin, 2008). With the establishment of the suite of facilities that we advocate in this paper these observations will become routinely available for vigorous observational tests of competing coronal magnetic field extrapolation tools and models.

Another approach to testing of extrapolations that avoids the issues of field–line selection is to generate synthetic images in EUV or X–ray bands, for comparison with observations, based on self–consistent determination of extrapolated coronal fields and emission measure distributions using a heating model, as in the calculations of Mok et al. (2005). However, robust coronal heating models have yet to be identified and this makes comparisons more a test of the heating mechanism than of the magnetic field model, but this approach may prove valuable in the future.

Resources

As noted earlier, the challenge of characterizing the magnetic field in the chromosphere and corona is being addressed by a combination of observational advances and improved numerical techniques. Observations of photospheric fields are not very useful for further progress unless they measure the full three–dimensional vector field, and even then they suffer from not being force–free. The

position endorsed by this paper is that a combination of a number of techniques is required in order to measure properties of magnetic fields in the corona that will help to address the questions raised in the Introduction. Results from the following methods are needed:

- Extrapolations of surface measurements are discussed in detail above. Chromospheric magnetic fields measured in a force–free region of the atmosphere are preferred over the non–force–free photospheric fields as the boundary condition. Testing and validating the extrapolation algorithms requires the actual measurements described in the following items as constraints.

- Polarization measurements of optically-thin magnetically-sensitive coronal lines such as forbidden Fe XIII near 10750 Å. Such lines cannot be observed polarimetrically in the presence of the bright continuum from the solar disk and therefore magnetic field measurements are limited to regions above the solar limb in projection, with the attendant line–of–sight issues. Nevertheless, the full polarization state can be measured and a two-dimensional image of the vector field in the corona in the plane of the sky containing sun center can be approximated from the data. The power of these measurements is demonstrated by the direct detection of Alfvénic waves in the corona (Tomczyk et al., 2007).

- Radio measurements are currently the only means by which magnetic field strengths in the corona can be measured against the solar disk. For this reason they have a unique role to play in constraining extrapolations, which invariably use disk regions where we can measure the surface fields, in addition to the science derived from coronal field strength measurements. An example of such science is searching for changes in coronal field strength produced by flares: changes associated with flares are seen in photospheric magnetic patterns, but the coronal changes should be more dramatic, if only because of the motion of field lines seen in eruptive events. Magnetograms of the absolute field strength at the base of the corona above active regions, in the range 200–2000 G, are straightforwardly obtained using the radio gyroresonance technique (White & Kundu, 1997). The gyroresonance data also contain information on the three–dimensional field but significant algorithm development is required to exploit this fully: the radio data at each frequency sample a narrow optically-thick constant-field-strength layer, but the height of each layer is not independently measured. Magnetic field measurements at lower field strengths (< 200 G) can also be made using the polarization of bremsstrahlung (Gelfreikh, 2004) and mode coupling observations of polarization reversals (Lee et al., 1998b).

- EUV and soft X–ray images are now available at very high spatial resolution and show those magnetic field lines that carry significant density, but due to the density–squared dependence of the emissivity at these wavelengths, which enhances contrast in density, we cannot "see" the bulk of the field lines since they do not carry sufficient density.

These four techniques complement each other. We argue that all four are needed simultaneously if we are to make significant progress in understanding coronal magnetic fields.

Finding: *spectropolarimetric measurements of optically thin coronal and chromospheric lines, broadband radio measurements of optically thick isogauss layers, and field–line images are all needed in combination to constrain and lead to improved extrapolation techniques.*

Existing and near–term facilities

Measurements of the vector magnetic field are now being carried out by numerous instruments, including the following:

- Ground–based telescopes that can make routine photospheric and chromospheric vector magnetogram observations with arcsecond resolution include the Facility Infrared Spectropolarimeter (FIRS) and the Spectro-Polarimeter for Infrared and Optical Regions (SPINOR) at Sacramento Peak, the Imaging Vector Magnetograph at Mees Solar Observatory, the MSFC Magnetograph now at the University of Alabama/Huntsville, and the Vector SpectroMagnetograph of the SOLIS project, together with other facilities overseas. An increasing number of instruments are able to measure the vector magnetic field on the solar disk in the preferred chromospheric lines (Judge, 2010; McIntosh et al., 2010).

- The Solar Optical Telescope on the Japan–U.S. satellite *Hinode* carries a Stokes Polarimeter that can make vector magnetograms over a field of $300''$ with a pixel size of $0.16''$ (110 km at the Sun) every 90 minutes.

- The Helioseismic and Magnetic Imager on the Solar Dynamics Observatory takes data full–disk vector magnetograms at high cadence with $1''$ resolution. However, as with all spectropolarimetric data, conversion of these observations to vector magnetograms is not a simple procedure and is still being developed for HMI.

- The HAO Coronal Multi-channel Polarimeter (CoMP), now deployed at Mauna Loa, combines a polarimeter and tunable filter to measure the complete polarization state of the Fe XIII and He I lines near 10800 Å in the solar atmosphere above the limb. CoMP has demonstrated the presence of Alfvénic waves in the corona (Tomczyk et al., 2007).

- The University of Hawaii SOLARC instrument is an 0.5m off–axis coronagraph equipped with a fiber-optic imaging spectropolarimeter located at Haleakala. SOLARC produced the first "Coronal Vector Magnetogram" from spectropolarimetric observations of the Fe XIII line (Lin et al., 2004).

- The 1m Sunrise telescope (operated by a large international consortium led by the Max Planck Institute for Solar System Research) carries a spectropolarimeter that can achieve very high spatial resolution measurements of the vector magnetic field during balloon flights.

Future facilities needed

Proposed facilities that will be important for advancing this field include:

- The Advanced Technology Solar Telescope (ATST) has now been approved for funding in the NSF Major Research Equipment and Facilities Construction line. It is a 4–meter optical/IR telescope optimized for low-scattering coronagraph observations with resolution of order $0.2''$ and field–of–view $300''$, and will be able to acquire high–resolution vector magnetograph observations above the limb in coronal lines.

- The COronal Solar Magnetism Observatory (COSMO) facility consists of a 1.5 m aperture coronagraph and a 20 cm aperture telescope to take regular synoptic spectropolarimetric measurements in chromospheric lines (He I, Hα) on the disk and coronal lines (Fe XIII 1074.7Å and 1079.8Å, He I 1083.0Å) above the limb with a cadence of order 10 minutes. COSMO will observe a one–degree field of view with a resolution of 5 arcseconds and a sensitivity of a few Gauss. In order to minimize scattered light from the solar disk, the COSMO coronagraph will use a fused–silica lens rather than a mirror. COSMO's wide–field synoptic measurements of chromospheric and coronal magnetic fields will complement user-designed observations with the narrower but higher-resolution field of ATST. More details are given in the White Paper by Tomczyk et al. (2010).

- The Frequency Agile Solar Radiotelescope (FASR) will exploit radio techniques to measure magnetic field strengths in the corona above active regions. FASR will be a full–disk solar–dedicated instrument with spatial resolution scaling with frequency, and hence measured magnetic field strength, continuously from $1''$ at 2000 G to $10''$ at 200 G. Routine measurement of coronal magnetic fields is a prime scientific goal of FASR, which is described in more detail in a White Paper by Gary et al. (2010). The solar radio array at Owens Valley operated by NJIT is currently undergoing an expansion that will prototype some technical aspects of FASR.

- The Expanded Very Large Array (EVLA) is currently being instrumented and, when complete, will be a valuable tool combining a wide frequency range (equivalent to sensitivity to a wide range of magnetic field strengths) with excellent spatial resolution and sensitivity. However, it is not solar–dedicated and suitable coronal magnetic field measurements are only likely to be obtained on a few days per year, at most.

Finding: *These facilities in combination with the ongoing algorithm development and vector magnetogram reconciliation efforts described earlier will provide a wealth of detail on coronal magnetic fields and permit us to move forward in understanding the detailed role of magnetic fields in atmospheric phenomena.*

The importance of coronal magnetic fields in solar phenomena that have relevance well beyond the Sun itself, in addition to their impact on the Earth, demands that research in this field should continue to be a high priority over the next decade.

References

Amari, T., Boulmezaoud, T. Z., & Mikic, Z. 1999, Astron. Astrophys., 350, 1051

Arge, C. N., Henney, C. J., Koller, J., Compeau, C. R., Young, S., MacKenzie, D., Fay, A., & Harvey, J. W. 2010, Twelfth International Solar Wind Conference, 1216, 343

Barnes, G., Leka, K. D., & Wheatland, M. S. 2006, Astrophys. J., 641, 1188

DeRosa, M. L., et al. 2009, Astrophys. J., 696, 1780

Gary, G. A. 1989, Astrophys. J. Supp., 69, 323

Gary, D. E., and the FASR Team, Heliophysics White Paper entitled “The Frequency Agile Solar Radiotelescope”.

Gelfreikh, G. B. 2004, in ”Solar and Space Weather Radiophysics”, eds. D. E. Gary and C. U. Keller, Kluwer Academic Publishers, p. 115

González Hernández, I., Hill, F., & Lindsey, C. 2007, Astrophys. J., 669, 1382

Judge, P. G., Heliophysics White Paper entitled “Measuring magnetic free energy in the solar atmosphere”.

Komm, R., Howe, R., & Hill, F. 2009, Solar Phys., 258, 13

Lee, J., McClymont, A. N., Mikić, Z., White, S. M., and Kundu, M. R. 1998a, Astrophys. J., 501, 853

Lee, J., White, S. M., Kundu, M. R., McClymont, A. N., and Mikić, Z. 1998b, Solar Phys. 180, 193

Lin, H., Kuhn, J. R., & Coulter, R. 2004, Astrophys. J. Lett., 613, L177

Liu, Y., & Lin, H. 2008, arXiv:0809.0729

McClymont, A. N. & Mikić, Z. 1994, Astrophys. J., 422, 899

McIntosh, S. W., et al, Heliophysics White Paper entitled “The Solar Chromosphere: The Inner Frontier of the Heliospheric System”.

Mok, Y., Mikić, Z., Lionello, R., & Linker, J. A. 2005, Astrophys. J., 621, 1098

Schrijver, C. J., Derosa, M. L., Metcalf, T. R., Liu, Y., McTiernan, J., Régnier, S., Valori, G., Wheatland, M. S., & Wiegelmann, T. 2006, Solar Phys., 235, 161

Schrijver, C. J., et al. 2008, Astrophys. J., 675, 1637

Tomczyk, S., et al, Heliophysics White Paper entitled “COSMO: The COronal SOlar Magnetism Observatory”

Tomczyk, S., McIntosh, S. W., Keil, S. L., Judge, P. G., Schad, T., Seeley, D. H., & Edmondson, J. 2007, Science, 317, 1192

Wang, Y.-M., Robbrecht, E., & Sheeley, N. R. 2009, Astrophys. J., 707, 1372

White, S. M. & Kundu, M. R. 1997, Solar Phys., 174, 31

Wiegelmann, T., Inhester, B., & Sakurai, T. 2006, Solar Phys., 233, 215

The Need for Explicit Basic and Applied Research Funding

T. G. Onsager
NOAA Space Weather Prediction Center

Recommendation:
The Decadal Survey Committee should advocate for distinct lines of funding by the U.S. civilian agencies for: 1. basic space physics research and 2. the development of space weather applications, maintaining distinct requirements for both. Under the current paradigm guiding research funding today, namely to favor research with societal benefit, neither basic research nor applications with value to society are being supported optimally.

Discussion:
Our Nation needs separate and explicit lines of funding for basic space physics research and for space weather applications. Because of the disparate nature of both the scientific effort and the time frames over which outcomes are expected for basic research and the development of applications, both activities suffer when funded under a single umbrella. Basic physics research is suffering from a perceived pressure on scientists to justify research in terms of space weather applications, and space weather applications with demonstrated utility are not being developed in a timely manner. Of course there is substantial overlap between basic research and applications, but there are important research and application areas in need of support that reside outside the overlap and consequently are being neglected.

Following the National Academy report on societal and economic impacts of space weather (2008), a number of government agencies around the world (including the Federal Emergency Management Agency and the Department of Homeland Security in the U.S.) have begun to take the threat of space weather on our electric power and communication/navigation infrastructures seriously. As our community has been saying since the inception of the U.S. National Space Weather Program, there are important national needs that must be met, including the dramatic improvement of space weather forecasts and services.

With a few exceptions, models that provide specific space weather forecasts have improved little over the past decade. More importantly, it has not been demonstrated that the improvements that have been made are sufficient to provide value to users who need forecasts and services, such as the electric power and navigation industries. While new knowledge has been gained from the excellent missions and research supported over the past decade, this knowledge has not been translated into capabilities with societal value.

This issue is not part of the often-described research-to-operations challenge; it precedes it. Given the lack of models with proven value, the U.S. space weather effort is not yet to the point in its development where the transition of research models to operations is the dominant issue. We must first develop forecast models and demonstrate their value to society before expense to employ the capabilities in operations can be justified.

Conducting basic space physics research and developing space weather applications have key differences, both in terms of the desired outcome and the time frame over which an outcome will

be achieved. Basic space physics research must be conducted with the goal of expanding knowledge, not creating something. The more our research is guided by a need to enable space weather capabilities, the less support there may be for basic research with unanticipated benefits. This sentiment was expressed at a recent meeting of the Decadal Survey Theory and Modeling Working Group (October 14-15, Boulder, CO). It was suggested that the recent emphasis on space weather has had a negative impact on basic space physics research. The perception that space physics research should have an identifiable tie to societal benefits may inhibit scientists from submitting proposals on basic research topics, and it may be detrimental to their proposals being supported. Interestingly, it was also pointed out at this meeting in a presentation by a different scientist that in spite of the focus of funding on space weather, few models today are capable of accurate forecasts. This combination of views from prominent scientists illustrates the gaps caused at both ends of the spectrum by our current approach to research funding: basic research funding has become difficult to obtain, and our stated goal of accurate forecast capabilities is not being achieved.

Just as basic research must be supported, research is also needed to create high-value applications that will address our near-term space weather needs. This effort is not engineering, but rather it involves focused scientific research to translate existing knowledge into the best possible applications. Similar to the parameterization of sub-grid-scale physics in today's weather models, sophisticated techniques need to be developed to optimize our data and models for space weather forecasts.

Unlike basic research, space weather research must be outcome-based with near-term goals. It has been demonstrated over the past decade that improved forecasting capabilities do not result automatically from research targeted toward space weather topics. NASA's Living With a Star Program and NSF's Space Weather funding are the civilian programs closest to supporting the development of space weather applications, and they have enabled excellent research on the science underlying space weather. These programs do not directly support the development of applications, but rather support the acquisition of knowledge needed for the eventual development of space weather applications. The fact that few forecast models with demonstrated societal value have been developed over the past decade indicates that while the acquisition of new knowledge is necessary for improved applications, it is not sufficient for the applications to be developed. Unless explicit funding is applied with clear requirements for near-term outcomes, applications with predictive value to mitigate space weather impacts will not be developed in a timely manner. We need targeted innovations and improvements in our capabilities to occur over the next three to five years, not in 10 to 20 years.

We must recognize that although there is overlap in knowledge content, the development of a basic physics understanding requires separate funding and distinct goals from the development of space weather applications. Funding for applications must be outcome based. The emphasis must be on near-term solutions to key problems and on the quantification of our capabilities and their limitations. Our forecast capabilities have not improved dramatically over the past decade, and they will not improve dramatically over the next decade unless it is made an explicit priority of dedicated funding. Our Nation needs explicit funding both for basic space physics research and for space weather forecasting.

Magnetic Properties of the Solar Atmosphere

(SolmeX Cosmic Vision Mission)

2010 NRC Decadal Survey White Paper

Submitted by J.D. Moses for the SolmeX Consortium

The structure and dynamics of the solar atmosphere is dominated by the magnetic field. Our current lack of measurements of the coronal magnetic field is a major drawback to advance our understanding of the solar atmosphere. It is now technologically feasible to overcome this observational gap and obtain comprehensive set of measurements of the magnetic field from the chromosphere to the outer corona through remote-sensing techniques over a multitude of spatial and temporal scales.

For each of the two ESA Cosmic Vision program calls for mission concepts, there has been a proposal submitted to achieve the first comprehensive set of measurements of the magnetic field in the upper atmosphere through spectro-polarimetic techniques. In order to provide to the NRC Decadal Survey a well-developed mission concept, this white paper will summarize the mission proposal being submitted to the 2010 ESA Cosmic Vision call entitled “Solar Magnetism Explorer – SolmeX”.

The specific science addressed by such an investigation includes the fundamental issues:

1. What determines the magnetic structure of the solar upper atmosphere?
2. What is the nature of the changes of the magnetic field over the solar cycle?
3. What drives large-scale coronal disruptions such as flares and coronal mass ejections?
4. How do magnetic processes drive the dynamics and heating of the outer atmosphere?
5. How is the magnetic field coupling the whole solar atmosphere, from the photosphere to the outer corona?

1 Structure of the coronal magnetic field

i. What is the magnetic structure of the corona on large scales?

During solar minimum the large-scale corona appears dipole-like. The polar coronal holes show a super-radial expansion, i.e., they expand faster than radial rays roughly following the edges of the helmet streamers at lower latitudes. However, in-situ observations suggest a more radial expansion, and radial rays are visible seemingly poking through the boundary between coronal hole and the quiet corona (Woo et al. 1999).

Coronagraphic observations of the magnetic field of the large-scale corona (e.g. of the edges of coronal holes and of helmet streamers) are necessary to understand the connection between the corona and solar wind.

ii. What is the internal magnetic structure of coronal loops?

Imaging observations show in great detail loop reaching high into the corona, with heights of 10% of the solar radius or more. It is generally accepted that these outline magnetic field lines which are loaded with hot plasma as a consequence of the coronal heating process. However, while the coronal loop cross sections appear constant in

observations, they should expand upwards according to the decrease of magnetic field strength with height. Furthermore the internal magnetic structure of coronal loops is unknown, e.g., their twist is an open issue.

These issues can be addressed by simultaneous measurement of the direction of magnetic field in a coronal loop and the structure of the coronal loop in EUV and visible light. This will provide results on the alignment of the observed structures with the magnetic field.

iii. What is the magnetic structure of the quiet Sun network?
On smaller scales of the order of 1Mm in the chromosphere and the transition region to the corona the Sun is highly structures and very dynamic. A multitude of small-scale short-lived phenomena can be seen, such as spicules, fibrils, explosive events, etc.

Spicules have been extensively studied with Hinode in chromospheric emission of Ca II. Indications of Alfven waves have been found to travel along the Spicules (De Pontieu et al. 2007). These conclusions are based on intensity images alone, and no reliable measurements of the magnetic field in spicules exist. Extending to somewhat higher temperatures, there are small transition region loops found crossing the chromospheric network. These are puzzling because there seems to be no connection from these small network loops to the underlying photosphere (Feldman et al. 2001).

To address these questions one must measure the magnetic field in lines formed in the chromosphere and the transition region. Measurements of the magnetic field the MgII k (279.6 nm) and in CIV (154.8 nm) provide magnetic information on spicules on the disk and above the limb as well as for small transition region loops above the chromospheric network.

2 The corona within the solar cycle

i. How does the magnetic field change in the corona on a global scale?

It has been known from eclipse images that the appearance of the corona changes dramatically within the 11-year solar cycle, i.e., with the cyclic change of the sunspot number. While having a dipole-like structure during minimum activity, the corona becomes very complex during maximum activity manifest through the numerous helmet streamers distributed all over the Sun.

During solar minimum coronal holes appear mainly near the poles, while during maximum activity smaller coronal holes appear at lower latitudes. These leave different imprints in the solar wind which are not understood. Coronagraphic polarimetric measurements of large-scale coronal structures over the solar cycle will reveal the differences in magnetic structure between minimum and maximum conditions through long-term observation over at least half a solar cycle, i.e., 6 years.

3 Large coronal disruptions

i. How is the magnetic field changing during large disruptions such as flares and CMEs?

In large disruptions of the corona, huge amounts of magnetic energy are converted into thermal energy and kinetic energy. The associated heating and acceleration process through magnetic reconnection are still poorly understood, and even the change of the magnetic structures is not well constrained. Currently extrapolations of the magnetic field from the surface level are used to investigate the change of the free energy of the magnetic field. However, during dynamic events the assumptions of these

extrapolations break down. Thus there is dire need for measurements of the coronal magnetic field in the course of the disruption. Simultaneous measurements of the magnetic field in the chromosphere, the transition region and the corona itself are necessary to address this problem. This provides vital data to constrain models and extrapolations of the magnetic field during the disruption and will identify regions of particular high potential for magnetic reconnection.

ii. What is the magnetic structure of a CME during an eruption?

Once a coronal mass ejection (CME) lifts of, it expands and through this some CMEs reveal their internal structure. Our current information is based on intensity images and this no information on the magnetic field is available.

This is unfortunate as different models for CME eruption make detailed predictions on their internal magnetic structure. As these CMEs often lift off with high speeds of the order of 1000 km/s, they cross a field-of-view of one solar radius in 10 minutes, which sets minimum requirements for the observations.

High-cadence infrared coronagraphic spectro-polarimetry can provide both the intensity structure of CMEs and the magnetic field vector. By this the different CME models can be tested and the physics of the CME eruption can be resolved.

4 Magnetic driving of coronal dynamics

i. How does the plasma respond to heating in current sheets?

In general the magnetic field will not be irrotational, which is the same as to say that currents are present, which are often organized in thin sheets. In-situ observations show these directly, e.g., the heliospheric current sheet, but they can also be found on small scales of 1Mm in the low atmosphere of the Sun (Solanki et al. 2003). According to theory they should exist on all scales, most prominently during reconnection events, when magnetic energy is converted to thermal and kinetic energy. While extensively studied theoretically, there are virtually no observations of current sheets in the upper solar atmosphere, simply because of the lack of magnetic field measurements. Many processes are believed to be based on current sheet reconnection, e.g. explosive events (Innes et al. 1997), but the link to the magnetic field structure is missing.

Spectro-polarimetric observations on all accessible scales combined with the measurement of Doppler shifts and changes in observed structures will determine the link between current dissipation during reconnection and the response of the plasma.

ii. Is wave action or field line braiding the dominant heating agent?

Certainly the driver of coronal dynamics and heating is to be found in the solar photosphere. However the agent to transport the energy out into the corona, against a temperature gradient, is not yet identified.

The convective motions of solar granulation shuffle around the magnetic field and through this will induce currents in the corona which are dissipated. This field line braiding (Parker 1972) indeed produces a hot corona in numerical models (Gudiksen & Nordlund 2002) which are consistent with observed spectral properties (Peter et al. 2004). Depending on the speed of the driver, the same process could also induce waves, which propagate upwards and then heat the corona through their dissipation. In order to distinguish between these processes, polarimetric observations in the chromosphere, transition region and corona with a cadence faster than the granular convective time (5 min) and sub-granular resolution (order 1”) are required.

iii. Which are the source regions of the wind?

Globally, fast solar wind streams are known to originate from large coronal holes. However, their detailed source regions in the lower corona and magnetic network remain poorly determined. In particular this is due to the lack of direct measurements of the coronal magnetic field in the chromosphere and the transition region to the corona, where the wind is initially accelerated (Tu et al. 2005). Probably all types of solar wind streams might ultimately originate through the interaction of coronal loops and open field regions, in which case the loops should open up intermittently to release plasma into the corona and solar wind. The combination of imaging of coronal plasma and spectro-polarimetric measurements of the magnetic field is required to determine the magnetic structure of the solar wind source region.

iv. What drives the wind in the upper corona?

Fast solar wind streams are permeated and dominated by large-amplitude Alfven waves, which may be generated by foot-point motions in the photosphere and magnetic reconnection in the chromospheric network. In the upper corona these waves can interact with the plasma, e.g., through the ion-cyclotron resonance, which leads to an asymmetric velocity distribution of the ions (Kohl et al. 1998).

Measurement of the magnetic field in the upper corona (where the temperature anisotropies are observed) will identify the gyration frequencies as a function of distance and determine the heating and acceleration of the wind on larger scales. This simultaneous spectroscopic analysis of the wind speed, e.g. through Doppler dimming, and of the magnetic field through EUV and IR spectro-polarimetry will constrain the models which are currently built on numerous assumptions.

5 Magnetic coupling through atmosphere

i. How does the atmosphere couple across the plasma beta=1 layer?

The plasma beta is the ratio of the thermal energy density of the gas to the energy density of the magnetic field, and it happens that beta is about unity somewhere in the upper chromosphere. Below, where beta>1, the plasma dominates the magnetic field and can push the field lines around. The situation is opposite in the corona, where beta<1.

Thus the upper chromosphere is of particular interest for plasma physics. Waves can convert from one mode to another resulting in different oscillatory behavior below and above beta~1 (e.g., McIntosh & Judge 2001). It is also unclear how the chromosphere and the transition region to the corona are magnetically connected through this region (Wikstøl et al. 1998). Spicules are among the most prominent features related to the beta~1 region still awaiting a conclusive explanation (De Pontieu et al. 2007). Simultaneous spectro-polarimetric observations of the magnetic field in the chromosphere and transition region will resolve the issue of magnetic coupling in this region.

ii. How are coronal structures rooted to the surface?

Certainly all structures in the corona have to be rooted on the solar surface. As soon as a structure detaches from the solar surface, it lifts off - as impressively seen during CMEs. Because of the exponential drop of the gas pressure with height the magnetic structures have to expand and ultimately fill the whole volume of the corona, while on the surface they are concentrated in small regions. As only a small fraction of the

magnetic flux in the photosphere is connected to the corona, the important question is which parts of the surface are connected to which structures in the corona and how the coronal structures are driven from below. Answering these questions requires measurement of the magnetic field not only in the photosphere, but also in the chromosphere and low corona in order to guide and constrain models and extrapolations of the magnetic field into the corona. This involves measurements of the magnetic field on all scales, from small magnetic patches on the surface to large-scale structures observed with coronagraphs.

SolmeX Observational Strategy

SolmeX instruments will measure the polarization of light from the chromosphere and the corona, on-disk as well as off-limb, from space. Thus they will provide a comprehensive set of measurements of coronal magnetic fields. The instrument suite comprises imaging as well as spectroscopic devices for coronagraphic observations above the limb as well as for measurements on the disk, namely:

Off-limb	(a)	VIRCOR (Visible light and IR coronagraph),
	(b)	CUSP (Coronal UV spectro-polarimeter),
On-disk	(c)	EIP (EUV imaging polarimeter),
	(d)	SUSP (Scanning UV spectro-polarimeter),
	(e)	ChroME (Chromospere magnetic explorer).

Magnetic field diagnostics: Zeeman effect and Hanle effect

The Zeeman effect is a manifestation of the interaction of the magnetic moment of the atom and the magnetic field. In the presence of a magnetic field spectral lines split into components with circular and linear polarization, depending on the angle between the line-of-sight and the magnetic field vector. The Zeeman effect is very prominent in the infrared (IR) as the splitting of the components increases with the wavelength squared.. Early attempts to measure the polarization in the transition region line CIV at 154.8 nm gave no conclusive results (UVSP on the Solar Maximum Mission launched in 1980; cf. Lites 2001). The most recent attempt with the SUMI rocket (Summer 2010) unfortunately had technical problems.

The *Hanle effect* modifies the linear polarization in the presence of a magnetic field. Excitation through anisotropic radiation can induce linear polarization of atomic levels and by this of the resulting emission. Quantum coherence causes the magnetic field to modify this linear polarization and by this imprints information on its strength and orientation. Depending on the scattering geometry, i.e., 90 degree scattering above limb or forward scattering in front of the disk, the modification consists in a reduction or enhancement of the linear polarization amplitude and in a rotation of the direction of linear polarization. At higher field strengths the Hanle effect saturates and is sensitive to the field orientation only.

As the Hanle effect is sensitive to weaker magnetic fields than the Zeeman effect, the diagnostics of these two effects is complementary

The SolmeX Science, Instrument and Mission design are presented in the following tables:

observational technique / science goals	(a)	(b)	(c)	(d)	(e)
	off-limb		*on-disk*		
	coronagraphic UV spectro-polarimetry	coronagraphic visible & IR polarimetry	EUV polarimetric imaging	UV slit spectro-polarimety	UV imaging spectro-polarimetry
Sect. 3.1.1 Structure of the coronal magnetic field	(i) What is the magnetic structure of the corona on large scales?			(iii) What is the magnetic structure of the quiet Sun network?	
		(ii) What is the magnetic structure of coronal loops?			
Sect. 3.1.2 The corona within the solar cycle	(i) How does the magnetic field change on a global scale?				
	(ii) Does the magnetic coupling through the atmosphere change through the cycle?				
Sect. 3.1.3 Large coronal disruptions		(ii) What is the magnetic structure of a CME?	(i) How is the magnetic field changing during large disruptions such as flares and CMEs?		
			(iii) How do changes on large scales relate to small scales during disruptions?		
Sect. 3.1.4 Magnetic driving of coronal dynamics	(i) How does the plasma respond to heating in current sheets?				
	(iv) What drives the wind in the upper corona?		(ii) Is wave action or field line braiding the dominant heating agent?		
			(iii) Which are the source regions of the wind?		
Sect. 3.1.5 Magnetic coupling through the atmosphere				(i) How does the atmosphere couple across the β=1 layer	
	(ii) How are coronal structures rooted to the surface?				
prime target structures	large-scale corona above the limb	off-limb corona above active regions	coronal structures on the disk	structures of the low corona on the disk	chromospheric magnetic structures
spectral lines or band	Ly-α, β, γ O IV (103 nm)	Fe XIII 1.07 μm vis:† ~ 400 nm	Fe X (17.4 nm)	120 – 150 nm incl, Ly-α, C IV	Mg II (279 nm) Fe I (525 nm)
field of view	> 1 R_{Sun}	> 2 R_{Sun}	full disk	300"	300"
spatial resolution	5"	IR: 5" visible: 2"	1" (imaging) 3" (polarimetry)	2"	0.3"
image sequence *	?? s (n-p) 30 min (pol)	30 s	10 s (n-p) 10 min (pol)	1 s (n-p) 3 min (pol)	5 s
velocities					
magnetic field					
measurement requirements	CUSP	VIRCOR	EIP	SUSP	ChroME

† visible: for observations of the K-corona * image sequence: (n-p): non-polarimetric, (pol): polarimetric mode

spectral lines or band	Ly-α, β, γ O IV (103 nm)	Fe XIII 1.07 μm vis: ~ 400 nm	Fe X (17.4 nm)	120nm–150 nm incl, Ly-α, C IV	Mg II (279 nm) Fe I (525 nm)

data rate	150	300	550	300	700
aperture	25 x 30 cm^2	⌀ 20 cm	⌀ 28 cm	15 x 10 cm^2	⌀ 25 cm
dimension cm^3	180 x 60 x 30	180 x 50 x 25	100 x 30 x 30	160 x 50 x 40	150 x 45 x 50
mass kg	70	60	40	68	52
power W	30	50	50	25	55

Mission Design

Launch Vehicle	Soyuz-Fregat 2b (TBC)
Orbit	Large-amplitude Halo orbit at Sun-Earth Lagrange Point 1 (L1)
Satellite System	
Satellite System Concept	Two-spacecraft formation, *Coronagraph Spacecraft* and *Occulter Spacecraft* at 100-m separation
Attitude Stabilization	3-axis stabilized, sun-pointing
Attitude Control	RPE of Line of Sight < 1 arcsec in 15 minutes
Payload Telemetry Rate, Data Volume	900 kbit/s, 78 Gbit/day
Mass Memory	100 Gbyte
Telecommunications	X-band, 3 Mbit/s telemetry rate, 8 hours/day
Power/Energy	1 kW, no eclipse
Propulsion	Hydrazine (AOC), Cold-gas (formation keeping)
Mass Budget	1400 kg with adapter, >> 20% margin wrt launcher capacity, including 260 kg payload mass with 20% margin
Ground System	35 m ground antenna, X band receive/transmit, 8 hours/day
Technology Developments	Formation keeping metrology (RF and optical): TRL = 3 (today), TRL = 9 (after Proba 3) Proportional thrusters: TRL = 4 (today), TRL = 9 (after GAIA)

The cost estimate for SolmeX is still under development at the time these NRC white papers are due. The cost estimate of the predecessor to SolmeX, COMPASS (Coronal Magnetism, Plasma and Activity Studies from Space), are provided for reference.

COMPASS Cost Estimate (2008)

ROM Cost estimate	**%**	**M€.**
Pre-Implement. phase	2%	6.0
Industrial activities	50%	150.0
Launch services	13%	40.0
Ground Segment (MOC+SOC)	12%	36.0
ESA Internal Costs	11%	33.0
Contingency	11%	35.0
Cost to ESA	**100%**	**300.0**
Payload (20% margin))		300.0
Cost at Completion		**600.0**

Fast UV Spectro-imagery for Solar Physics

J.-C. Vial
Institut d'Astrophysique Spatiale, Orsay, France

The needs of UV solar physics in terms of fast spectro-imagery

In order to diagnose the plasma in the solar atmosphere, in 3 spatial dimensions and time, physical quantities such as temperature, densities, ionization degree, abundance, along with magnetic and velocity fields must be known. In the remote-sensing approach, these quantities are derived from the measurements of the radiation field. The high chromosphere, corona and the transition region emit essentially in the VUV and EUV, which imposes some constraints upon the instrumentation.

The velocity field is crucial for the determination of the momentum and the kinetic energy in the solar plasma. The prime candidate for this measurement is the Doppler effect which allows access to the field component along the line of sight. This requires a rather high spectral resolution (see below) and in some cases, the combination of measurement of apparent motions (in a time-dependent imaging mode) and Doppler velocity field can provide the full velocity vector.

With the advent of the SOHO, TRACE, STEREO, Hinode and now SDO missions, exceptional sequences of images have been obtained in the UV and EUV. They show important dynamics at all spatial and temporal scales. For instance, the so-called EIT waves have been observed with EIT on SOHO at a scale which can reach a solar radius, but the lack of spectral information hampers the identification of these waves. Spectrographs have attempted to identify them: CDS on SOHO observed blue shifted profiles (Harra & Sterling, 2001), which is the possible signature of outflowing material in a coronal mass ejection. But because of the limited field of view (FOV) and the duration of the observing sequence, the measurement has not been confirmed.

At a smaller scale (let us say twice the size of an average active region, 200″×200″), rapid events such as mini or micro-flares can be completely missed: a typical CDS scan lasts 10 minutes while the event front lasts about two minutes. The same is true for the dynamics within coronal loops at high and intermediate temperatures (see the debate about the steady vs. burst heating which cannot be resolved because of the lack of combined spatial, spectral and temporal resolution). At even small scales (a few arc seconds), the solar atmosphere is highly dynamic (e.g. Wang et al., 2000) and a limited spatial scan does not allow us to put the small events in a more general picture.

This means that the combination of imaging and spectral capabilities is necessary. Note that the need of fast spectro-imagery for the study of the transition region is discussed in Gary et al. (2008).

Fast spectro-imagery in the visible

This has been recognized for a long time and many attempts have been made, until now in the visible domain (see Mein 1977 and more recently Mein 1995, for reviews). Starting with the Hα line and later extended, e.g. to the Ca II triplet, the Multi-Channel Subtractive Double Pass spectrograph (MSDP) has obtained important performance (bandpass = 0.13nm @ 656.3nm, FOV 30″ × 250″, spectral resolution = 0.025 nm) and various observatories have implemented it (Meudon, Themis, VTT, Wroclaw, Pic du Midi). The reason for its success lies in the fact that this instrument provides a good compromise between FOV and spectral/spatial/temporal resolutions. (Note that the importance of this instrumentation has been demonstrated by the publication of about 280 papers related to the MSDP). In particular, the various monochromatic images are obtained simultaneously. But this instrument can "only" observe the photosphere and the lower chromosphere.

Other instrumentation has been implemented earlier, such as the UBF at Sac Peak or IPM at THEMIS. Let us recall that they have (as all Fabry-Perots) very interesting properties in the case of narrow

bandpasses and isolated lines and benefit from excellent spectral resolution (e.g. IBIS at Sak Peak). But they have a limited field of view and are definitely limited to visible lines, i.e. to the study of the photosphere and low chromosphere.

Fast spectro-imagery in the UV

The SIMURIS proposal (Damé et al. 1994) included an Imaging Fourier Transform Spectrometer (FTS) with a beam splitter supposed to work above 115 nm. In the EUV, recently, Kankelborg et al. (2006) developed the MOSES concept which flew on 8 February 2006, the instrument, working in He II (30.4nm). It seems that the intensity image was easily recovered but the precise spectral information and the derived velocity field appeared to be more difficult to recover (see section 3.4 of Fox et al. 2010) but was recovered. Let me quote the conclusion of the authors of the paper : "Our observation of this event's three-dimensional structure and dynamics is only possible because of the simultaneous imaging and spectroscopy provided by *MOSES*. With an imager only, we would not know this is a transition region EE. With a slit spectrometer, we would have missed pieces of the full event, perhaps only seeing one of the jets, or only the core, depending on slit alignment, and would not have observed the peculiar, non-collinear geometry of the jets."

In between the EUV and the FUV range, no fast spectro-imaging instrumentation has been proposed (see however Millard 2005, Millard et al. 2006 & Ruiz de Galarreta Fanjul et al. 2010). In particular, the need for a fast spectro-imaging device working in the VUV (say, between 60 nm and 140 nm) has still not been fulfilled, probably because this range extends below 121.5 nm (and in particular 102.5 nm), precluding the use of MgF2, LiF or other dielectrics which could have been potential candidates for use in filters, especially Fabry-Perots.

However, note that Gary et al. (2007, 2008) report very positive results on the development of a Fabry-Perot Interferometer (FPI) dedicated to the C IV doublet around 155nm.

So we are left with all-reflective optics schemes with the further constraint of a minimal number of reflections in order to optimize the throughput. We propose to focus on two different all-reflection optical architectures based on interferometric techniques: Spatial Heterodyne Spectroscopy (SHS), with and without Integral Field Unit (IFU); and Imaging Transform Spectrometer (IFTS). In order to properly compare the pros and cons of these two types, it is important to select a wavelength range and the expected performance in terms of spectral, spatial and temporal resolutions and size of the FOV.

Lyman α

A first step from both observational and instrumental standpoints could be devoted to the Lyman α line at 121.67 nm. The Hydrogen Lyman α doublet is, in some respect, the UV "workhorse" for the hot solar atmosphere, equivalent to the Hα line in the visible. There are many reasons for that: it is the strongest solar UV emission line, it has a large absorption coefficient and H is the most abundant element. It is also a solid tool for diagnostic and a strong contributor to radiation losses in many solar features. But the conditions of its formation are not easy to handle. Lyman α is very optically thick: at a temperature of about 10^4K, the absorption coefficient is about 4.2×10^{-14}cm^{-1}, which means that a layer with a density of about 10^{10}cm^{-3} and a thickness of 1000 km has an opacity of about 4×10^4. Actually this drawback can be turned into an advantage, since each part of the line profile corresponds to a given depth in the atmosphere: Lyman α is formed all through the chromosphere and the bottom of the transition region.

In the case of the solar chromospheres(s) and transition region (TR), profiles are rather broad (at least 0.1 nm wide with very extended wings a few nm away) and, when observations take place at low spatial and temporal resolution, this only requires a rather low spectral resolution (about 0.01 nm) to resolve the reversed profile (see e.g. profiles obtained in coronal holes by Hui Tian et al. 2010). But in the case (e.g.) of prominences, profiles obtained at high spatial resolution (Vial et al. 2007) display features at the scale

of the SUMER resolution (0.005 nm); unfortunately, because of the slit-mode observation, they were difficult to localize in the overall prominence. As for the beautiful VAULT images they don't have any spectral information (Korendyke et al. 2001).

A comparison between two candidates : (I)FTS vs (I)SHS

One can summarize the comparison as follows (for more details, see Ruiz de Galarreta Fanjul et al. 2010) : since detectors are bi-dimensional, the three dimensions cannot be recorded simultaneously. FTSs acquire a panchromatic image and scan through the interferogram in the temporal domain. This can limit the temporal resolution in the study of solar activity. Spectral Heterodyne Spectroscopy (SHS) is an alternative solution, based on the Fourier Transform which simultaneously records all optical path differences. Compared to a conventional FTS, it is insensitive to a change in the intensity of the incoming light (e.g. Damiani 2009). In order to run it in an imaging mode (since it is a slit device), at least one of the spatial dimensions of the data cube must be sequentially acquired over time. One could think of an Integral Field Unit (IFU) which works very well in the VIS-IR and for moderately large FOV. But the mirror slices should have a micron width which is technically unfeasible.

One is left with an IFTS being – as far as we know – the only realistic candidate, and whose supplementary advantage lies in its flexibility: variable resolution enables either moderate-resolution wide spectral range or very accurate line profile acquisition (Millard 2006), a feature which could even be used in on-board operations.

In the case of Lyman α, a preliminary study (Ruiz de Galarreta Fanjul et al. 2010) shows that a 200″ × 200″ field can be measured with a 0.01 nm and 10s resolution. Such performance is rather similar to the performances of the above-mentioned MSDP working in the visible.

Other candidate lines :

We have focused on the study of the Lyman α line, because of its importance and its high output. Of course, an all-reflection IFTS could work on a large bandpass (actually, FTSs were designed essentially for this use), e.g. in the Lyman continuum, the Lyman series (from Lyman β down to 91.2 nm) or windows including optically thin lines representative of the Transition Region. Close to the Lyman β line, one could record the O VI lines between 103 and 104 nm. Above the Lyman α line, SUMER has observed a large variety of emission lines (e.g. N V at 124 nm, Si IV at 139 nm or C IV at 155 nm) emitted at temperatures around 10^5 K. A spectral resolution of about 0.005 nm would be sufficient to build velocity maps with a cadence depending on the effective bandpass. This "free spectral range" could be limited with the help of dedicated selective multilayer mirrors (see e.g. Welsh et al. 1996). With the same type of selective multilayers, still lower wavelengths (< 40 nm) could be reached (Delmotte et al. 2005, Ménesguen et al. 2009).

Technical feasibility of an IFTS in the UV

The concepts have been proposed as early as 1972 by the Wisconsin group (Kruger et al. 1972) and later on by Thorne et al. (1987) but working in the VUV requires special designs and techniques ("all-reflective mirrors") which have hampered major successes until recently. With the advent of ultra-polishing techniques, efficient coatings and multilayers, space-qualified PZT and better UV detectors, it is now possible to build a realistic and space-qualified UV instrument (see e.g. Thorne 1998). In physics laboratories (such as the SOLEIL Synchrotron ring), VUV (non-imaging) FTS are routinely used (de Oliveira et al. 2009). In terms of the (difficult) metrology and stabilization, one can think of transposing the techniques used "on the ground" to instrumentation designed to fly in space. Before envisaging such a possibility, it is necessary to test the feasibility on mock-ups, to test the performance (especially in terms of stability of the optical path difference) and design a flyable instrument.

We think that such an effort is worth the effort, in view of its astrophysical benefits.

Some references

N.B. This note is not comprehensive about the various past efforts. It simply relies upon some work that I am more familiar with.

Damé, L., Rutten, R.-J., Thorne, A.-P., and Vial, J.-C., "The SIMURIS interferometric mission: Solar physics objectives and model payload," *Advances in Space Research* **14**, 167 (1994).

Damiani, C., 2009, Etude en laboratoire des propriétés d'un spectromètre de Fourier Hétérodyne Imageur pour applications spatiales, PhD Thesis, Observatoire de Paris

Delmotte, F., Gautier, J., Roulliay, M., et al., 2005, *B4C/Mo/Si multilayers for 20 – 40 nm wavelengths : application to broadband mirrors,* Proc. SPIE 5963, 59631V, Advances in Optical Thin Films II, Claude Amra, Norbert Kaiser, H. Angus Macleod, Editors

de Oliveira, N., Joyeux, D., Phalippou, D., Rodier, J.-C., Polack, F., Vervloet, M., and Nahon, L., "A Fourier
transform spectrometer without a beam splitter for the Vacuum Ultraviolet range: From the optical design to the first UV spectrum," Rev. Sci. Instrum. 80, 043101 (2009)

Fox, J.L., Kankelborg, C.C., Thomas, R. J., 2010, *A transition region explosive event observed in He II with the MOSES sounding rocket*, The Astrophysical Journal, Volume 719, Issue 2, pp. 1132-1143

Gary, G.A., West, E.A., Rees, D., Zukic, M., & Herman, P., 2007, *Solar CIV vacuum-ultraviolet Fabry-Pérot interferometers*, A & A, 461, 707-722.

Gary, G. A., Davis, J.M., & West, E. A., 2008, *Future Possibilities for Doppler and Magnetic Field Measurements in the Extended Solar Atmosphere*, COSPAR General Assembly, Montreal

Harra, L.-K. and Sterling, A.-C., 2001, "Material Outflows from Coronal Intensity "Dimming Regions" during
Coronal Mass Ejection Onset," *The Astrophysical Journal* **561**, L215–L218 (2001)

Hui Tian, L. Teriaca, W. Curdt, J.-C.Vial, 2009, *Hydrogen Lyman-alpha and Lyman-beta radiances and profiles in polar coronal holes,* Astrophysical Journal Letters 703, L152

Kankelborg, C., Fox, J.-L., and Thomas, R.-J., "First Results from the MOSES Rocket Flight," AAS Solar Physics Division Meeting 37 (2006)

Kruger, R.A., Anderson, L.W., and Roesler, F.L., 1972, *All reflection interferometer for use as a Fourier-Transform Spectrometer*, JOSA (1917-1983), 62, 938

Korendyke, C.~M. , Vourlidas, A., Cook, J.W., et al., 2001, High-resolution Imaging of the Upper Solar Chromosphere: First Light Performance of the Very-high-Resolution Advanced ULtraviolet Telescope, Solar Physics, 200, 63

Mein, P., "Multi-Channel Subtractive Spectrograph and Filament observations," *Solar Physics* **54**, 45–51 (1977)

Mein, P., "Solar Imaging Spectroscopy: Multichannel Subtractive Double Pass Instruments," in [*Tridimen-*
sional Optical Spectroscopic Methods in Astrophysics], series, A. C., ed., *Proc. of IAU* **71**, 350 (1995)

Ménesguen T ., de Rossi S., Meltchakov, E., Delmotte F., 2009, *Aperiodic multilayer mirrors for efficient broadband reflection in the extreme ultraviolet*, Appl. Phys. A, DOI 10.1007/s00339-009-5400-2

Millard, A., 2005, IFTSUV : un spectromètre imageur à transformée de Fourier dans l'ultraviolet pour les prochaines missions spatiales solaires, PhD Thesis, Université Paris-Sud

Millard, A., Lemaire, P., and Vial, J.-C., "IFTSUV: an imaging Fourier transform spectrometer in UV for the next solar space missions," in [*Space Telescopes and Instrumentation II: Ultraviolet to Gamma Ray*], Turner, Martin, J.-L., and Hasinger, G., eds., *Proc. SPIE* **6266** (2006)

Ruiz de Galarreta Fanjul, C., Philippon, A., Vial, J.-C. et al., 2010, *Definition of an imaging spectrometer meeting the needs of UV solar physics*, Space Telescopes and Instrumentation: Ultraviolet to Gamma Ray. Edited by Arnaud, Monique; Murray, Stephen S.; Takahashi, Tadayuki. Proceedings of the SPIE, Volume 7732, pp. 773236-773236-12

Thorne, A.P., Harris, C.J., Wynne-Jones, I., et al., 1987, *A Fourier transform spectrometer for the vacuum ultraviolet : design and performance*, Journal of Physics E : scientific instruments, 20, 54

Thorne, A., "High resolution Fourier transform spectrometry in the visible and ultraviolet regions," *J. Anal.*
At. Spectrom. **110**, 1205–1215 (1998)

Vial, J.-C., Ebadi, H., and Ajabshirizadeh, A., "The Lyα and Lyβ profiles in solar prominences and the issue
of the fine structure," *Solar Physics* **246**, 327 (2007)

Wang, J., Li, W., Denker, C., Lee, C., Wang, H., Goode, P.R., McAllister, A., and Martin, S.F., "Minifilament
Eruption on the Quiet Sun. Observations at Hα Central Line," *The Astrophysical Journal* **530**, 1071–1084 (2000)

Welsh, B. Y., Siegmund, O. H., Jelinsky, P. N., 1996, *Lyman Imaging Telescope Experiment*, Proc. SPIE Vol. 2807, p. 11-19, Space Telescopes and Instruments IV, Pierre Y. Bely; James B. Breckinridge; Eds.

A White Paper advocating a *Heliophysics Theory Mission*

Principle Authors: J. R. Jokipii(1) and Gary P. Zank(2)

(1)Distinguished Professor of Astrophysics, University of Arizona

(2)Pei-Ling Chan Chair of Physics –and- Director, Center of Space Plasma and Aeronomic Research –and- Chair, Physics Department, University of Alabama in Huntsville

1. Introduction

Theory and modeling, together with observations are the foundations of a sound, balanced scientific program in any discipline. The past few decades have witnessed a shift from a strongly exploratory and discovery-driven science to a more mature explanatory science. With its maturing, solar and space physics places increasing demands on theory. Theory provides a meaningful context for basic space physics observations, often revealing that seemingly disparate observed phenomena correspond to the same physical processes in a system(s). Critically, besides organizing and understanding observations, theory drives the prediction of new or unexpected important phenomena that in turn drive new missions.

Some 34 years ago it was realized that, within what was then called the Space Physics Division of NASA, theory and modeling was not adequately funded. At that time the Colgate Committee was formed. It found that, compared to peer disciplines such as fusion plasmas, space physics theory and modeling was significantly under-supported. The committee recommended that theory funding be increased and attention be given to the encouragement of critical-size theory groups. This recommendation led directly to the creation of the Space Physics Theory Program and indirectly raised the general awareness of the importance of a balanced program. This program still exists as the Heliophysics Theory Program, but at a significantly reduced fraction of the overall budget. This program has been an important factor in the significant progress made in the area of Heliophysics science of the past few decades, but the number of proposals supported now numbers 9 with a typical grant size of some $400k or less per year. This has decreased from 11 or 12 funded proposals just 20 years ago, and these were funded at levels that were comparable to today's (i.e., ~$400k in 1990 dollars)! There has therefore been a considerable erosion in both the number of theory groups supported in the US and in the level of individual support. With the reduction in the theory program, most theory is now being funded through guest investigator programs associated with missions or through the SR&T program. Unfortunately, the level of these programs is also being reduced, to the point that the last GI program was completely cancelled. The deleterious effect of the erosion in theory support can scarcely be over-estimated, especially in academic environments where it is now almost impossible to develop a major spacecraft or experimental program that can sometimes indirectly support the development of theory.

In significant measure, this erosion of theory and modeling funding appears to be the result of significant budget constraints combined with Mission cost growth. With respect to the latter, there appear to be several factors at work. Proposed budgets for a mission may not always reflect accurately the real costs – this may be due to overly aggressive or optimistic budgeting to ensure

the selection of an instrument or mission, or it may be due to the shifting of the mission schedule and dates, or unexpected development costs, etc. NASA has frequently responded by adding further support to the mission (or perhaps de-scoping the mission, but often still with an increased budget), and the increase in mission funding has all too often resulted in a corresponding decrease in the level of support for theory and modeling related programs.

The purpose of this white paper is two-fold. 1) We recommend that theory and modeling be funded at a significantly enhanced level that would allow sizable teams of researchers to pursue major science projects using a combination of theory, modeling, and data analysis. We suggest that an individual grant could total as much as $1M/year, and that as many as 10 groups should be supported. 2) We recommend that the theory program be given the status of a Mission and be similarly protected against the vagaries of NASA funding, particularly in maintaining the level of funding and adjusting for inflationary increases.

2. Current state of Theory and Modeling

The Colgate report listed 6 specific problem areas in basic theory as important for the progress of Space Physics:

(1) Magnetic-field reconnection
(2) Interaction of turbulence with magnetic fields
(3) Behavior of large-scale flows and their interactions
(4) Acceleration of energetic particles
(5) Particle confinement and transport
(6) Collisionless shocks.

Although this list was compiled some 34 years ago, it is remarkably applicable to the present time. Significant progress has been made.

We now have a significantly improved understanding of each of the problems listed, due to more and better observations allied with theoretical and modeling advances, but in none of them can we claim to have achieved a more-or-less definitive understanding of the problem. Of course, the richness of each of these problems has spawned both increased understanding in related fields and new fields of inquiry.

Some case studies provide an interesting perspective on the development of theory, the associated modeling and simulations, and the role of data analysis. One example, a subset of the collisionless shock example, illustrates both our increasing understanding and the further fundamental questions that are opened up by theory.

Collisionless Shocks Our current understanding of perpendicular shocks is rightfully regarded as one of the major achievements of modern collisionless nonlinear plasma physics. A perpendicular shock wave has the upstream magnetic field aligned perpendicular to the shock normal. Deviations up to nearly 45^0 for the magnetic field from the shock normal are described as quasi-perpendicular, and a clear physical picture can be drawn since reflected ions will gyrate back to the shock front in this case. Nearly perpendicular shocks have received considerable

attention because of their relatively clean, laminar appearance in the time series data. The particle gyromotion in the quasi-perpendicular magnetic field acts to prevent particles behaving excessively diffusively (i.e., the particles tend to be bound together), thereby simplifying the collisionless processes responsible for thermalization of the plasma, and the injection and acceleration of energetic non-thermal particles. Despite much effort, key questions remained unanswered or open to interpretation, and prior to Cluster, single, and dual, spacecraft studies were unable to place quantitative limits on the spatial scales, or address non-stationarity of the overall shock transition. By taking advantage of the sharp, quasi-perpendicular shock transitions and four spacecraft techniques, Cluster studies can probe internal shock scales and physics, including energetic particle origin, of the Earth's quasi-perpendicular bow shock. This has led to a resurgence in theoretical studies of the perpendicular shock, with many of the formerly accepted assumptions now being questioned. For example, the role of dimensionality in kinetic simulations and the importance of the electron/ion mass ratio are now recognized as introducing different behavior in the simulations depending on the assumptions. In all of this, the role of magnetic-field-line mixing must also be considered.

The capacity for the Voyager Interstellar Mission to surprise remains undiminished, with the recent crossing of the heliospheric termination shock (HTS) by the Voyager 2 spacecraft. With a working plasma instrument, the Voyager 2 observations revealed a quasi-perpendicular heliospheric termination shock that appears to be considerably different in character from quasi-perpendicular shocks observed in the inner heliosphere. The Voyager 2 observations suggest the possibility of a broad structure, or a velocity slowdown before the shock, possibly considerable fine-scale structure, many partial termination shock crossings, very little heating of the solar wind protons, to the extent that the downstream flow appears not to be subsonic, and a density spike appears to be present. These results suggest that the heliospheric termination shock is unlike any heliospheric shock observed elsewhere. Nonetheless, the observations were not entirely unexpected theoretically. It had been argued that the primary dissipation mechanism at a quasi-perpendicular heliospheric termination shock would be reflected pickup ions, and that the solar wind ions would be heated very little. This basic concept appears to have been borne out by the Voyager 2 observations.

Thus, theory appears to be making headway in our basic understanding of quasi-perpendicular shock physics but fundamental questions remain unresolved, and have not been clarified observationally, in part because without an accepted theoretical structure, it is difficult to know what we should be looking for. For example, even the question of the scale length over which changes in the electric field occur and its relation to the scale size over which changes in the magnetic field occur is unclear theoretically and observationally.

Transport and acceleration of energetic particles The acceleration of energetic particles and cosmic rays is thought by many to be caused by shock waves. Perpendicular shocks, such as much of the termination shock, are efficient and fast accelerators. With the crossings of the termination shock, it has become clear that the early theoretical models were too simple. The energetic particles did not behave as expected. In response, a number of approaches have been suggested. Non-shock acceleration mechanisms such as reconnection and statistical acceleration have been resurrected and the effects of pre-existing turbulence have been examined. Thus, theory drove certain expectations of particle acceleration at the heliospheric termination shock,

some of which were met and others not by the Voyager observations. Clearly, a major theory effort is now needed to address these recent developments.

All of these acceleration issues require an increased understanding of particle transport, both in formulating the theories and in interpreting spacecraft observations which are most-often remote from the acceleration site. These theoretical and modeling issues are of basic importance throughout physics, ranging from thermonuclear plasmas to supernova blast waves. Thus, increased knowledge here will have broad impact.

Magnetohydrodynamic turbulence in the solar wind Magnetohydrodynamic (MHD) turbulence is characterized by nonlinear interactions among fluctuations of the magnetic field and flow velocity over a range of spatial and temporal scales. It plays an important role in plasma heating, the transport of energetic particles, such as galactic and anomalous cosmic rays, and radiative transfer and is ubiquitous in the solar and interplanetary plasma. The solar wind exhibits turbulent behavior, as can be observed from in situ data. Solar wind fluctuations are observed over spatial scales that range from many AUs to electron kinetic scales. These observations have stimulated an ongoing effort to develop theoretical treatments of MHD turbulence. Large-scale fluctuations can be described by fluid models, but the smaller than proton scales requires Hall MHD or a kinetic description. The coupling of large-scale and kinetic-scale processes in the solar wind is mediated by the turbulence. Energy is transferred from large-scales to small-scales through turbulent fluctuations, including eddies and waves, that interact. The cascade is eventually halted by kinetic processes and heats the plasma. With an explicitly turbulence-based model that includes appropriate source terms, the radial evolution of turbulence intensity and the temperature of the plasma has been computed successfully from 1 to > 50 AU.

Solar wind turbulence remains incompletely understood despite considerable progress in the last 20 years. It is an important problem since turbulence mediates the complex dynamical couplings between large and small scales, slow fluid motions and fast kinetic processes, and low energy and high energy charged particles. Numerous important theoretical problems remain unresolved, ranging from compressible effects, passive scalar transport, the dimensionality and symmetries of interplanetary turbulence, etc.

While the Colgate Report list of fundamental plasma physics problems remains comprehensive, at least one additional area of fundamental research has assumed particular importance for solar and interplanetary plasmas. In the examples described above, a common element was the interaction of the heliosphere with interstellar material, especially the neutral gas (hydrogen). This took the form of a new shock dissipation mechanism based on interstellar pick-up ions, the acceleration of anomalous cosmic rays, and even the heating of the solar wind by pick-up ions providing the energy (through a resonant instability associated with pick-up) that cascades to kinetic scales and heats the solar wind. A more detailed example related to the formation of the "hydrogen wall" is given below. Thus, to the Colgate Report list, we would add the additional fundamental problem of ***Partially-Ionized Plasmas***.

Structure of the heliosphere The physics of the outer heliosphere and the large-scale structure of the heliosphere is determined fundamentally by its interaction with the partially ionized local interstellar medium (LISM). To illustrate, an important development in outer-heliospheric

research was the prediction and discovery of the "hydrogen wall." Theoretical models predict that the partially ionized LISM and the solar wind are separated by a complex set of plasma and neutral-atom boundaries, of which the termination shock has now been observed by the Voyager spacecraft. A specific theoretical prediction of the models is that a wall of interstellar neutral hydrogen should exist in the upstream region owing to the relative motion of the heliosphere and the LISM. This hydrogen wall is predicted to have a number density slightly more than twice the interstellar density, to be hotter than interstellar hydrogen, and to be some 100 AU wide. The physical reason for the hydrogen wall is the deceleration and diversion of the interstellar plasma flow about the heliosphere leading, through charge-exchange coupling, to a pile-up and heating of interstellar neutral hydrogen. The result is the formation of a giant wall, which acts to filter neutral hydrogen as it enters the heliosphere. Confirmation of the hydrogen wall's existence was not expected for decades, but a serendipitous convergence of predictive theoretical modeling, observations to place limits on the cosmological deuterium/hydrogen ratio, and a multi-disciplinary investigation spanning space physics and astrophysics led to the detection of the hydrogen! This was the first of the boundaries separating the solar wind and the LISM to be discovered, and it offers a glimpse into the global structure of the three-dimensional heliosphere.

The research leading to the discovery of the hydrogen wall is an excellent example of theory driving the frontiers of space science and motivating the development of new observational techniques and methodology to complement traditional space physics tools.

3. *Theory and Modeling*

It is important to emphasize that the terms theory and modeling are not synonymous, and th must both be present in a balanced program. We view theory as the extension of our basic knowledge of phenomena. Theory attempts to find new physical descriptions of poorly understood phenomena. Modeling, on the other hand, involves the implementation of theory to complicated, realistic situations and scenarios. This often involves considerable investment in computer hardware and software. The end results are quantitative predications that can be directly compared to observations. An illustrative example is the creation of a synthesized time dependence of the predicted magnetic field or energetic particles along a real spacecraft trajectory to compare with an observed timeline.

Noting this distinction between theory and modeling allows us to consider several essential elements that a theory program needs to address.

Formulation of theoretical models In recognizing that multiple scales, regions, processes, and plasma populations are intrinsic to the challenging space physics problems of today, the correct mathematical and physical formulation is critical. In this, there is no substitute for time-honored analytical approaches to theoretical developments in plasma physics, fluid dynamics, and applied mathematics. Progress on highly nonlinear, coupled plasma problems may be made using techniques that range from the relatively standard to nonlinear, low-order reductive approaches and statistical methods. Current agency funding programs are not adequate to even support these basic theoretical efforts, and unfortunately innovative and bold ideas, approaches, and techniques in proposed research are often not encouraged and rewarded. Computation is no substitute for the development of rigorous theories and well-conceived models. Basic theory must be regarded as a

critical component of the funding profile for either NASA or the NSF and a Mission-level perspective must be encouraged

Invariably, many of the problems listed in the "extended" Colgate Report list impose significant computational demands in terms of CPU power and the concomitant development of sophisticated and efficient codes. Two challenges face the community. The first is to further develop existing codes and algorithms, such as three-dimensional MHD codes that incorporate adaptive mesh refinement, for example, or three-dimensional hybrid codes with improved electron/ion mass ratios or improved codes for data exploration. These problems do not demand the inclusion of new physics but demand instead substantial progress in current research areas. The second challenge lies in developing and implementing new computational approaches for both model solving/simulation and data exploration that exploit advances made by numerical mathematicians, statisticians, and computer scientists.

The coupling of different physical processes, scales, and regimes and the self-consistent incorporation of multiple scales, physical processes, and distinct regions into models will be the main challenge to theorists and modelers in the coming decade, demanding the formulation and development of sophisticated models and theory, the development of new and innovative algorithms, access to sophisticated computational resources, and the opportunity to test model predictions and validate theories against existing and future observations. We know that theory will demand sophisticated new measurements, which will in turn drive and define new space and ground-based missions (in situ, multipoint, remote, etc.). A theory program must focus on the investigation of well-chosen, theoretical problems and the development of coupled global models. For major advances to be made in space physics, fundamental theoretical analysis, sophisticated computational tools, and state-of-the-art data analysis must all be coupled intimately under a single umbrella program. Theoreticians working with pen and paper, computational space physicists, and data analysts will be needed collectively to achieve the major advances expected of space physics. Only by creating and maintaining major groups of this sort can a strong and vital connection between basic science, computation, and observations be achieved.

A well-balanced discipline would be in a position to leverage more fully the remarkable observational possibilities to increase our understanding of our home in space.

4. *Recommendation and Implementation*

The above specific examples and issues document the real decrease in support of theory, both absolutely and relative to the missions. We therefore recommend that theory and modeling be protected from further erosion and be restored to the relative position it had in the aftermath of the Colgate report.

We propose the creation of a New Theory MISSION. What is needed for a successful Theory Program? The answers are rather simple:

- Long-term, stable funding;

• Synergistically interacting groups of students, postdoctoral associates, research scientists, and several university or institutional faculty who are able to integrate research and education at some level.

To guard theory and modeling from continuing erosion, we propose that a theory mission be created with a specific goal or goals and a specified time line, in analogy with a hardware mission. One possible scenario would involve selecting one of the issues discussed above as the target of the Mission. For example, one might select a 'Mission to Understand the Interaction of the Heliosphere with the Local Interstellar Medium", or perhaps 'The Sources of the Interplanetary Magnetic Field as a focus for a 5-year effort. Within such a mission concept, one would fund separate groups (in analogy to separate instruments), some to work on aspects of basic theory, others to develop models and simulations and perhaps others to work on the connections with observations.

Such a mission should NOT be funded at the expense of the usual SRT or GI programs.

In summary,
1) we recommend that theory and modeling be funded at a significantly enhanced level that would allow sizable teams of researchers to pursue major science projects using a combination of theory, modeling, and data analysis. We suggest that an individual grant could total as much as $1M/year, and that as many as 5-10 groups should be supported at any one time, for up to 5 years; and 2) we recommend that the theory program be given the status of a Mission and be similarly protected against the vagaries of NASA funding, particularly in maintaining the level of funding and adjusting for inflationary increases.

Role of the National Science Foundation ATM/GEO Directorate in promoting and supporting Space Physics

A white paper for the 2010 Solar and Space Physics Decadal Survey

Principal Author: G.P. Zank
Pei-Ling Chan Chair of Physics –and- Director, Center of Space Plasma and Aeronomic Research –and- Chair, Physics Department, University of Alabama in Huntsville

1. Introduction

The NSF ATM/GEO Directorate has played a fundamental role historically in supporting space physics. Traditionally, this has taken a broad systems perspective, recognizing that our understanding of geo-spatial processes at 1 Astronomical Unit is informed by physical processes and couplings from regions as diverse as the solar corona and the boundaries of the heliosphere. *Unfortunately, there is an effort within the NSF ATM/GEO Directorate to significantly restrict such a broader perspective by limiting the space physics that is supported to that within 1 AU or which has an obvious demonstrable impact on the Earth.* Such thinking represents both a significant step backwards in our discipline for many reasons, and may in fact not be in compliance with NSF policy. Specifically, the "NSF has long recognized the value of interdisciplinary research in pushing fields forward and accelerating scientific discovery. Important research ideas often transcend the scope of a single discipline or program. NSF also understands that the integration of research and education through interdisciplinary training prepares a workforce that undertakes scientific challenges in innovative ways." [National Science Foundation Investing in America's Future: Strategic Plan FY 2006-2011.] In light of this, the ATM/GEO Directorate appears to be in violation of basic NSF policy by limiting funded research to within 1 AU, and providing no other home for this work.

In this White Paper, we provide just a few examples with which we are particularly familiar to illustrate why such a restriction by the ATM/GEO Directorate is shortsighted.

- With this White Paper, we are requesting that a recommendation be made to the NSF ATM/GEO Directorate that a broad perspective of space physics be adopted and that the Directorate return to compliance with NSF policy. The NSF should support solar and heliospheric physics research within one section and one directorate without imposing an artificial boundary to the heliosphere at 1 AU and without insisting that, to be supported by that directorate, one needs to make the case that the research is of direct relevance to the Earth.
- That the restriction of funding to support studies only within 1 AU or that are directly relevant to the Earth be eliminated;
- And that a panel be established to evaluate the broad compliance of these goals as reflected in funding trends made to individual researchers.

The region within 1 AU is not an environment isolated from the rest of the heliosphere. Several examples point directly to why an understanding of the physics of the interplanetary environment beyond 1 AU has a direct influence on our understanding of geospace, some of which are discussed below. The development of techniques and models and the elucidation of physical

processes for regions of the heliosphere beyond 1 AU can frequently shed light on problems within 1 AU. Finally, in attempting to promote the cross-disciplinary and multi-process approach to space physics that is undoubtedly necessary to make the break-through discoveries necessary to advance the field, it is imperative that we not compartmentalize our science. This is especially true if we are to encourage students to enter the field and post-docs and researchers to remain part of the space physics community. This implies that we cannot restrict NSF ATM/GEO supported science to an arbitrary and unphysical 1 AU boundary in the interplanetary medium.

2. Importance of physics beyond the Earth

A major element of space weather is understanding the particle radiation environment at 1 AU, especially solar energetic particles. Much of our current understanding of particle energization derives from diffusive shock acceleration. It hardly needs noting, however, but the entire theory of diffusive shock acceleration was developed without any reference to particle acceleration in the inner heliosphere. Initial papers on the subject focused on particle acceleration at supernova remnants and then later at the heliospheric termination shock. It should be added that this latter application was in fact funded through ATM/GEO many years ago. It was only some 10 years or more later that these ideas migrated into the field of SEPs. Today, such science would not be funded under the current policy because it would not have been immediately obvious how it would have relevance to the Earth. While it is not necessary for ATM/GEO to now consider funding particle acceleration studies at supernova remnants, a better understanding of how particles are accelerated at e.g., the heliospheric termination shock may ultimately shed much light on the general physics of diffusive shock acceleration, and hence on the SEP problem, probably in ways that are difficult to anticipate.

Consider another example which pertains to space weather but which would not be funded under the current policy. The evolution of turbulence in the outer heliosphere, beyond 1 AU, would appear to be of little consequence to our understanding of the transport of solar energetic particles. It has been established that energetic SEPs accelerated within 1 AU and observed at the Earth typically propagate well beyond 1 AU before being scattered back by turbulent magnetic field fluctuations at several 10's of AU. Turbulence beyond some 5 – 8 AU is very strongly influenced and driven by instabilities associated with the pick up of interstellar neutral hydrogen. Thus, to understand the transport of SEPs and their fluence at 1 AU not only requires ultimately an understanding of turbulence evolution throughout the heliosphere but also a knowledge of the heliospheric structure, properly taking into account the role neutral hydrogen.

A further example, attracting much discussion, is the ubiquity of a v^{-5} spectrum for energetic particles, but this was first discussed in the context of puzzling Voyager observations made in the heliosheath. Besides exploring related observations in the outer heliosphere, theoretical models were developed to describe the observations and these are now finding application in the inner heliosphere. The migration of ideas developed for application to one part of the heliosphere to another part of the heliosphere is almost inevitable and to somehow bound the heliosphere to 1 AU is completely contrary to the way that science is done.

Consider a final example that would certainly not have any possibility of funding under the current policy. Clearly, the historical behavior of the sun and solar wind is of considerable interest to ATM/GEO, especially its potential impact on long-term climate trends. A much cited

paper by Wood et al., 2001 presented the Sun's mass loss history, showing how this may be responsible for the loss of the Martian atmosphere, for example. The methodology by which these results were arrived at required sophisticated modeling of nearby stellar astrospheres of G-class stars in partially ionized clouds. The models used were based on simulations developed to understand the interaction of our heliosphere with the local interstellar medium. The heliospheric models predicted the existence of the hydrogen wall, which was subsequently observed using Lyman-alpha absorption measurements. The prediction and subsequent measurement of similar hydrogen walls about the astrospheres of other G-class stars at different stages in their evolutionary history allowed Wood et al. to determine a the history of the Sun's mass loss and related coronal activity. The development of the necessary heliospheric models is not supported under the current funding policy.

A number of additional examples can be provided quite easily from other areas of heliophysics but the point has been made adequately that we should not artificially introduce a boundary of 1 AU in our discipline. It should be evident that understanding the physics of one region of the heliosphere frequently informs our understanding of other aspects and regions. We cannot afford a blinkered approach to the science of heliophysics and we must recognize that the heliosphere is a single, coupled, integrated system. The commonality of physical processes within the heliosphere and the key role of in situ observations, whether by ACE or Voyager, distinguishes the science of heliophysics from astrophysics. We cannot allow the physics of the heliosphere to be shunted off to astrophysics because, for whatever reason, it has been determined that heliophysics ends at 1 AU.

3. Impact on young researchers

A final important point to be made concerns the impact that the current NSF ATM/GEO policy is having on building our field. NSF policy for workforce training suggests the necessity of researchers being trained in all areas of heliospheric research. ATM/GEO policy as implemented will limit the training and career development of the scientists who will be required to conduct heliospheric research in future. The bringing of young researchers into faculty positions in research universities may be the most critical component that will ensure the long-term survival of heliophysics as a discipline. At this point, a young assistant professor whose work investigates for example the structure of the heliosphere would not be able to propose for a CAREER award through the NSF ATM/GEO Directorate (nor indeed for any other form of research support). Unfortunately, heliospheric studies beyond 1 AU without an immediately obvious connection to geo-space cannot be accommodated in the NSF Astronomy section, since astronomers do not regard the Sun as being a particularly special star.

As is typical of many of us in the field of heliophysics, most of us move from one area or region to another within the heliosphere since many of the ideas are interchangeable and offer insight elsewhere. What will happen now is that good young heliospheric scientists will see few opportunities beyond the very narrow compartmentalized options that are currently available and move to different disciplines. Our field will be the ultimate loser, and the research universities will continue to hire in fields that are more promising than heliophysics with a 1 AU boundary.

A Radio Observatory on the Lunar Surface for Solar studies (ROLSS)

R. J. MacDowall[1,6], T. J. Lazio[2,6], S. D. Bale[3,6], J. Burns[4,6], W. M. Farrell[1,6], N. Gopalswamy[1], D. L. Jones[2,6], and K.W. Weiler[5]

(1) NASA/GSFC, Greenbelt, MD, USA, (2) JPL, Pasadena, CA, USA, (3) SSL, Berkeley, CA, USA, (4) U. Colorado, Boulder, CO, USA, (5) NRL, Washington, DC, USA, (6) NASA Lunar Science Institute, NASA Ames Research Center, CA, USA

1. Introduction

By volume, more than 99% of the solar system has not been imaged at radio frequencies. Almost all of this space (the solar wind) can be traversed by fast electrons producing radio emissions at frequencies lower than the terrestrial ionospheric cutoff, which prevents observation from the ground. To date, radio astronomy-capable space missions consist of one or a few satellites, typically far from each other, which measure total power from the radio sources, but cannot produce images with useful angular resolution. To produce such images, we require arrays of antennas distributed over many wavelengths (hundreds of meters to kilometers) to permit aperture synthesis imaging. Such arrays could be free-flying arrays of microsatellites or antennas laid out on the lunar surface. In this white paper, we present the lunar option. If such an array were in place by 2020, it would provide context for observations during Solar Probe Plus perihelion passes. Studies of the lunar ionosphere's density and time variability are also important goals.

This white paper applies to the Solar and Heliospheric Physics study panel.

2. Scientific Goals

High-energy particle acceleration occurs in diverse astrophysical environments including the Sun and other stars, supernovae, black holes, and quasars. A fundamental problem is understanding the mechanisms and sites of this acceleration, in particular the roles of shock waves and magnetic reconnection. Within the inner heliosphere solar flares and shocks driven by coronal mass ejections (CMEs) are efficient particle accelerators.

Low frequency observations provide an excellent remote diagnostic because electrons accelerated in these regions can produce intense radio bursts. The intensities of these bursts make them easy to detect, as well as providing information about the acceleration regions. The radio burst mechanisms discussed here occur at the local plasma frequency, $f_p \approx 9\ n_e^{1/2}$ kHz, or its harmonics, where n_e is the electron density in cm^{-3}. With a model for n_e, f_p can be converted into a height above the corona, and changing f_p can be converted into radial speed. Observations by widely-separated spacecraft permit triangulation.

Solar radio bursts are one of the primary remote signatures of electron acceleration in the inner heliosphere and our focus is on two emission processes, referred to as Type II and Type III radio bursts. Type II bursts originate from suprathermal electrons (E > 100 eV) produced at shocks. These shocks generally are produced by CMEs as they expand into the heliosphere with Mach numbers greater than unity. Emission from a Type II burst drops slowly in frequency as the shock moves away from the Sun into lower density regions at speeds of 400–2000 km s^{-1}. Type III bursts are generated by fast (2–20 keV) electrons from magnetic reconnection. As the fast electrons escape at a significant fraction of the speed of light into the heliosphere open along magnetic field lines, they produce emission that drops in frequency rapidly (see Figure 1).

Electron densities in the inner heliosphere yield relevant frequencies below 10 MHz. Observations must be conducted from space because the ionosphere is opaque in this frequency range. Figure 1 illustrates the active low-frequency radio environment in space, including terrestrial radio frequency interference

(RFI), as seen by the WAVES instrument on the Wind spacecraft (Bougeret et al. 1995). Solar radio observations would necessarily take place in the gaps between the RFI.

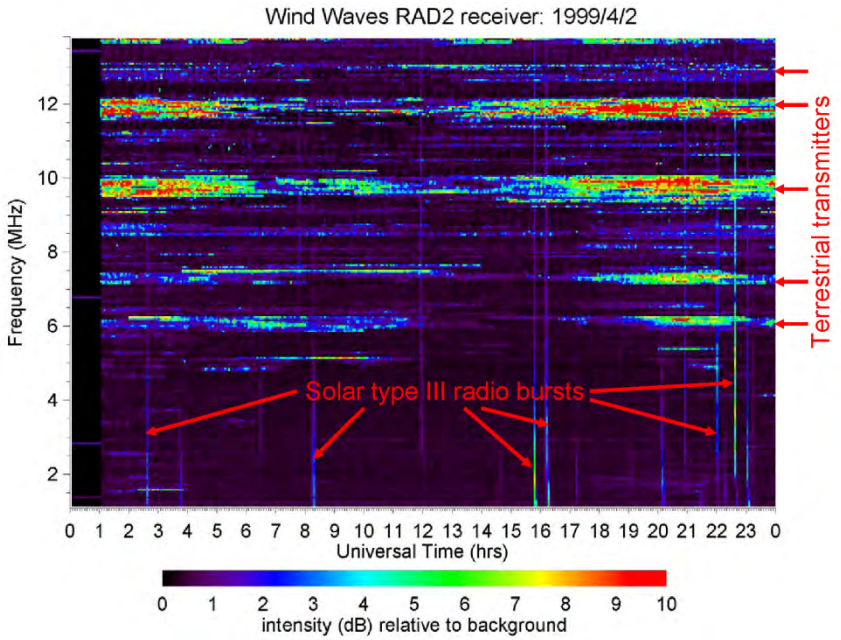


Fig. 1 - Observed low-frequency radio environment when the Wind spacecraft was near the moon. The 24-hr dynamic spectrum covers 1–14 MHz, with the intensity on a logarithmic color scale.

Acceleration at Shocks: Observations of CMEs near Earth suggest electron acceleration generally occurs where the shock normal is perpendicular to the magnetic field (Bale et al. 1999), similar to acceleration at planetary bow shocks and other astrophysical sites. This geometry may be unusual in the corona, where the magnetic field is largely radial, as shown schematically in Figure 2a. However, geometric arguments suggest that the shock at the front of a CME generally has a quasi-parallel geometry (Q-||). Acceleration along the flanks of the CME, where the magnetic field-shock normal is quasi-perpendicular (Q-⊥) would seem to be a more likely location for the electron acceleration and Type II emission. The radio array needs ~2º resolution to localize the acceleration site(s) and geometry (Q-|| vs. Q-⊥) around CMEs.

Electron and Ion Acceleration: Observations at 1–14 MHz made with the Wind spacecraft showed that complex Type III-L bursts are highly correlated with CMEs and intense (proton) solar energetic particle (SEP) events observed at 1 AU (Cane et al. 2002; Lara et al. 2003, MacDowall et al. 2003). While the association between Type III-L bursts, proton SEP events, and CMEs is now secure, the electron acceleration mechanism remains poorly understood. Two competing sites for the acceleration have been suggested: at shocks in front of the CME or in reconnection regions behind the CME, see Figure 2b. For typical limb CMEs, the angular separation of the leading edge of the shock and the hypothesized reconnection region behind the CME is approximately 1.5° when the CME is 3–4 $R_{\odot}$ from the Sun.

CME Interactions and Solar Energetic Particle (SEP) Intensity: Unusually intense radio emission can occur when successive CMEs leave the Sun within 24 hours, as if CME interaction produces enhanced particle acceleration (Gopalswamy et al. 2001, 2002). Statistically associated with intense SEP events (Gopalswamy et al. 2004), this enhanced emission could result from more efficient acceleration due to changes in field topology, enhanced turbulence, or direct interaction of the CMEs. Lack of radio imaging makes it difficult to determine the nature of the interaction. Images with <2 degree resolution would give Type II locations and permit identification of the causal mechanism and the relation to intense SEPs.

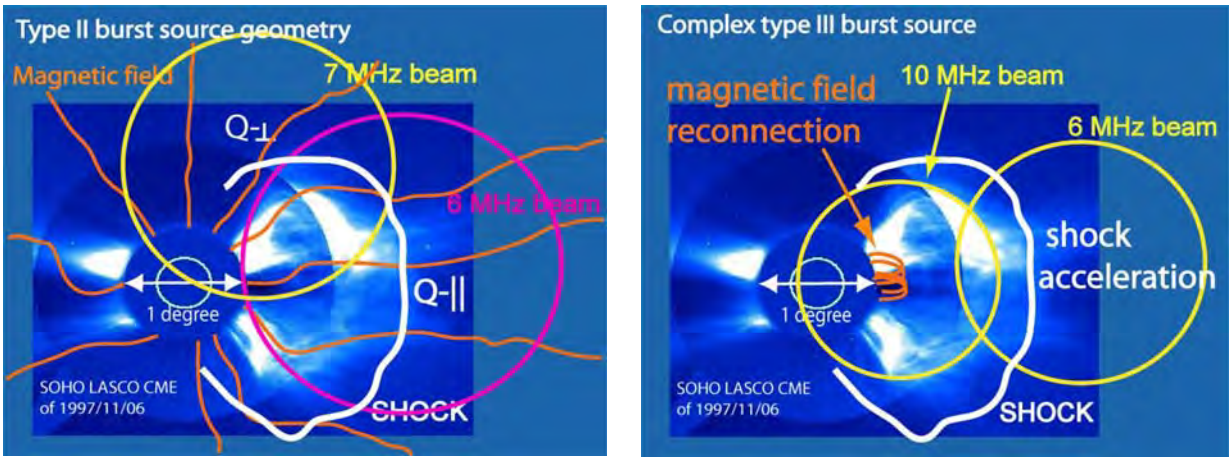


Fig. 2 – (Left) Where on the shock does electron acceleration occur, yielding type II radio emission? (Right) Are complex type III-L bursts produced by shock acceleration or reconnection?

Relevance to Solar Probe Plus (SPP): The ROLSS mission would image radio sources in the volume from ~2 to 10 solar radii, corresponding to the region sunward of SPP perihelion passes. Remote radio imaging would give locations as a function of time of flaring and (fast) CMEs in the environment of SPP, yielding a context for understanding solar energetic particle acceleration - one of SPP's key goals. These observations would provide a large scale perspective, complementary to those of the all-sky imager.

Lunar ionosphere: The lunar ionosphere is a dynamic component of the tenuous atmosphere, and one that can have a significant effect on plans to use low frequency imaging arrays to study solar and interplanetary radio sources. The existing data, mainly from dual-frequency radio occultation measurements, suggest that the density of free electrons is highly variable but can exceed 2000 cm^{-3} (Figure 3). The implied local plasma frequency is about 0.4 MHz.

The interpretation of these data is model dependent. Bauer (1996) concluded that the Luna data were consistent with no significant lunar ionosphere. However, ALSEP measurements during Apollo found a photoelectron layer near the surface with electron densities up to 10^4 cm^{-3} (Reasoner & O'Brien 1972), implying a plasma frequency of 0.9 MHz. Such a high plasma frequency would affect radio imaging at frequencies up to a few MHz.

Below the plasma frequency, radio waves cannot propagate. Consequently, as ROLSS tracks solar radio bursts to lower frequencies, it will also be able to search for a lunar ionosphere. If a low frequency cutoff (e.g., around a few MHz) in the radio burst spectra were observed, this would be an indication of lunar ionospheric absorption and would place a constraint on the total electron column of the lunar ionosphere.

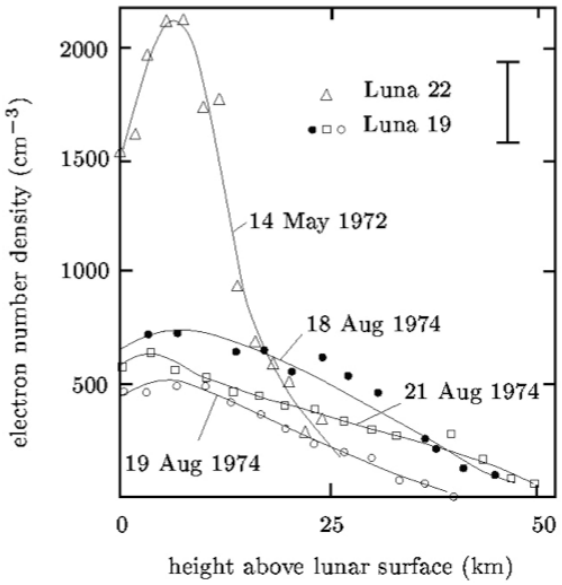


Fig. 3 - Lunar ionosphere electron densities derived from dual-frequency radio occultation measurements during the Luna 19 and Luna 22 missions (Vyshlov 1976; Vyshlov & Savich 1978).

3. Implementation

Overview of the ROLSS concept: The array consists of 3 arms arranged in a Y configuration,

Table 1 – Summary of ROLSS Parameters		
Parameter	**Value**	**Comment**
Wavelength (Frequency)	30–300 m (1–10 MHz)	• Matched to radio emission generated by particle acceleration in the inner heliosphere • Obtain spectra of Galactic and extragalactic sources at longest wavelengths • Probe lunar ionosphere • Operate longward of Earth's ionospheric cutoff
Angular Resolution	2º (at 10 MHz)	• Localize particle acceleration sites in CMEs and Type III solar bursts • Order of magnitude improvement
Bandwidth	100 kHz	Track (time-)evolution of particle acceleration
Lifetime	1 yr	Obtain measurements during several solar rotations

subject to local topographic constraints. Each arm is 500 m long, providing ~2º angular resolution at 30-m wavelength (10 MHz). The arms themselves consist of a polyimide film (PF) on which electrically-short dipole antennas are deposited, and they hold the transmission system for sending the electrical signals back to the central electronics package (CEP), located at the intersection of the arms. The CEP performs the requisite filtering and digitization of the signals, then downlinks them to ground for final imaging and scientific analysis.

The array operates over the wavelength range 30–300 m (1–10 MHz), with a selectable, variable frequency sub-band that can be placed anywhere within the operational wavelength range. During the course of the concept study, the NASA/GSFC Instrument Design Laboratory (IDL) provided an intensive engineering study of the ROLSS concept. We refer to the conclusions of that study as output from the "IDL run."

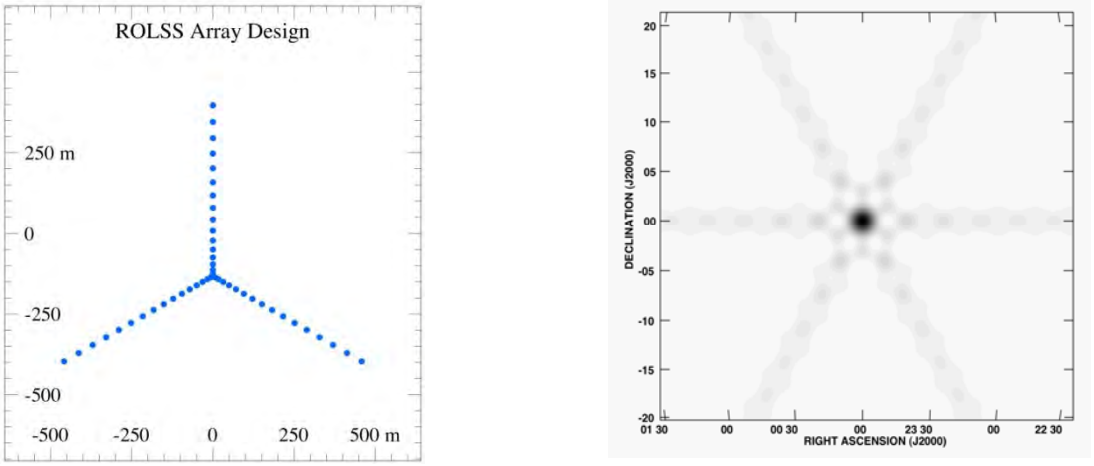


Fig. 4a - (Left) Nominal science antenna distribution along the antenna arms. Fig. 4b – (Right) The resulting point-spread function ("beam") for a snapshot image. The maximum sidelobe is at –5.9 dB and the rms sidelobe level is –15 dB, sufficient to achieve ROLSS goals.

Figure 4a shows a nominal layout of 16 science antennas along each arm. Standard radio astronomical software within the Astronomical Image Processing Software (AIPS) package of the National Radio Astronomy Observatory (NRAO) was used to simulate the instantaneous or "snapshot" point spread

function (PSF) or "beam" of the ROLSS array, assuming 16 antennas on each arm and logarithmic spacing along the arm. The 6-arm "star" pattern of the beam in Figure 4b reflects the Y shape of the ROLSS array. The dynamic range in the beam - defined as the ratio between the peak and the rms level - is 15 dB. This dynamic range is consistent with a simple estimate that the rms level in an interferometric image should be of order 1/*N* if the array consists of *N* antennas. For $N \sim 50$ ($\sim 16 \times 3$), the expected rms level is 2% (-17 dB) of the peak.

Antenna and substrate design: The ROLSS array consists of multiple science antennas. Each science antenna is a single polarization, electrically short dipole, deposited on a PF; also deposited on the PF are the transmission leads to the CEP. The PF is flexible enough to be stored in a roll during transit and deployed directly on the lunar surface by unrolling.

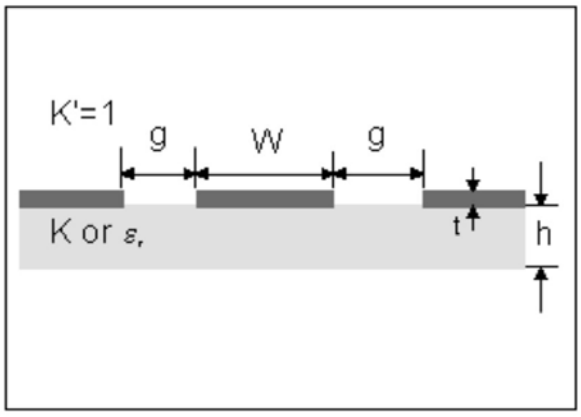


Fig. 5 - Cross section of the transmission line. The light gray indicates the substrate. The dark gray shows the deposited metal. (Figure from IDL run.)

The lunar regolith has a typical relative dielectric permittivity of 2–3 (for lunar soils), high resistivity, and a large skin depth at low radio frequencies. Consequently, it is possible for a low frequency radio antenna to operate when placed directly on the lunar surface. Modeling of the PF antennas suggest using a nominal length of 14 meters. The antenna resonant frequency will be near the high end of the ROLSS observing frequency range. At the low end of this range, the antenna will be electrically short and will have a small feed point resistance and high capacitive reactance. The radiation resistance, efficiency, and fractional bandwidth will all be very small, thereby reducing the sensitivity, but the ROLSS science goals can still be met.

The baseline design has 16 antennas per arm spaced logarithmically in order to provide a range of Fourier spacings for good imaging. The nominal topology is a wide dipole, with ~1.5-m width, which helps provide a broad wavelength response. Each of the dipoles is connected to the CEP at the end of the PF by transmission lines deposited on the film.

Signal transmission to the central hub: Transmission lines carry the signals from the antennas to the CEP. Like the antennas, they consist of a conductor deposited on the PF. Each transmission line is a coplanar waveguide structure (62 mm wide), see Figure 5. Simulations, conducted as part of the IDL run, indicate that the gap between the lines reduces cross coupling and suggest that the lines should be placed symmetrically to reduce antenna pattern distortions. These simulations suggest that ~25 dB of signal loss will occur on the longest transmission lines, which must be compensated for at the receiver. Investigation of using active elements on the PF and alternate signal transmission line designs is ongoing.

The width of the transmission lines is determined in part by the requirement of mitigating micrometeorite damage. During the IDL run, estimates of the micrometeorite damage were determined by scaling from experience with the Hubble Space Telescope (HST) Wide Field/Planetary Camera (WF/PC) radiator. The estimated bombardment rate at the Moon's surface is 4.9 m^{-2} yr^{-1} for impacts producing craters up to 1 mm wide. Larger craters are sufficiently rare that none were seen in a 3.6-yr exposure. The longest transmission lines are likely to receive a few impacts per year (~ 3), but none are likely to be so large that the transmission line will be completely severed.

Preamplifiers, receiver, data processing, and data storage: The ROLSS CEP (shown in Figure 6) houses all of the science antenna receivers including preamplification, data acquisition hardware, power supply electronics, data and telemetry communications electronics, and the thermal management system for these components. The CEP is designed to be carried in a single lunar lander carrier. When deployed,

the CEP stands ~1 m above the lunar surface in order that the thermal management system is not subject to dust contamination.

Data downlink: ROLSS uplink telemetry and commanding would be handled by an S-band uplink. A nominal configuration has a 9-m diameter or larger antenna at a site such as GSFC Wallops. The ROLSS CEP would contain an S-band transceiver, which could also be used for data downlink in emergency situations.

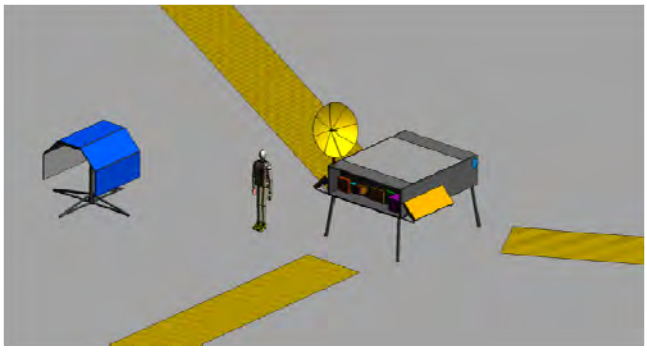


Fig. 6 – ROLSS Central Electronics Package (CEP) showing relative sizes of electronics box with thermal louvers, high-gain antenna, and solar arrays. (Figure from IDL run.)

Data downlink would be handled by a Ka-band system. The ROLSS CEP would host a 0.6-m diameter parabolic gimbaled antenna. The conclusion of the IDL run was that a Ka-band downlink via such an antenna provided the current optimum based on downlink data requirements, weather (rain) conditions, and likely ground-station antenna availability.

Siting and deployment: The prime science mission for ROLSS is solar observations, which favors an equatorial site. A polar site would require that one or more arms of ROLSS be longer than the nominal 500 m length in order to compensate for the foreshortening of the array and maintain the same angular resolution.

The site itself should be of relatively low relief topography (e.g., the surface of a *mare* vs. the highlands). We emphasize that the entire site does not have to be flat as the fraction of the area occupied by the antenna arms is relatively small—for a circular area with a radius of 500 m, the antenna arms occupy only 0.3% of the total area. Further, there is no requirement on the absolute orientation of the arms. Rotating the arms about the center of the array merely rotates the beam pattern on the sky. The *relative* orientation of the arms is required to be $120^{o} \pm 6^{o}$, with the error determined by requiring that the nominal antenna position not vary by more than 1 wavelength at the extreme end of an arm. This level of alignment was demonstrated during the Apollo missions. Only modest requirements are set on the smoothness of the antenna arm locations. The shortest operational wavelength for ROLSS would be 30 m, and an individual science antenna would be 15 m in length. Requiring that the antenna shapes are not distorted by more than $\lambda/10$ implies that deployment requires linear extents of order 500 m with elevation variations not larger than ~3 meters along the intended antenna arm positions. We assume that on the timescale of likely deployment for ROLSS that any site will have better than meter-scale resolution images available (post-Lunar Reconnaissance Orbiter).

The ROLSS deployment itself could be done completely robotically, by astronauts on a crewed rover, or with a mix of these two modes. A completely robotic deployment would allow ROLSS to be deployed prior to any human sortie missions. Scaling from existing rovers, a ROLSS rover could easily be designed to carry and deploy tens of kilograms of antenna mass.

Other ROLSS concept elements: The ROLSS mission concept calls for lunar day operations only, when the solar arrays provide adequate power. There are no nighttime operations. The baseline ROLSS design assumes a solar panel assembly for power generation. Multiple load profiles were generated during the IDL run, with the final load profile suggesting that an average 125 W load (peak 167 W load) could be generated by solar panels. Only daytime operation is assumed, per the prime science mission.

Thermal management is a key requirement for lunar-based instrumentation. The antenna arms are entirely passive and require no thermal management. The CEP is equipped with a set of radiators in order to maintain the internal electronics at a temperature below their assumed maximum operating temperature (80° C). These radiators are equipped with thermal louvers, outfitted with actuators that sense radiator

temperature and rotate thin aluminum blades to open or closed positions. When open, the louvers allow the radiators to dissipate heat.

Other radio imaging options: In this white paper, we have focused on the polyimide film antenna system, but other viable approaches for antennas on the lunar surface are also being studied. Also, a constellation of microsats carrying dipole antennas is an alternative for low-frequency radio imaging (MacDowall et al 2006).

4. Cost Estimate

The IDL run did not involve cost estimation, but did provide masses for the ROLSS components, assuming implementation using present day technology. In the table below, we list the components, estimated masses assuming additional technology development, and ROM cost estimates for a mission similar to the IDL run, with the modification of deployment by robotic rover. This well-documented, but conservative, concept design would be easy to price. Its major weakness is high mass. Alternatively, additional technology development could result in mass savings using advanced techniques like a phased array for the downlink, ultra-low temperature/ultra-low power electronic components, and advanced battery, thermal, and power components. The mass and cost of the rover are likely to be inversely related, with reliability also entering into the trade space.

Component	Mass Estimate (kg)	ROM Cost Estimate (M$)	Details of ROM estimate
Science antenna array, including development	90	10	Primarily development cost
Central Electronics Package	140	12	Development cost
Lithium Ion Batteries	120	5	Hardware cost
CEP Thermal system	25	3	Development & hardware cost
RF/Comm System	25	5	Hardware cost
Solar Panel Assembly	25	2	Actual cost
Deployment rover system	50	20	Hardware & software development + testing
Operations	n/a	12	Labor cost + facilities
Systems Engineering	n/a	4	Labor cost
Project Management	n/a	16	Based on costs for other missions
Total cost excluding LV and lander	475	89	

5. Conclusion

Investigation of particle acceleration and propagation using radio bursts as diagnostic tools has been performed to date without imaging capability below ~20 MHz. A radio observatory such as ROLSS, located on the lunar surface, would provide the first images of radio bursts in the outer solar corona and interplanetary space, yielding new information on solar flaring and shock acceleration of particles. ROLSS observations would be of particular interest to establish the context for perihelion data from the SPP mission. The complement of ROLSS images, white light data from the SPP all-sky imager, and ground-based solar radio observations would constrain the large-scale space weather environment in the vicinity of SPP and sunward of the spacecraft.

References

Bale, S. D., Reiner, M. J., Bougeret, J.-L., et al. 1999, "The Source Region of an Interplanetary Type II Radio Burst," Geophys. Res. Lett. 26, 1573

Bauer, S. J., 1996, "Limits to a Lunar Ionosphere," Sitzungsberichte und Anzeiger, Abt. 2, 133, 17

Bougeret, J.-L., Kaiser, M. L., Kellogg, P. J., et al. 1995, "Waves: The Radio and Plasma Wave Investigation on the Wind Spacecraft," Space Sci. Rev., 71, 231

Cane, H. V., Erickson, W. C., & Prestage, N. P. 2002, "Solar Flares, Type III Radio Bursts, Coronal Mass Ejections, and Energetic Particles," J. Geophys. Res. (Space Physics), 107, 1315

Gopalswamy, N., Yashiro, S., Krucker, S., Stenborg, G., & Howard, R. A. 2004, "Intensity Variation of Large Solar Energetic Particle Events Associated with Coronal Mass Ejections," J. Geophys. Res. (Space Physics), 109, 12105

Gopalswamy, N., Yashiro, S., Kaiser, M. L., Howard, R. A., & Bougeret, J.-L. 2002, "Interplanetary Radio Emission due to Interaction Between Two Coronal Mass Ejections," Geophys. Res. Lett., 29, 080000-1

Gopalswamy, N., Yashiro, S., Kaiser, M. L., Howard, R. A., & Bougeret, J.-L. 2001, "Radio Signatures of Coronal Mass Ejection Interaction: Coronal Mass Ejection Cannibalism?" Astrophys. J., 548, L91

Lara, A., Gopalswamy, N., Nunes, S., Mu˜noz, G., & Yashiro, S. 2003, "A Statistical Study of CMEs Associated with Metric Type II Bursts," Geophys. Res. Lett., 30, 120000-1

MacDowall, R., J., M.D. Desch, et al., 2006, "Microsat and lunar-based imaging of solar radio bursts", *Planetary Radio Astronomy VI, Proc. of the Sixth International Workshop*, H.O Rucker, W.S. Kurth, and G. Mann, eds., Austrian Academy of Sciences Press, Vienna

MacDowall, R. J., A. Lara, P. K. Manoharan, *et al.*, 2003, "Long-duration hectometric type III radio bursts and their association with solar energetic particle (SEP) events," *Geophys. Research Let*., **30** (12), 8018, doi:10.1029/2002GL016624

Reasoner, D. L., & O'Brien, B. J. 1972, J. Geophys. Res., 77, 1292

Vyshlov, A S., 1976, "Preliminary Results of Circumlunar Plasma Research by the Luna 22 Spacecraft," Space Research XVI, Proc. Open Meetings ofWorkshop Groups of Physical Sciences (Akademie-Verlag) p. 945

Vyshlov, A. S., & Savich, N. A., 1978, "Observations of Radio Source Occultations by the Moon and the Nature of the Plasma near the Moon," Cosmic Res., 16 (transl. Kosmicheskie Issledovaniya, 16), 551

Data Mining for Heliophysics

Nikunj C. Oza, Shawn Wolfe, David Thompson, Jeffrey Scargle
NASA Ames Research Center

Introduction

Like many areas of science and engineering, vast quantities of relevant data are available for heliophysical studies. Analysis of this data is essential to increasing our knowledge of the sun. Unfortunately, the volume of data presents unprecedented challenges: no longer can the data be examined manually; the scientists' current techniques to process do not scale to the size; and even the most advanced information technologies are not designed for such volume. New techniques are needed to support the next breakthroughs in heliophysics.

Making the Best Use of the Voluminous Data

A key issue in data mining for heliophysics is how to find regions (or events) in the data that are worthy of closer scrutiny, instead of analyzing all of the data. This can be organized by the degree to which the scientist can express desiderata to the information system. We present these in order of decreasing precision.

Database queries

If the scientist knows what s/he is looking for in precise terms, then database technology can be employed. Some preliminary work has been done in the area of efficient multivariate time series search on massive databases. However, such work needs to be enhanced in several ways.

1. The amount of data available is so massive that even seemingly narrow queries may return a substantial amount of matching data. This amount of data may be too high to be sent over computer networks from the data repository to the user. This requires a search capability that can run entirely on the system that hosts the data repository, which needs to be powerful enough to handle queries from many scientists.
2. On the other hand, seemingly broad queries may not return any matches. This is because scientists often want to search for phenomena on the edge of human knowledge. When searching on the fringes, it is easy to step over the boundary into a regime where no data has been gathered. Tools that assist scientists in understanding and visualizing the data are needed to remedy this problem.
3. The indices for the database must be identified and created. In many cases, it is not sufficient or practical to create indices only in terms of the raw data measurements (e.g., magnetic flux). Instead, higher order events (e.g., solar

flare) are needed. This may require progressive processing to produce higher and higher level concepts, as well an ability to name events as they become recognized in the literature. It may also be challenging to determine the most reusable (across investigations) and valuable (essential for specific investigations) indices.

4. Along with processing to produce higher level concepts, it is important to identify relationships among these concepts, and to be able to query the database based on these semantic relationships. This enables the scientist to use a more abstract query language to interrogate the data based on the data context and, and to detect data that satisfy complex contextual constraints.

Ad hoc retrieval

When the scientist has some idea of what is wanted, but cannot specify it exactly, ad hoc retrieval is appropriate. This may be because of several reasons: the scientist has a precise data need, but must approximate it in the search language of the system; the scientist is thinking of an idealized event which does not exist, but would like to see things similar to it; or the scientist only has a rough idea of what might be interesting. With the growth of the internet, search engines have become very sophisticated and capable of dealing with enormous volume. Nonetheless, several challenges remain.

1. Internet search engines are intended to work with textual content, and make assumptions well-suited for text. These same assumptions do not hold for data. Indeed, the core concept of content is ill-defined in scientific data. New relevance models must be developed that are appropriate for the unique properties of data.
2. Likewise, the search interfaces built for textual content are not appropriate for data. Can the query be specified in the familiar "search box", is something else needed? How can each matching result be summarized so the scientist can tell what is of interest? Should the matching data be retrieved as requested, or should it be prefetched, or should the entire search process be interactive?
3. New paradigms for searching may also need to be explored. Given the data volume and the complexity of the search parameters, it may not be possible to provide results on the fly. One possibility is to dynamically populate or refine result list, or to compute it offline. A complementary technique is filtering, which dynamically monitors incoming data for items matching the scientists' interests.
4. Methods by which the user's search criteria and results can be visualized and interactively examined are needed.

Anomaly detection

The desire to search on the fringes of human knowledge can be assisted by methods that automate the process of finding anomalous data – potentially unknown,

interesting phenomena. These methods are able to operate without requiring input from the scientist because they look for anomalies using properties of the data themselves. For example, density-based anomaly detection methods assume that normal data tend to form clusters of many points that presumably represent common modes and that anomalous data tend to appear in low density regions. Distance-based anomaly detection methods assume that normal data has neighbors that are relatively nearby whereas anomalous data points have relatively distant neighbors. Once these methods identify anomalies based on the data, the scientist can check if the anomalies are scientifically significant. Anomaly detection methods could be improved on several fronts:

1. As mentioned earlier, anomaly detection methods assume that normal data points and abnormal data points have certain characteristics (e.g., high density vs. low density). Techniques to combine multiple anomaly detection methods may focus the scientist on the most anomalous data points (according to multiple models), which are presumably the most interesting.
2. Anomaly detection methods often return too many false positives, increasing the amount of uninteresting data that the scientist must examine. More complex specifications of normality from the scientist could greatly reduce these false positives, for instance, when encountering spatiotemporal phenomena and changing data.
3. Efficiency concerns and the segmentation of very large datasets over multiple machines points to distributed solutions. However, most anomaly detection methods cannot operate in a distributed manner, and not all that do guarantee to return the same anomalies as what the centralized algorithm would return. How to parallelize these algorithms remains an open question.
4. Just as with retrieval, with anomaly detection, interactive visualization of user's inputs and results is needed.

Next Steps

Advances in database query, ad hoc retrieval and anomaly detection technology for massive datasets should also enable studying data across missions. Currently, most data mining is done within the context of a single mission. Working with data from multiple missions will greatly increase the scope of work and the potential to gain new knowledge.

There are concrete steps that can be taken in the near term, of which two are the following.

1. There is the need to look at data mining efforts within other fields (such as particle physics and Earth sciences) to draw lessons learned.
2. There is the need to identify some concrete challenge problems for which we can examine the literature to determine what has been done, identify what needs to be done, and then find ways (e.g., funding vehicles) to get that work done. We selected on the problem of devising a (universal?) metric of

similarity for Heliophysics events out of several possibilities. The main difficulty in this problem is identifying the (accepted?) key features that represent a pattern within the space of Heliophysics events.

THEORETICAL RESEARCH ON SOLAR WIND TURBULENCE

A White Paper Submitted to the NRC Decadal Survey of Solar and Space Physics

Benjamin D. G. Chandran, Eliot Quataert, Jean Perez, Aveek Sarkar, Steve Cranmer, John Podesta, & Prateek Sharma

Fluctuations in the flow velocity, magnetic field, and electric field are ubiquitous in the solar wind.[13,36] These fluctuations are turbulent, in the sense that they are disordered and span a broad range of scales in both space and time. The study of solar wind turbulence has been motivated by a number of factors. Not only are these fluctuations a fundamental feature of the solar wind, but they may hold the key to understanding the SW's origin and thermodynamics, as described in section I below. Moreover, an understanding of solar wind turbulence will help astronomers to understand magnetized, astrophysical plasmas more generally, as described in section IV.

Research into solar wind turbulence and magnetohydrodynamic (MHD) turbulence has seen a number of seminal developments since the 1960s. For example, Iroshnikov and Kraichnan showed that Alfvén waves (AWs) interact nonlinearly with one another only when they propagate in opposite directions in the plasma rest frame.[19,23] Thus, AWs propagating away from the Sun do not interact with one another, but they do interact with AWs propagating towards the Sun. As shown by Barnes, at wavelengths exceeding the proton gyroradius and in collisionless plasmas, AWs are virtually undamped while fast and slow magnetosonic waves undergo significant damping.[3] As a result, solar wind turbulence likely consists primarily of non-compressive Alfvénic fluctuations. In addition, when counter-propagating AWs interact, energy cascades rapidly to larger $k_\perp$ but only weakly to larger $k_\parallel$, where $k_\perp$ and $k_\parallel$ are the wavevector components perpendicular and parallel to the background magnetic field $\mathbf{B}_0$.[11,12,29,33]

In the following sections, we describe several questions and research areas in which further research promises to be particularly fruitful during the period 2013-2023 covered by the National Research Council Decadal Survey of Solar and Space Physics. We conclude in section V with some policy recommendations for strengthening the nation's space physics theory program.

I Does Turbulence Play an Important Role in Generating the Solar Wind?

The origin of the solar wind is one of the central problems in the field of space physics. Early studies by Parker and others demonstrated that some type of non-thermal energy flux is required in order to explain the large solar wind speeds and large proton temperatures that are measured near Earth.[14,24,30] However, the form of this non-thermal energy flux and the way in which it interacts with thermal plasma are still not understood.

One possibility is that AWs power the solar wind by transporting energy from the solar surface to the solar atmosphere and then on into the solar wind, as illustrated schematically in Fig. 1. Alfén waves can be launched from the Sun in a number of ways, including the continual shaking of the the footpoints of open magnetic field lines by convective motions in the photosphere. Some of the outward propagating waves reflect due to the spatial variation in the Alfvén speed, and the interaction between outward-propagating waves and inward-propagating waves causes AW energy

to cascade from large scales to small scales. At sufficiently small scales (comparable to the proton gyroradius), the AWs dissipate, heating the ambient plasma and accelerating the solar wind.

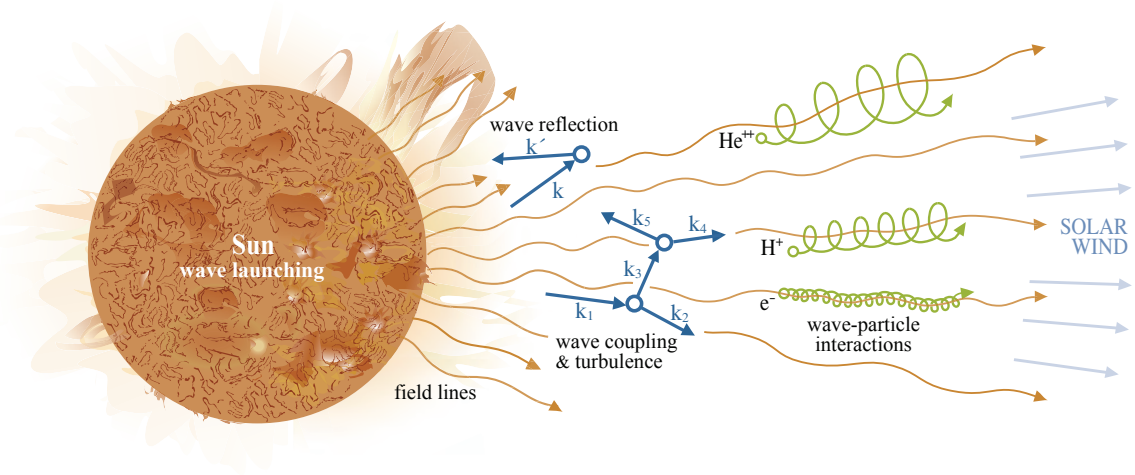


Figure 1: The possible role of waves and turbulence in generating the solar wind.

The above scenario has been the subject of numerous studies,[8,27,37,39] and is consistent with a wide range of observations, including *Hinode* observations of spicule motions,[9] ground-based observations of wave motions in the corona,[34,35] Faraday-rotation observations of magnetic fluctuations in the corona and inner solar wind,[16] and measurements of wind speeds, temperatures, and wave amplitudes near Earth.[8,39] On the other hand, there are at least two possible problems with this scenario:

1. It is not clear that AW turbulence can explain the perpendicular ion heating that is observed in coronal holes[1,10,22] and low-β fast-solar-wind streams.[15,26] The traditional explanation for perpendicular ion heating is via a cyclotron resonance between ions and high-frequency Alfvén/ion-cyclotron waves.[17] However, anisotropic turbulence theories and simulations show that AW energy remains at comparatively low frequencies as it cascades to small scales, so that a cyclotron resonance does not arise.[12,18,33]

2. Satellite observations show that the solar wind undergoes substantial heating and acceleration within a few solar radii of the Sun. It is not clear that the energy of the large-wavelength AWs launched by the Sun can cascade to small scales rapidly enough to explain this near-Sun heating.

A combination of theoretical and computational investigations will be needed to address these problems, including direct numerical simulations of AW turbulence in the inhomogeneous solar atmosphere as well as analytical and numerical investigations of stochastic heating and cyclotron heating by AW turbulence. These types of investigations are discussed further in the sections below. The importance of solving the above problems is underscored by the preparations underway for NASA's Solar Probe Plus (SPP) Mission. The primary science objective of this landmark mission will be to uncover the origin of the solar wind. Substantial progress in our understanding of the

turbulence and its dissipation within the solar wind will be needed to interpret the measurements that SPP will provide and maximize the returns from SPP's historic visit to the Sun.

II Inertial Range Physics

Wave energy is injected into the solar wind (e.g., by the Sun) at a "stirring scale" or "outer scale" that greatly exceeds the length scales at which the fluctuations dissipate. The range of scales intermediate between the outer scale and the dissipation scale in a turbulent system is known as the inertial range. A widely adopted hypothesis in AW turbulence research is that fluctuations at inertial-range length scales are only weakly influenced by the details of the large-scale forcing or small-scale dissipation, and instead evolve primarily due to interactions among intertial-range structures. The result of such interactions is a cascade of AW energy from large scales to small scales. The inertial-range energy cascade is of critical importance in the solar wind because it controls the rate at which fluctuation energy is dissipated, and hence the rate at which turbulence heats the ambient plasma. The inertial range energy cascade also controls the anisotropy of the fluctuations in wavenumber space.

Over the last decade, a focus of turbulence research has been to determine the inertial range power spectrum of homogeneous, incompressible, MHD turbulence. Although a great deal of progress has been made on this problem, its relevance to the solar wind is limited, because the solar wind is neither homogeneous nor incompressible. A priority for the coming decade will be to extend our theoretical understanding of the inertial range by accounting for the following key effects.

Inhomogeneity and Wave Reflection. As mentioned in section I, some of the outward propagating AWs launched by the Sun are reflected by the gradient in the Alfvén speed, providing the mix of outward-propagating AWs and inward-propagating AWs that is needed to generate an AW energy cascade. As shown by Velli, Verdini, and others, AW turbulence driven by wave reflection differs fundamentally from AW turbulence driven by mechanical forcing, because wave reflections enhance the coherence between the inward and outward AW fields.[37,38] Direct numerical simulations of this type of turbulence will be critical for determining whether the AW energy cascade is fast enough to dissipate AW energy close to the Sun, before the AW energy escapes far into the solar wind.

Cross Helicity and "Imbalanced Turbulence." Cross helicity is the correlation between the velocity and magnetic field vectors in a turbulent flow. In the solar wind, it measures the excess of outward propagating AWs over inward propagating AWs. MHD or AW turbulence with nonzero cross helicity is also referred to as "imbalanced." Solar wind turbulence in the inner heliosphere is imbalanced, with an excess of outward-propagating waves. Most of the points of agreement within the community about the inertial range of MHD turbulence are restricted to "balanced" turbulence with zero cross helicity. There are now at least five mutually contradictory theories of imbalanced turbulence in the literature, and also several mutually inconsistent numerical studies of imbalanced turbulence.[4,5,25,31,32] A priority for turbulence research in the coming decade will be to delineate the different regimes of applicability (if any) of these different theories and simulations, and to

determine the correct theoretical framework(s) for understanding turbulence in the solar wind.

Pressure Anisotropy. As the solar wind flows away from the Sun, its expansion acts to decrease the plasma temperature. However, because the solar wind is nearly collisionless, expansion has a different effect on thermal motions parallel and perpendicular to the magnetic field, which are measured by the parallel and perpendicular temperatures $T_\parallel$ and $T_\perp$. Between the outer corona and the Earth, expansion reduces $T_\perp$ to a greater degree than it reduces $T_\parallel$. As a result, much of the solar wind plasma near 1 AU satisfies $T_\perp < T_\parallel$. However, if $T_\perp/T_\parallel$ becomes too small, then the plasma becomes unstable to firehose instabilities. Measurements from the *Wind* spacecraft show that the values of $T_\perp/T_\parallel$ found in the solar wind are approximately bounded from below by the firehose-instability threshold.[2,21] Presumably, the excitation of (kinetic) firehose instabilities acts to limit temperature anisotropy, preventing the plasma from migrating far past the marginal stability point. Even when the plasma is stable, temperature anisotropy leads to fundamental changes in AWs and other wave modes in a plasma, which may in turn have important consequences for nonlinear interactions between waves, and hence for solar wind turbulence. To understand solar wind turbulence in detail, it will thus be necessary to move beyond traditional fluid descriptions to account for temperature anisotropy in both theoretical analyses and numerical simulations.

III Dissipation of Alfvén-Wave Turbulence

As AW energy cascades from large scales to small scales, it eventually reaches length scales comparable to the proton gyroradius, at which point the waves begin to dissipate. Because collisions are very weak in the solar wind and corona, the small-scale waves dissipate via collisionless processes rather than viscosity or resistivity. The collisionless dissipation of the AW cascade at small scales is not yet understood, but is critical for understanding the thermodynamics of the solar wind. For example, collisionless dissipation determines the fraction of the turbulent heating power that goes to each particle species, as well as the fractions that go to particle motions parallel to the magnetic field and particle motions perpendicular to the magnetic field. As mentioned in section I, one of the critical unsolved problems for the origin of the solar wind is whether the dissipation of low-frequency AW or kinetic Alfvén wave (KAW) turbulence can explain the perpendicular ion heating that is seen in the corona and low-β fast solar wind streams. Previous theoretical investigations suggest that low-frequency AW/KAW turbulence in the low-β corona may dissipate through "stochastic heating" of ions, in which low-frequency fluctuations in the electrostatic potential distort ion gyromotion and lead to perpendicular heating.[6,7,20,28,40] Further theoretical and numerical investigations are needed to test this suggestion, and to explore collisionless dissipation of AW turbulence more generally.

IV Broader Impact within Astrophysics

The study of heliospheric plasma physics can have a tremendous impact on a wide range of astrophysics problems. This is because the heliospheric environment is a critical laboratory in which physical processes can be studied in detail that are important in many astrophysical systems. In

doing so, we believe that it will be critical to draw on the understanding of key physical processes observed and studied in detail in the solar corona and the broader heliospheric environment. Historically, this has been common: many ideas first developed in the solar context were later applied extensively to astrophysics problems. For example, the theory of the spindown of the sun by magnetic fields was later developed into models for jets from accretion disks around black holes and neutron stars.

Current studies of turbulence in the solar wind are impacting and/or have the potential to impact astrophysics modeling in a number of ways. We highlight two examples here:

1. There are many low collisionality astrophysical systems, including some accretion disks onto black holes, the intracluster medium of galaxy clusters, winds from other stars, outflows from accreting neutron stars and black holes, and highly energetic particles in the interstellar medium of galaxies. Pitch angle scattering by turbulent fluctuations self-generated from pressure anisotropies and velocity streaming (e.g., firehose, mirror) are believed to make the physics of these low collisionality systems more fluid-like, but this is not well understood in detail. Studying these fluctuations and their role in the dynamics of the solar wind will provide critical insight into the physics of a wide variety of low-collisionality astrophysical plasmas.

2. The turbulent heating of electrons, protons, and ions has been directly measured in the solar wind near Earth (and in the corona using spectroscopic diagnostics). These data are still not fully understood, but are very important for certain models of accretion onto black holes, where the same science question is the biggest uncertainty in our theoretical modeling. As the data and theory improve in the solar context, this will thus enable us to make more realistic models of black hole accretion. This is particularly important for the black hole at the center of our own Galaxy, which is the best case for a black hole in astrophysics.

V Recommendations

Theoretical investigations, including the types of studies described in this White Paper, will be essential if the community is to reap the full science return from existing and future missions. We thus strongly recommend increased support to NASA's Heliophysics Theory Program, the NSF SHINE Program, and NASA's Solar and Heliospheric Physics Supporting Research Program. In addition, just as the overall space science program is optimized when there is a balanced portfolio of experimental, observational, and theoretical studies, the theory program itself is optimized when there is a balance between numerical simulations and analytical theory. The support for computer simulations in space physics research has grown steadily over the past two decades, but support for analytical theory has declined to a level that is perilously low. We recommend that program officers within NASA and NSF help maintain a healthy mix of numerical and analytical research in the field by viewing the balance between analytical and numerical research as a programmatic factor when selecting proposals for funding.

References

[1] E. Antonucci, M. A. Dodero, and S. Giordano, "Fast solar wind velocity in a polar coronal hole during solar minimum," *Solar Physics*, **197**, pp. 115–134, 2000.

[2] S. D. Bale, J. C. Kasper, G. G. Howes, E. Quataert, C. Salem, and D. Sundkvist, "Magnetic Fluctuation Power Near Proton Temperature Anisotropy Instability Thresholds in the Solar Wind," *Physical Review Letters*, **103**, pp. 211101–+, 2009.

[3] A. Barnes, "Collisionless damping of hydromagnetic waves," *Physics of Fluids*, **9**, pp. 1483–1495, 1966.

[4] A. Beresnyak and A. Lazarian, "Structure of Stationary Strong Imbalanced Turbulence," *Astrophysical Journal*, **702**, pp. 460–471, 2009.

[5] B. D. G. Chandran, "Strong Anisotropic MHD Turbulence with Cross Helicity," *Astrophysical Journal*, **685**, pp. 646–658, 2008.

[6] B. D. G. Chandran, "Alfvén-wave Turbulence and Perpendicular Ion Temperatures in Coronal Holes," *Astrophysical Journal*, **720**, pp. 548–554, 2010.

[7] B. D. G. Chandran, B. Li, B. N. Rogers, E. Quataert, and K. Germaschewski, "Perpendicular Ion Heating by Low-frequency Alfvén-wave Turbulence in the Solar Wind," *Astrophysical Journal*, **720**, pp. 503–515, 2010.

[8] S. R. Cranmer, A. A. van Ballegooijen, and R. J. Edgar, "Self-consistent Coronal Heating and Solar Wind Acceleration from Anisotropic Magnetohydrodynamic Turbulence," *Astrophysical Journal Supplement*, **171**, pp. 520–551, 2007.

[9] B. De Pontieu, S. W. McIntosh, M. Carlsson, V. H. Hansteen, T. D. Tarbell, C. J. Schrijver, A. M. Title, R. A. Shine, S. Tsuneta, Y. Katsukawa, K. Ichimoto, Y. Suematsu, T. Shimizu, and S. Nagata, "Chromospheric Alfvénic Waves Strong Enough to Power the Solar Wind," *Science*, **318**, pp. 1574–7, 2007.

[10] R. Esser, S. Fineschi, D. Dobrzycka, S. R. Habbal, R. J. Edgar, J. C. Raymond, J. L. Kohl, and M. Guhathakurta, "Plasma properties in coronal holes derived from measurements of minor ion spectral lines and polarized white light intensity," *Astrophysical Journal Letters*, **510**, pp. L63–L67, 1999.

[11] S. Galtier, S. V. Nazarenko, A. C. Newell, and A. Pouquet, "A weak turbulence theory for incompressible magnetohydrodynamics," *Journal of Plasma Physics*, **63**, pp. 447–488, 2000.

[12] P. Goldreich and S. Sridhar, "Toward a theory of interstellar turbulence. 2: Strong Alfvénic turbulence," *Astrophysical Journal*, **438**, pp. 763–775, 1995.

[13] M. L. Goldstein, D. A. Roberts, and W. H. Matthaeus, "Magnetohydrodynamic Turbulence In The Solar Wind," *Annual Reviews of Astronomy and Astrophysics*, **33**, pp. 283–326, 1995.

[14] R. E. Hartle and P. A. Sturrock, "Two-fluid model of the solar wind," *Astrophysical Journal*, **151**, p. 1155, 1968.

[15] P. Hellinger, P. Trávníček, J. C. Kasper, and A. J. Lazarus, "Solar wind proton temperature anisotropy: Linear theory and WIND/SWE observations," *Geophysical Research Letters*, **33**, pp. 9101–+, 2006.

[16] J. V. Hollweg, S. R. Cranmer, and B. D. G. Chandran, "Coronal faraday rotation fluctuations and a wave/turbulence-driven model of the solar wind," *Astrophysical Journal*, **722**, pp. 1495–1503, 2010.

[17] J. V. Hollweg and P. A. Isenberg, "Generation of the fast solar wind: A review with emphasis on the resonant cyclotron interaction," *Journal of Geophysical Research (Space Physics)*, **107**, pp. 1147–+, 2002.

[18] G. G. Howes, S. C. Cowley, W. Dorland, G. W. Hammett, E. Quataert, and A. A. Schekochihin, "A model of turbulence in magnetized plasmas: Implications for the dissipation range in the solar wind," *Journal of Geophysical Research (Space Physics)*, **113**, pp. 5103–+, 2008a.

[19] P. S. Iroshnikov, "Turbulence of a Conducting Fluid in a Strong Magnetic Field," *Astronomicheskii Zhurnal*, **40**, pp. 742–+, 1963.

[20] J. R. Johnson and C. Z. Cheng, "Stochastic ion heating at the magnetopause due to kinetic Alfvén waves," *Geophysical Research Letters*, **28**, pp. 4421–4424, 2001.

[21] J. C. Kasper, A. J. Lazarus, and S. P. Gary, "Wind/SWE observations of firehose constraint on solar wind proton temperature anisotropy," *Geophysical Research Letters*, **29**, pp. 170000–1, 2002.

[22] Kohl, J., et al., "UVCS/SOHO empirical determinations of anisotropic velocity distributions in the solar corona," *Astrophysical Journal Letters*, **501**, p. L127, 1998.

[23] R. H. Kraichnan, "Inertial-range spectrum of hydromagnetic turbulence," *Physics of Fluids*, **8**, p. 1385, 1965.

[24] E. Leer and T. E. Holzer, "Energy addition in the solar wind," *Journal of Geophysical Research*, **85**, pp. 4681–4688, 1980.

[25] Y. Lithwick, P. Goldreich, and S. Sridhar, "Imbalanced strong MHD turbulence," *Astrophysical Journal*, **655**, pp. 269–274, 2007.

[26] E. Marsch, X.-Z. Ao, and C.-Y. Tu, "On the temperature anisotropy of the core part of the proton velocity distribution function in the solar wind," *Journal of Geophysical Research (Space Physics)*, **109**, pp. 4102–+, 2004.

[27] W. H. Matthaeus, G. P. Zank, S. Oughton, D. J. Mullan, and P. Dmitruk, "Coronal heating by magnetohydrodynamic turbulence driven by reflected low-frequency waves," *Astrophysical Journal Letters*, **523**, pp. L93–L96, 1999.

[28] J. M. McChesney, R. A. Stern, and P. M. Bellan, "Observation of fast stochastic ion heating by drift waves," *Physical Review Letters*, **59**, pp. 1436–1439, 1987.

[29] C. S. Ng and A. Bhattacharjee, "Interaction of shear-Alfvén wave packets: implication for weak magnetohydrodynamic turbulence in astrophysical plasmas," *Astrophysical Journal*, **465**, p. 845, 1996.

[30] E. N. Parker, "Dynamical theory of the solar wind," *Space Science Reviews*, **4**, p. 666, 1965.

[31] J. C. Perez and S. Boldyrev, "Role of Cross-Helicity in Magnetohydrodynamic Turbulence," *Physical Review Letters*, **102**, pp. 025003–+, 2009.

[32] J. J. Podesta and A. Bhattacharjee, "Theory of Incompressible Magnetohydrodynamic Turbulence with Scale-dependent Alignment and Cross-helicity," *Astrophysical Journal*, **718**, pp. 1151–1157, 2010.

[33] J. V. Shebalin, W. Matthaeus, and D. Montgomery, "Anisotropy in MHD turbulence due to a mean magnetic field," *Journal of Plasma Physics*, **29**, p. 525, 1983.

[34] S. Tomczyk and S. W. McIntosh, "Time-Distance Seismology of the Solar Corona with CoMP," *Astrophysical Journal*, **697**, pp. 1384–1391, 2009.

[35] S. Tomczyk, S. W. McIntosh, S. L. Keil, P. G. Judge, T. Schad, D. H. Seeley, and J. Edmondson, "Alfvén Waves in the Solar Corona," *Science*, **317**, pp. 1192–, 2007.

[36] C. Tu and E. Marsch, "MHD structures, waves and turbulence in the solar wind: Observations and theories," *Space Science Reviews*, **73**, pp. 1–210, 1995.

[37] M. Velli, R. Grappin, and A. Mangeney, "Turbulent cascade of incompressible unidirectional Alfven waves in the interplanetary medium," *Physical Review Letters*, **63**, pp. 1807–1810, 1989.

[38] A. Verdini, M. Velli, and E. Buchlin, "Turbulence in the Sub-Alfvénic Solar Wind Driven by Reflection of Low-Frequency Alfvén Waves," *Astrophysical Journal Letters*, **700**, pp. L39–L42, 2009.

[39] A. Verdini, M. Velli, W. H. Matthaeus, S. Oughton, and P. Dmitruk, "A Turbulence-Driven Model for Heating and Acceleration of the Fast Wind in Coronal Holes," *Astrophysical Journal Letters*, **708**, pp. L116–L120, 2010.

[40] Y. Voitenko and M. Goossens, "Cross-field heating of coronal ions by low-frequency kinetic Alfvén waves," *Astrophysical Journal Letters*, **605**, pp. L149–L152, 2004.