

A MISSION TO STUDY THE COUPLING OF ATMOSPHERIC REGIONS BY PRECIPITATING ENERGETIC PARTICLES

Cora E. Randall, University of Colorado

Scott M. Bailey, Virginia Tech

Charles H. Jackman, Goddard Space Flight Center

A White Paper Submitted to the 2010 Heliophysics Decadal Survey

INTRODUCTION

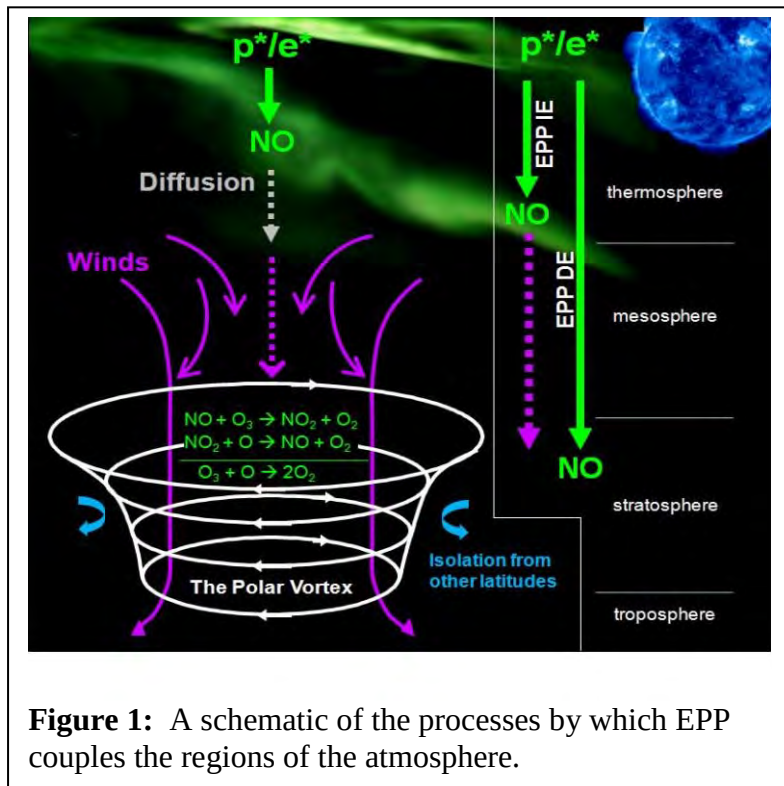
Energetic Particle Precipitation (EPP) is a primary mechanism that couples the different layers of the Earth's atmosphere. In particular, energetic particles ionize the neutral atmosphere, triggering the production of reactive odd nitrogen ($\text{NO}_x = \text{NO} + \text{NO}_2$) and odd hydrogen (HO_x). Both NO_x and HO_x catalytically destroy O_3 [e.g., Weeks et al. 1972; Heath et al. 1977; Solomon et al. 1983; Jackman et al. 1990; Lary, 1997], which can affect middle atmosphere temperatures and circulation [Gabriel et al., 2007, and references therein]. Variations in the temperature structure of the middle atmosphere can be communicated to the lowermost atmosphere, impacting weather and climate [Baldwin et al., 2007; Haynes, 2005]. In turn, variations in the troposphere, such as might originate from anthropogenic climate change, can be communicated to the middle and upper atmosphere [e.g., Fomichev et al., 2007], thereby affecting the response to EPP.

The pathways by which the coupling described above occurs have not yet been adequately elucidated and are ultimately linked to heliophysical processes that control both EPP and the basic dynamical and chemical structure of the Earth's atmosphere. A comprehensive description of Sun-Earth connections requires quantification of the atmospheric coupling pathways relevant to EPP. **A NASA mission is required to explore and understand the atmospheric response to energetic particles, specifically the coupling of atmospheric regions, as the particle energy is redistributed via dynamical, chemical, and radiative processes.**

Although both HO_x and NO_x are of interest, it is most imperative to determine the atmospheric effects of NO_x produced by EPP. Whereas the HO_x produced in the mesosphere by EPP does cause short-term O_3 reductions [Solomon et al., 1981], HO_x -induced changes do not persist beyond the initial days after particle precipitation. We refer to NO_x produced by energetic particle precipitation as EPP- NO_x .

PRODUCTION OF NO_x BY PRECIPITATING ENERGETIC PARTICLES

Production of EPP- NO_x varies spatially and temporally according to the energy, flux, and origin of the precipitating energetic particles. For example, precipitating protons with energies greater than about 30 MeV lead to direct stratospheric NO_x production [e.g., Crutzen et al., 1975; Reid, 1986]. Such very high-energy proton precipitation occurs sporadically, on the order of several times per solar cycle, and is generally associated with solar eruptions. It also tends to occur fairly uniformly over the polar cap (geomagnetic latitude $> 60^\circ$) [e.g., Reagan et al., 1981; Jackman et al., 2001]. Because the EPP- NO_x is produced directly in the stratosphere, it is immediately available for catalytic destruction of stratospheric O_3 . This process is thus referred to as the **direct effect** of EPP on the stratosphere (EPP DE).



production in the lower thermosphere near 100-110 km. Such precipitation is generally associated with auroral and ring current processes and occurs routinely, although with varying intensity, over the solar cycle [e.g., Solomon, 1991 and references therein]. The EPP- NO_x resulting from this low-energy precipitation is distributed over the auroral and subauroral regions. Particles with energies between the two extremes described above will lead to NO_x production in the mesosphere, with varying spatial and temporal distributions. For EPP- NO_x in the mesosphere/lower thermosphere (MLT) to influence stratospheric O_3 distributions, it must be transported downward; this process is thus referred to as the **indirect effect** of EPP (EPP IE). As described below, efficient transport to the stratosphere occurs only during or near the polar winter.

Quantifying the EPP IE and DE requires knowledge of the EPP- NO_x distributions from the stratosphere to the lower thermosphere throughout the polar winter. At the current time, observations of NO_x during the polar night are available only in the stratosphere and lower mesosphere [e.g., Lopez-Puertas et al., 2005; Seppala et al., 2004]. Thus, to date, the complete distribution of EPP- NO_x during the polar winter has never been measured and represents a fundamental gap in our understanding of the pathways to atmospheric coupling.

Measurements of NO_x in the polar night are a key requirement for any mission to study EPP effects on the atmosphere.

It is important to link NO_x production to specific precipitating particle populations and through these to the space weather disturbances that produced them. Such knowledge will ultimately be needed to predict the solar cycle variation of EPP- NO_x -related O_3 loss. The coupling between space weather and stratospheric meteorology places spatial and seasonal constraints on O_3 -depleting NO_x production; thus, certain types of space weather disturbances may be more effective than others because of their characteristic spatial and temporal variations. For example, new observations indicate that auroral precipitation makes a significant contribution to NO_x descending into the stratosphere in the southern hemisphere [c.f., Randall et al., 2007]. In fact, peaks in descending NO_x are well correlated with peaks in auroral energy input. However, these

From analyses of fairly recent solar energetic particle (SEP) events [Jackman et al., 2009], it has been estimated that the EPP DE contributes 0.01 to 5.5 Gmoles of NO_y (total odd nitrogen, including NO, NO_2 , N_2O , N_2O_5 , HNO_3 , ClONO_2 , etc.) to the stratosphere annually with highest amounts occurring near solar maximum [e.g., Vitt and Jackman, 1996]. The largest stratospheric source of NO_y results from the oxidation of nitrous oxide $\{\text{N}_2\text{O} + \text{O}(^1\text{D}) \rightarrow 2\text{NO}\}$ and amounts to about 52-58 Gmoles NO_y annually. Thus, the EPP DE in recent decades would contribute a maximum of about 10% to the annual total production of NO_y .

On the other extreme, precipitating particles at much lower energies lead to NO_x

peaks did not all occur in years of high solar activity at the maximum phase of the solar cycle; some occurred in the descending phase in association with peaks in high-speed solar wind stream activity [Kozyra et al., 2006]. This raises the question of whether the intense, expanded auroral precipitation pattern during day-long major magnetic storms or the contracted, weaker auroral precipitation pattern associated with long-lived (up to 10-day duration) high speed solar wind streams is more effective at creating NO_x that will ultimately enter the stratosphere. In fact, for major magnetic storms, if the precipitation pattern extends well equatorward of the polar vortex, a significant portion of the NO_x that is produced may ultimately be transported to lower latitudes and be destroyed by photodissociation, reducing the magnetic storm impact on O_3 variability. A further complication is introduced by the fact that high speed streams also produce elevated fluxes in the outer radiation belts; thus, in addition to long-lived periodic auroral activity, enhanced radiation belt electron precipitation also contributes to NO_x production and O_3 loss.

TRANSPORT OF NO_x

Quantifying the EPP IE requires knowledge of the amount of EPP- NO_x descending into the stratosphere and thus an understanding of how EPP- NO_x produced in the MLT is redistributed. The relevant processes are illustrated in Figure 1, which depicts a polar winter situation where NO_x created by EPP in the polar lower thermosphere will rapidly diffuse downward to the mesopause at a rate of ~ 0.05 m/s. The large-scale meridional circulation is such that air in the mesosphere is transported from the summer to the winter hemisphere and descends at the winter pole [Brewer, 1949]. As air descends through the mesosphere, it can be transported to lower latitudes by horizontal winds; if EPP- NO_x in the MLT is transported to sunlit latitudes, it will be photodissociated with a lifetime of hours to days [Barth, 1992; Bailey et al., 2002]. Thus horizontal transport to lower latitudes represents a loss process for EPP- NO_x .

Air that is not transported to lower latitudes will descend into the lower mesosphere, at some point encountering the polar vortex. A strong vortex will act to confine any EPP- NO_x in the descending air to polar latitudes, increasing the probability that it reaches the stratosphere. Estimates of descent velocities in the polar winter range from about 13 km/month in the upper mesosphere [Forkman et al., 2005] down to less than 1 km/month in the lower stratosphere [Juckes et al., 2007]. Engel et al. [2005] estimated that the fraction of mesospheric air in a layer near 25 km at the end of winter was between 35% and 100%. Once in the stratosphere, the NO_x photochemical lifetime increases to months-years, which thus defines the timescale over which it will participate in catalytic O_3 destruction.

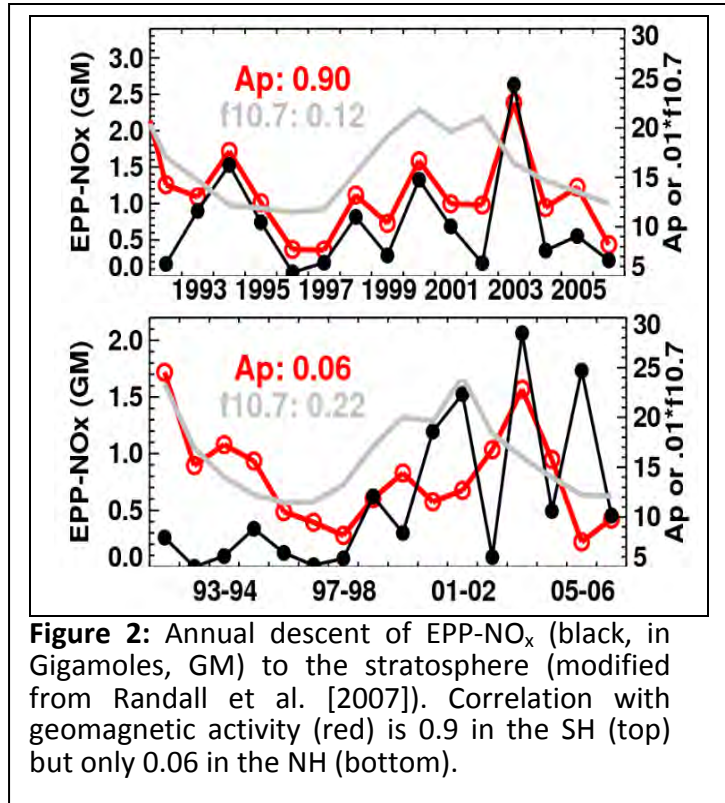
Ideally, the processes depicted in Figure 1 would be observed. Historically, however, most observations applied to such investigations have required sunlight, precluding observations of NO_x transport in the polar night. The most dramatic example of the EPP IE in recent years occurred in Mar-Apr of 2004, when upper stratospheric NO_x mixing ratios increased by more than a factor of four and were accompanied by O_3 reductions of up to 60% [Natarajan et al., 2004; Randall et al., 2005; Rinsland et al., 2005]. Because of a lack of data in the polar night MLT, however, we still do not know whether or what fraction of the NO_x enhancements were caused by NO_x created during periods of high geomagnetic activity in Oct-Dec 2003 or during more quiescent times in late Dec-Jan.

Indirect estimates of the amount of EPP-NO_x descending into the stratosphere over the last solar cycle have large uncertainties because they are based on solar occultation data, which is both limited in geographic extent and restricted to sunlit periods. Without nighttime data, we do not know the altitude at which most of the EPP-NO_x originated. Nevertheless, by relying on measurements of atmospheric tracers to identify air with NO_x enhancements that originated in the MLT, Randall et al. [2007] estimated that EPP-NO_x descending into the Southern Hemisphere (SH) stratosphere contributes up to about 40% (10%) of the total annual source of NO_x in the polar (global) stratosphere. The other 60% (10%) can ultimately be traced back to oxidation of N₂O originating in the troposphere.

Randall et al. [2007] also showed that the interannual variability in the amount of EPP-NO_x descending to the SH stratosphere since 1992 is well-correlated with variations in the Ap index, with a correlation coefficient of 0.9; similar excellent correlations were found with other measures of geomagnetic activity, including auroral and medium-energy hemispheric power. This is shown in Figure 2 (top) and suggests that variations in EPP-NO_x entering the SH stratosphere are controlled by variations in the MLT NO_x production. While this might seem obvious, the same situation does not hold true for the Northern Hemisphere (NH), results of which are shown in the bottom panel of Figure 2. The correlation coefficient between EPP-NO_x reaching the NH stratosphere and the Ap index is only 0.06. The poor correlation arises because interannual variations in transport are at least as important for the NH as are variations in EPP-NO_x production. For example, 3 to 6 times more EPP-NO_x was observed descending into the NH stratosphere in Feb-Mar 2006 than ever before, aside from spring 2004 [Randall et al., 2006]. Yet EPP itself was very low in the preceding months, suggesting that only nominal production of EPP-NO_x had occurred in the MLT. Siskind et al. [2007] showed that unusual wave activity that winter forced enhanced descent in the mesosphere, explaining the large amount of EPP-NO_x descending into the stratosphere even in the absence of excess production. A similar situation prevailed in 2009 [Randall et al., 2009].

Clearly, we must quantify the extent to which horizontal and vertical transport control EPP-NO_x distributions and thus the significance of changing meteorology on the contribution of EPP-NO_x to stratospheric odd nitrogen.

The transport of EPP-NO_x from the MLT to the stratosphere must be observed and quantified. These observations must be made in the polar night, where most of the transport is expected to occur.



THE IMPACT OF EPP-NO_x ON THE MIDDLE ATMOSPHERE 4

Although substantial work has been done in recent decades to understand stratospheric O₃, we still cannot identify all of the sources contributing to its chemical loss. The key missing element is knowledge of NO_x produced by EPP, since this has never before been measured and tracked as it is transported from its production region to the stratosphere. The NO_x catalytic cycle is the primary O₃ loss mechanism from ~24-40 km [Garcia and Solomon, 1994]; NO_x also interferes with halogen-induced O₃ loss at lower altitudes.

Dramatic O₃ loss of 30 to 60% in the middle to upper polar stratosphere has been attributed to the EPP IE, as shown in Figure 3 [Randall et al., 1998; 2001; 2005]. The two panels on the left from Randall et al. [2005] show NO_x (top) and O₃ (bottom) mixing ratios at an altitude of 40 km for every measurement made by the Halogen Occultation Experiment (HALOE) in the NH from 1991-2003 (purple) and 2004 (green). The large excursions in 2004 were shown by Randall et al. [2005] to result from EPP-NO_x. Note that the enhancements resulted in up to 60% reductions in O₃. Panels on the right compare SH O₃ measured by the Polar Ozone and Aerosol Measurement (POAM) instrument in 1994 and 1996 (top; [Randall et al., 1998]) and in 2000 and 1999 (bottom; [Randall et al., 2001]) near an altitude of 30 km. Randall et al. [1998; 2001] showed that EPP-NO_x enhancements at this altitude were very small in 1996 and 1999, but were ~10 ppbv in 1994 and 2000; this correlates well with the much larger O₃ loss in 1994 and 2000. These variations are suggestive, but not conclusive.

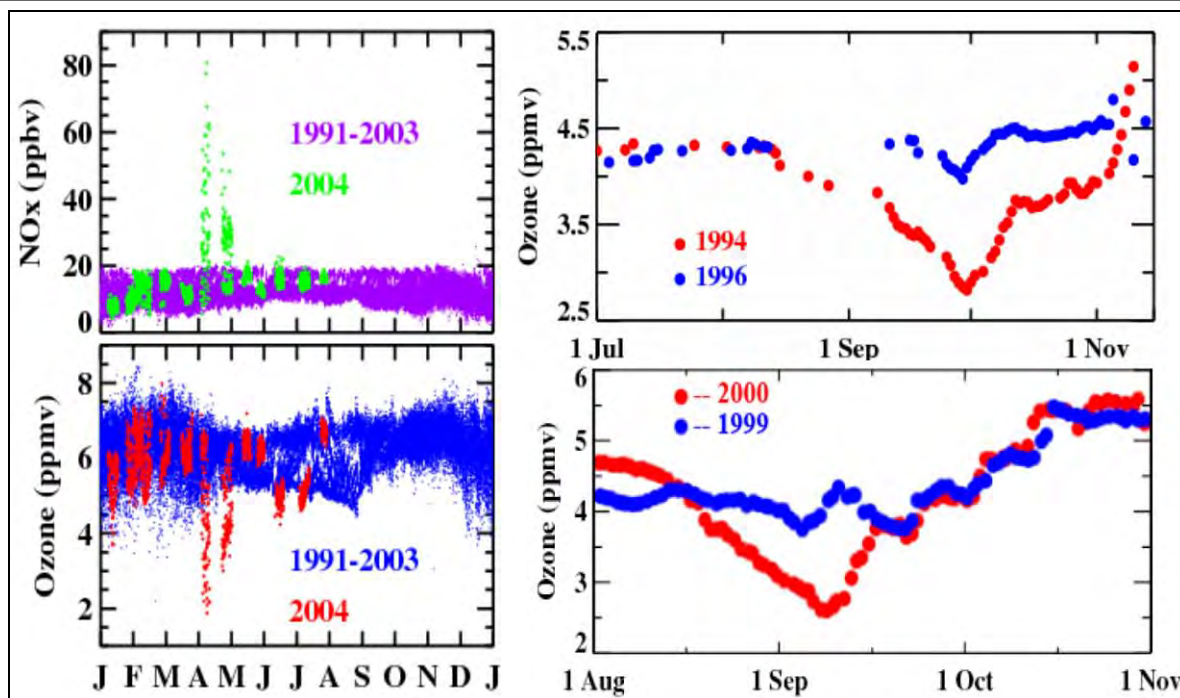


Figure 3: Substantial O₃ loss occurs when EPP-NO_x descends into the stratosphere (see text).

We must quantify for the first time the contribution to stratospheric O₃ loss from EPP-NO_x and relate the loss back to the heliophysical processes controlling the NO_x production.

SUMMARY

With the advent of sophisticated general circulation models that couple atmospheric and geospace regions, and with proven advanced remote sensing technologies at our disposal, we are poised for major new discoveries in Sun-atmosphere coupling. We are ready to explore and understand the atmospheric response to energetic particle precipitation, specifically the coupling of atmospheric regions as the particle energy is redistributed via dynamical, chemical, and radiative processes. In particular, we are ready to explore the remaining unexplored region of the MLT, the polar night. We strongly suggest that NASA pursue such a mission. The mission would have as its primary observations measurements of nighttime NO_x and the transport of that NO_x from the MLT to the stratosphere.

References

- Bailey, S. M., C. A. Barth, and S. C. Solomon, A model of nitric oxide in the lower thermosphere, *JGR*, v.107, n. A8, 1205, 10.1029/2001JA000258, 2002.
- Barth C.A., Nitric-Oxide in the Lower Thermosphere, *Planetary and Space Science* 40 (2-3): 315-336, 1992.
- Brewer, A. W., Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, *Quart. J. Roy. Meteor. Soc.*, 75, 351-363, 1949.
- Crutzen, P. J., I. S. A. Isaksen, and G. C. Reid, Solar proton events: Stratospheric sources of nitric oxide, *Science*, 189, 457-459, 1975.
- Engel, A., et al., On the observation of mesospheric air inside the arctic stratospheric polar vortex in early 2003, *Atmos. Chem. Phys. Discuss.*, 5, 7457-7496, 2005.
- Fomichev, V.I., et al., Response of the middle atmosphere to CO₂ doubling: results from the Canadian Middle Atmosphere Model, *J. Climate*, vol. 20, 1121-1144, 2007.
- Forkman P., P. Eriksson, D. Murtagh, P. Espy (2005), Observing the vertical branch of the mesospheric circulation at latitude 60°N using ground-based measurements of CO and H₂O, *J. Geophys. Res.*, 110, D05107, doi:10.1029/2004JD004916.
- Gabriel, A., D. Peters, I. Kirchner, and H-F. Graf, Effect of zonally asymmetric ozone on stratospheric temperature and planetary wave propagation, *Geophys. Res. Lett.* 34, L06807, doi:10.1029/2006GL028998, 2007.
- Garcia, R.R., and S. Solomon, A new numerical model of the middle atmosphere, 2, ozone and related species, *J. Geophys. Res.* 99, 12,937-12,951, 1994.
- Haynes, P., Stratospheric Dynamics, *Annu. Rev. Fluid Mech.* 37, 263-293, doi:10.1146/annurev.fluid.37.061903.175710, 2005.
- Heath, D. F., A. J. Krueger, and P. J. Crutzen, Solar proton event: influence on stratospheric ozone, *Science*, 197, 886-889, 1977.
- Jackman, C. H., et al., Effect of solar proton events on the middle atmosphere during the past two solar cycles as computed using a two-dimensional model, *J. Geophys. Res.*, 95, 7417-7428, 1990.
- Jackman, C.H., et al., Northern hemisphere atmospheric effects due to the July 2000 solar proton event, *Geophys. Res. Lett.*, 28, 2883-2886, 2001.
- Jackman, C.H., et al., Long-term middle atmospheric influence of very large solar proton events, *J. Geophys. Res.* 114, D11304, doi:10.1029/2008JD011415, 2009.
- Juckes, M. N., An annual cycle of long lived stratospheric gases from MIPAS, *Atmos. Chem. Phys.* 7 (7): 1879-1897, 2007.

Kozyra, J. U., et al., Response of the upper/middle atmosphere to coronal holes and powerful high-speed solar wind streams in 2003, in *Recurrent Magnetic Storms: Corotating Solar Wind Streams*, edited by B. T. Tsurutani, et al., AGU monograph, 2006.

Lary, D.J., Catalytic destruction of stratospheric ozone, *J. Geophys. Res.* 102, 21,515-21,526, 1997.

López-Puertas, M., et al., Observation of NO_x enhancement and ozone depletion in the Northern and Southern Hemispheres after the October-November 2003 solar proton events, *J. Geophys. Res.*, 110, A09S43, doi:10.1029/2005JA011050, 2005.

Natarajan, M., E.E. Remsberg, L.E. Deaver, and J.M. Russell, III, Anomalously high levels of NO_x in the polar upper stratosphere during April, 2004: Photochemical consistency of HALOE observations, *Geophys. Res. Lett.* 31, doi:10.1029/2004GL020566, 2004.

Randall, C. E. et al, Stratospheric effects of energetic particle precipitation in 2003-2004, *Geophys. Res. Lett.*, 32, L05802, doi:10.1029/2004GL022003, 2005.

Randall, C. E., V. L. Harvey, C. S. Singleton, P. F. Bernath, C. D. Boone, and J. U. Kozyra, Enhanced NO_x in 2006 linked to strong upper stratospheric Arctic vortex, *Geophys. Res. Lett.*, 33, L18811, doi:10.1029/2006GL027160, 2006.

Randall, C. E., et al., Energetic particle precipitation effects on the Southern Hemisphere stratosphere in 1992-2005, *J. Geophys. Res.*, 112, D08308, doi:10.1029/2006JD007696, 2007.

Randall, C.E., et al., NO_x descent in the Arctic middle atmosphere in early 2009, *Geophys. Res. Lett.* 36, L18811, doi:10.1029/2009GL039706, 2009.

Reagan, J.B., et al., Effects of the August 1972 solar particle events on stratospheric ozone, *J. Geophys. Res.*, 86, 1473-1494, 1981.

Reid, G. C., Chapter 12 – Solar energetic particles and their effects on the terrestrial environment, in *Physics of the Sun*, Vol. III, edited by P. A. Sturrock, pp. 251-278, D. Reidel Publishing Company, 1986.

Rinsland, C.P., et al., Atmospheric Chemistry Experiment (ACE) Arctic stratospheric measurements of NO_x during February and March 2004: Impact of intense solar flares, *Geophys. Res. Lett.* 32, L16S05, doi:10.1029/2005GL022425, 2005.

Siskind, D. E., S. D. Eckermann, L. Coy, J. P. McCormack, and C. E. Randall, On recent interannual variability of the Arctic winter mesosphere: Implications for tracer descent, *Geophys. Res. Lett.*, 34, L09806, doi:10.1029/2007GL029293, 2007.

Solomon, S., et al., The effect of particle precipitation events on the neutral and ion chemistry of the middle atmosphere: 2. Odd hydrogen, *Planet. Space Sci.*, 29, 885– 892, 1981.

Seppala, A., et al., Solar proton events of October–November 2003: Ozone depletion in the Northern Hemisphere polar winter as seen by GOMOS/Envisat, *Geophys. Res. Lett.*, 31, L19107, doi:10.1029/2004GL021042, 2004.

Solomon, S., G. C. Reid, D. W. Rusch, and R. J. Thomas, Mesospheric ozone depletion during the solar proton event of July 13, 1982, 2, Comparison between theory and measurements, *Geophys. Res. Lett.*, 10, 257-260, 1983.

Solomon, S.C., Optical Aeronomy, *Rev. Geophys. Supplement* (U.S. National Report to International Union of Geodesy and Geophysics 1987-1990), 1089-1109, 1991.

Vitt, F. M., and C. H. Jackman, A comparison of sources of odd nitrogen production from 1974 through 1993 in the Earth's middle atmosphere as calculated using a two-dimensional model, *J. Geophys. Res.*, 101, 6729-6739, 1996.

Weeks, L. H., R. S. CuiKay, and J. R. Corbin, Ozone measurements in the mesosphere during the solar proton event of 2 November 1969, *J. Atmos. Sci.*, 29, 1138-1142, 1972.

The Importance of Distributed Measurements of the Ionospheric Electric Field for Advancement of Geospace System Science and Improved Space Weather Situational Awareness

Joseph B.H. Baker, Virginia Tech

William A. Bristow, University of Alaska, Fairbanks

Raymond A. Greenwald, Virginia Tech

J. Michael Ruohoniemi, Virginia Tech

Simon G. Shepherd, Dartmouth College

Elsayed Talaat, The Johns Hopkins University Applied Physics Laboratory

Summary

As we move toward the next solar maximum there are strong scientific and technological imperatives for improving our understanding of the processes that lead to severe magnetosphere-ionosphere disturbances and associated space weather hazards. The electric field is one of the most important measurable parameters in the magnetosphere-ionosphere system because it drives currents and the bulk motion of plasma. Continuous, multipoint observations of the electric fields and plasma convection within the coupled magnetosphere-ionosphere system are a necessary component of any strategy to fully understand the geospace system response to solar wind forcing. Progress in the study of the near-Earth space environment will increasingly rely on the availability of ionospheric measurements on global scales. In particular, ground-based radars are an effective tool for monitoring the earthward boundary of the Sun-Earth connection. Widespread measurements of ionospheric electric fields from ionospheric radars are valuable to the space physics community because they: (1) enable validation of physics-based magnetosphere-ionosphere models; (2) are an important input for data assimilative techniques for ionospheric space weather specification; (3) provide a valuable ground context for ongoing and future spacecraft and rocket missions; (4) provide a framework for interpretation of other more localized ground-based measurements; (5) foster international scientific cooperation; and (6) allow students to gain hands-on experience with large geoscience engineering projects and therefore provide training for the next generation of geospace scientists and engineers. We recommend that networks of ground-based radars continue to be developed to facilitate further progress on the most pressing outstanding questions in space science, and enhance the capabilities for space weather specification and situational awareness.

Background

From the solar atmosphere outwards through interplanetary space and into the Earth's magnetosphere, ionosphere and upper atmosphere, matter is largely found in an ionized state known as plasma. Over the years, scientific spacecraft have made many *in situ* observations of the plasmas of the solar wind, as well as those of the different regions of the Earth's magnetosphere and ionosphere. From these data sets, and additional data sets obtained with sounding rockets, scientific balloons, and ground-based instrumentation, the average properties of the solar wind and different regions of the Earth's magnetosphere and ionosphere have been determined. We now know that mass and energy flow outward from the sun in the form of enhanced plasma density, flow velocity, energetic particles, and magnetic and electric fields. We know that the character of this solar wind can change dramatically over short time scales, particularly in response to major solar disturbances such as flares and coronal mass ejections. We know that the disturbed solar wind interacts strongly with the Earth's magnetosphere and ionosphere, severely distorting both of these regions and creating conditions that can damage spacecraft and commercial power grids, disrupt communications and navigation systems, as well as harm astronauts in low Earth orbit and spacecraft inside the inner magnetosphere. Finally, we have learned that almost all disturbances in the solar wind produce coherent global changes to the Earth's magnetosphere and ionosphere system. What we do not understand fully is how these coherent global changes come about or

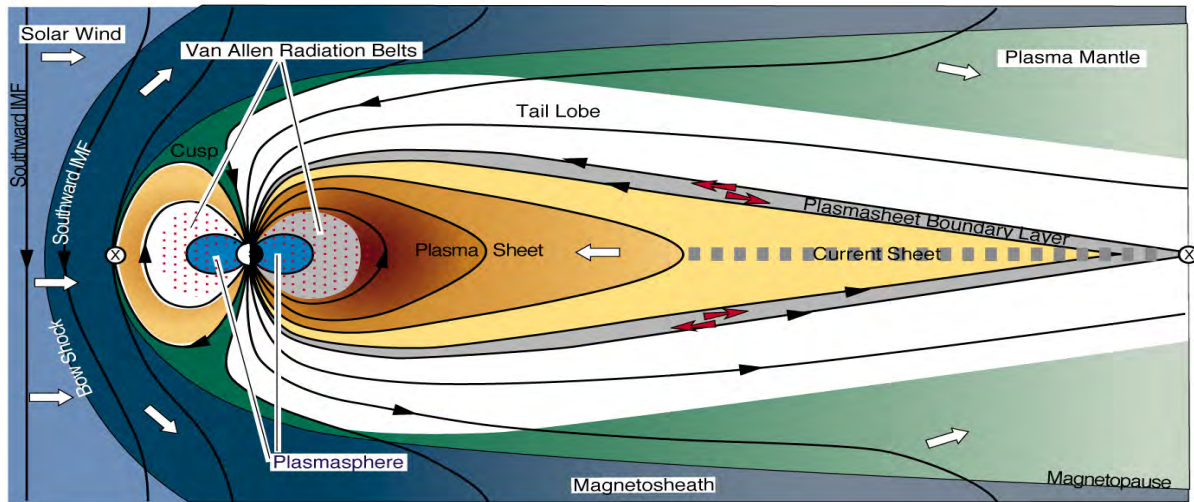


Fig. 1 Meridional cut through the Earth’s magnetosphere revealing the magnetic topology (black lines), primary plasma domains, and large scale convection of plasma (white arrows). Ground-based radars at high latitudes make observations at the footpoints of magnetic field lines that extend into the outer magnetosphere, whereas radars at middle latitudes make observations pertinent to the Van Allen Radiation Belts.

what determines the intensity of their impacts on the magnetosphere-ionosphere system. Figure 1 shows a diagram of the magnetosphere and identifies some of the dominant plasma regions.

Obtaining improvement in our understanding of geomagnetic disturbances requires widely distributed measurements of the near-Earth space environment during changing conditions in the solar wind. These measurements can be made *in situ* with spacecraft or by using remote sensing techniques from the Earth’s surface. It should be emphasized that the two approaches are complementary and both ground-based and space-based measurements are required to achieve the desired goal of improving our physical understanding of the geospace environment. Whether obtained from the ground or in space, the electric field is one of the most important measurable parameters in the magnetosphere-ionosphere system because it drives currents and the bulk motion of plasma. For the most part, geomagnetic field lines can be treated as electrostatic equipotentials because charged particles can easily spiral along the magnetic field and equalize any potential difference that might arise between its footpoints in the conjugate ionospheres. The net result is to produce a common electrostatic potential structure for the entire magnetic flux tube and the plasma contained on the flux tube moves collectively. This means that ground-based measurements of ionospheric plasma motion can be effectively used to infer the pattern of plasma convection further out in space. This is particularly true at middle to high latitudes where the magnetic field lines map outward to the furthest reaches of the magnetosphere and some of them are actually connected to the interplanetary magnetic field. In particular, by using ground-based radars to observe the plasma drift velocity in the ionosphere it is possible to “image” the large-scale circulation of plasma over vast volumes of the outer magnetosphere (see Figure 2). Looking forward, it is essential that we develop a better physical understanding of the fundamental plasma processes that are generated in the magnetosphere-ionosphere system during space weather disturbances. This can only be achieved with widely distributed measurements of ionospheric electric fields obtained by ground-based radars.

Ionospheric Radar Systems

Broadly speaking, there are two primary radar techniques for probing electric fields in the ionosphere. The first is the incoherent scatter radar (ISR) technique that uses high power signals to scatter from thermally-excited clusters of electrons in the ionosphere. A combination of advanced signal processing and fitting to empirical models is then used to obtain a full suite of parameters, including ion drift velocity; electron and ion densities and temperatures; as well as neutral atmosphere properties, such as the

thermospheric wind. Examples of ISRs include the NSF-funded facilities at Sondre Stromfjord, Greenland; Millstone Hill, Massachusetts; Jicamarca, Peru; and Arecibo, Puerto Rico. The second technique is coherent scatter radar, in which the radar signals are backscattered from instability-excited irregularities in ionospheric plasma structure, rather than thermal fluctuations. The Doppler motion of the irregularities is then used to infer the ion drift velocity and electric field. The best example of the coherent radar technique is the high frequency (HF) radars of the Super Dual Auroral Radar Network (SuperDARN). SuperDARN is an international collaborative radar network with funding provided by more than a dozen countries. At the present time there are more than 24 radars in operation worldwide. The coherent radar technique has two significant advantages over ISR that are particularly relevant for collaborations with spacecraft missions: (1) it requires relatively low power and can run continuously at low cost, and (2) the radar range is significantly larger because HF signals experience refraction in the ionosphere and can be ducted into propagation modes extending to several thousand kilometers. Doppler measurements from multiple SuperDARN radars are routinely combined using an assimilative technique to generate instantaneous maps of the hemispheric pattern of ionospheric convection (see Figure 2), which are widely used by space scientists to study the physics of the coupled magnetosphere-ionosphere system, as well as by representatives from aerospace corporations and government agencies for space weather specification.

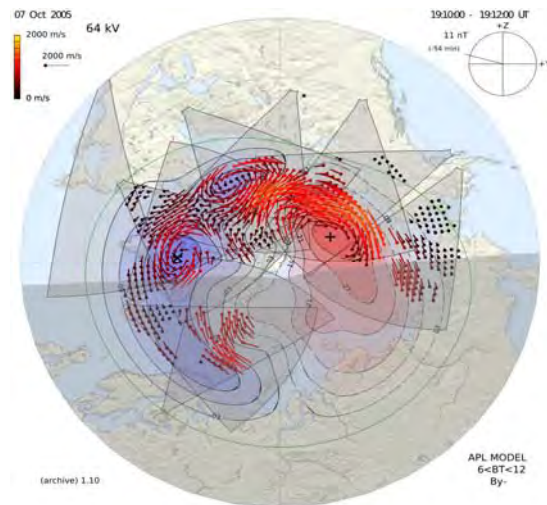


Figure 2: A SuperDARN space weather convection pattern. The north magnetic pole is at the center, the sun is directed upward, and the outer boundary is at 60° magnetic latitude.

It should be emphasized that the two radar systems are very much complementary: incoherent scatter radars provide a comprehensive dataset of ionosphere-thermosphere parameters (with appropriate assumptions) but over a relatively localized region; whereas coherent radars provide less information about the ionosphere-thermosphere parameters but obtain measurements of the ionospheric electric field over a more widely distributed spatial area and with continuous operation.

Current Status of Ionospheric Radar Infrastructure

In recent years there have been several new additions to the complement of radars for ionospheric research. First, there has been the deployment of the Advanced Modular Incoherent Scatter Radar (AMISR) systems at Poker Flat, Alaska, and Resolute Bay, Canada. AMISR is the first incoherent scatter radar to employ phased array technology, and as such, represents the next generation of ISRs. Not to be outdone, the Europeans are in the planning stages for constructing their own advanced ISR facility, the European incoherent scatter imaging radar (EISCAT-3D) system, which will comprise a distributed network of phased array instruments.

The past 5 years has also seen significant expansion in the SuperDARN network. First generation SuperDARN radars were built near latitudes of ~60° magnetic latitude to maximize the coverage of the auroral zone during quiet to moderate levels of geomagnetic activity. In recent years the network has been expanding to higher and lower latitudes with the construction of two PolarDARN radars in the polar cap (Rankin Inlet and Inuvik) and seven radars at middle latitudes in Virginia, Kansas, Oregon, and Japan. In the southern hemisphere, there have been recent installations of SuperDARN radars in the polar cap at McMurdo, Antarctica, and at mid-latitudes in the Falkland Islands.

These recent expansions of ionospheric radar infrastructure into the polar cap and middle latitudes are

providing new opportunities to study: (1) the dynamics of ionospheric flows in the highest latitudes of the polar cap, (2) the expansion of geomagnetic effects from high-latitudes to middle latitudes during magnetic storms, and (3) the nature of electric fields in the subauroral ionosphere and inner magnetosphere during quiet geomagnetic conditions.

Geospace Science Activities Enabled by Distributed Electric Field Measurements

A full explanation of the influence of the Sun on the near-Earth space environment requires continuous monitoring of the fundamental processes responsible for solar-terrestrial coupling. In particular, there should be monitoring of global current systems and plasma flows; the energization and loss of energetic magnetospheric particles; and the transport of mass, energy, and momentum throughout the global magnetosphere-ionosphere-thermosphere system. The magnetosphere-ionosphere electric field plays a major role in all of these interactions, and the most efficient method for obtaining widely distributed measurements is to use ground-based radars.

The distributed dataset of ionospheric electric fields provided by ground-based radars allows researchers to continually monitor solar-terrestrial processes with high temporal resolution over a wide variety of spatial scales. To fully understand the geospace system we need to understand the external driving influences (e.g. the solar wind and interplanetary magnetic field) but we also need to understand the system response of the magnetosphere-ionosphere-thermosphere (M-I-T). This will require a wide range of datasets, including ground-based radars. Recent evidence suggests that the system response may vary depending on the pre-existing state (i.e. preconditioning). Advancement of geospace science requires the coupling between the components of the M-I-T system to be examined over the full spectrum of activity levels, from quiet to very disturbed. Again, ground-based radar measurements are vital to this effort because of their global reach and “always on” capability.

Ground-based radar measurements are also valuable for advancing the capabilities of physics-based numerical models and as inputs for data assimilative techniques. The past decade has seen the development of sophisticated numerical models in the context of the NSF Geospace Environment Modeling (GEM) and Center for Integrated Space-weather Modeling (CISM) programs and the NASA Community Coordinated Modeling Center (CCMC). There have also been significant advancements in data assimilation techniques applied to the ionosphere-thermosphere system. A key component of the GEM effort is an ongoing dialog between modelers and data providers. The models provide a means of interpreting distinctive features that may arise in the datasets, and the data can be used to effectively test the model performance and make model refinements. Distributed measurements of ionospheric electric fields have played a large role in these efforts and will continue to do so.

In addition, the relatively low cost of a coherent scatter radar system, when compared to a typical spacecraft mission, allows researchers from around the globe to participate in space science research at relatively low cost. There are also substantial educational benefits provided by ground-based radar projects. They provide hands-on exposure to space science and associated engineering for graduate students and post-doctoral scientists, thus training the geospace principal investigators of the future.

The Relationship of Ground-Based Radars to Other Geospace Instruments

As mentioned previously, ground-based and space-based measurements are very much complementary to each other, and there are numerous examples of spacecraft missions that have included a robust ground-based instrument program at the earliest stages of planning. Recent examples include the NASA TIMED (Thermosphere Ionosphere Mesosphere Energetics and Dynamics) spacecraft, the NASA THEMIS (Time History of Events and Macroscale Interactions During Substorms) mission, and the European Space Agency (ESA) Cluster-II mission. At the present time, ionospheric radar measurements are also contributing to new space-based research initiatives, such as the NSF AMPERE (Active Magnetosphere and Planetary Electrodynamics Response Experiment) project, and the NASA RBSP (Radiation Belt Storm Probes) program, among others. Radar measurements of ionospheric electric fields are also

routinely used to provide context to rocket missions and have in some cases, been used to determine whether a particular launch should proceed on a given night or be delayed for more geophysically active conditions. Looking forward to the future, distributed measurements of ionospheric electric fields from ground-based radars will continue to complement and extend the capabilities of the next generation of space-based research instruments by providing a global context within which to understand the in situ observations.

Similarly, ground-based radar measurements provide a context for interpreting other more localized ground-based measurements. Mesoscale maps of the ionospheric convection electric field can be valuable for characterizing the wider geomagnetic state of activity, similar to large-scale maps of the auroral distribution obtained from space-based imaging systems like the NASA Polar spacecraft Ultraviolet Imager (UVI) and Visible Imaging System (VIS), and the NASA IMAGE (Imager for Magnetopause-to-Aurora Global Exploration) spacecraft Far Ultraviolet Imager (FUV). Over the past two decades there have been significant advancements in instrument development for studying the magnetosphere-ionosphere system from the ground. For the most part, these advancements have been driven by needs to understand different aspects of solar-terrestrial science at the subsystem level. For example, advances in atmospheric lidar technologies have provided new improved means for monitoring neutral temperatures and winds in the middle atmosphere. In recognition of these instrument developments, the previous Decadal Survey for Solar and Space Physics recommended that the next major ground-based instrumentation initiative for space science research should be the deployment of Distributed Arrays of Small Instruments (DASI). It was anticipated that these arrays would provide continuous real-time observations with the resolution needed to resolve mesoscale phenomena and their dynamic evolution. It was expected that affordable high-speed computing would facilitate the consolidation of these local deployments into elements of a global instrument system.

Ten years later, the DASI concept has not yet been realized. However, recent expansion in ionospheric radar infrastructure (in particular, the newer mid-latitude SuperDARN radars located in the continental United States) may provide a unifying framework to compare measurements between instruments, and stimulate renewed interest in the DASI concept. For many years, the NSF Geospace Facilities program has fostered clustering of geophysical instruments in the vicinity of incoherent scatter radars. This approach provides significant cost savings to small instrument PIs by virtue of leveraging the operations costs. Another advantage is the improved opportunities for doing multi-instrument event studies based on the collocated measurements. A similar approach could be applied to the new mid-latitude SuperDARN radars. DASI instruments could be clustered at locations close to the “sweet-spot” for maximum backscatter returns in the SuperDARN radar fields-of-view. Some of these DASI clusters could be located at minority educational institutions, similar to the model adopted for the NASA THEMIS EPO (Education and Public Outreach) ground-based magnetic observatories. The wider geophysical context provided by the distributed electric field measurements from the SuperDARN radars would provide a basis for the various DASI PIs in the cluster to inter-compare their measurements with each other, and perhaps with other DASI clusters located inside other radar fields-of-view.

Recommendations

In summary, it is recommended that the report of the Decadal Survey should include language that:

- Highlights the important role that ground-based radars should play in an integrated research strategy for advancing understanding of the Sun-Earth Geospace system;
- Reaffirms the Distributed Arrays of Small Instruments (DASI) concept with ground-based radars providing a unifying context within which DASI measurements can be interpreted and compared;
- Urges NASA to consider including ionospheric radars as a ground component when planning for upcoming spacecraft missions;
- Nurtures the education of graduate students and training of early career scientists and engineers at ionospheric radar facilities.

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. Omid, 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 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 $\sim$ 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 $L \sim 4$. 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 plasmopause, 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 plasmopause 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 ~ 10 s 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 from 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 performed 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>).

Understanding Tropospheric Influences on the Mesosphere/Thermosphere/Ionosphere Region

Authors: Dr. Rebecca Bishop, The Aerospace Corporation,
Dr. Doug Rowland, Goddard Space Flight Facility
Dr. William Farrell, Goddard Space Flight Facility

During the extraordinarily quiet recent solar minimum period, the influence of lower altitude forcing on the thermosphere/ionosphere region was clearly observed. Observational evidence implies that the troposphere can couple electrically, dynamically, and chemically to the upper atmospheric regions. Figure 1 shows a sampling of different forcing sources such as volcanic

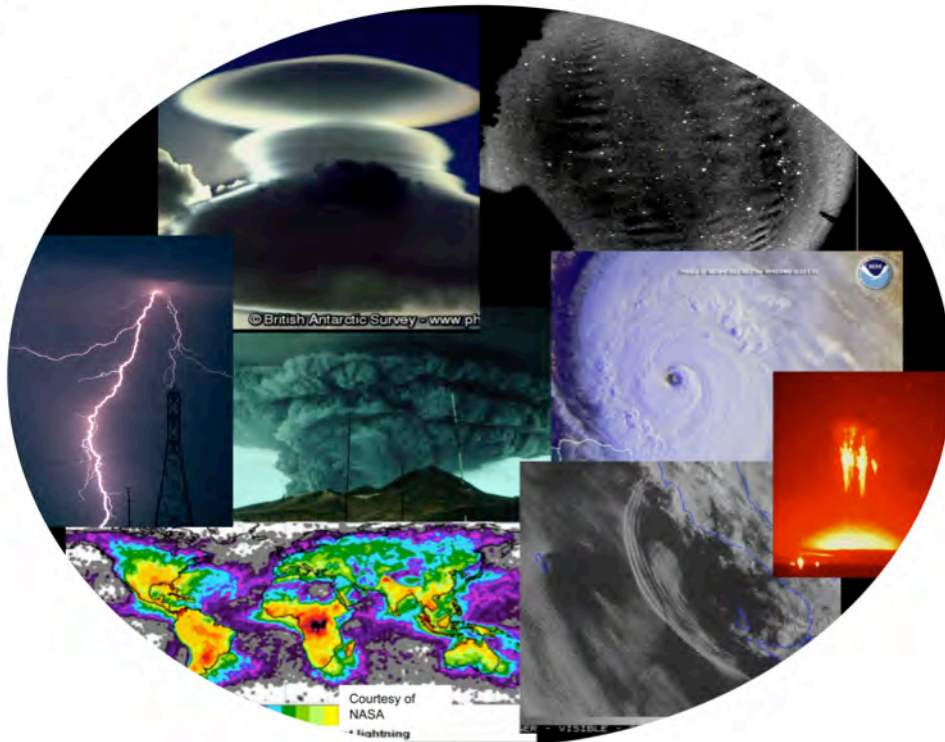


Figure 1. Collage of forcing sources in the troposphere influencing the upper atmospheric regions.

eruptions, atmospheric gravity waves, sprites, lightning, and tropical storms that produce upper atmospheric effects [Laštovička, 1998]. Additionally, it has been demonstrated that power lines and other man-made radiation sources can modify the upper atmosphere [Farrell et al., 1998].

One potentially strong coupling source between the troposphere and ionosphere/thermosphere is tropical storms, hurricanes, and typhoons. These large storms can potentially couple both electrically and dynamically. Tropical cyclones are prodigious generators of gravity waves and are suspected of affecting the ionosphere through upward propagation of this wave momentum and modification of plasma velocities. Observations have shown that the passage of tropical storms can lead to perturbations in the plasma drift [Bishop et al., 2006] as shown in Figure 2. One of the specific challenges related to tropical storms is determining the electric fields produced above the storm.

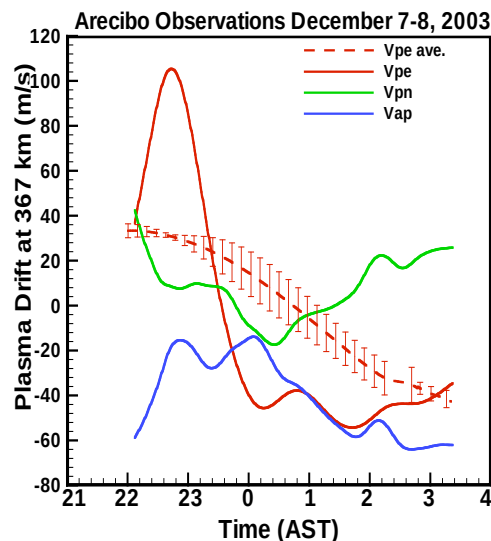


Figure 2. Anomalous plasma drifts observations coincident with passage of Tropical Storm Odette.

Another example of coupling is the signature of gravity waves in ionospheric velocity and density measurements. The DE-2 satellite observed suspected gravity waves in neutral and ion measurements at 260 km at night consistent with upward propagation from below [Earle et al., 2008].

In another study, Lieberman *et al.* [2007] showed significant reduction of non-migrating diurnal tides, and a concurrent enhancement in the migrating diurnal mode during the strong 1998 ENSO event.

Although not as strong or consistent, anthropogenic effects may also modify the lower ionosphere. For example, urban areas could be an intense source of ULF emissions that are strongly attenuated and/or absorbed by the lower ionosphere locally, thereby creating a highly localized region of heating [Farerell et al., 1998].

The primary difficulty to studying coupling between the lower and upper regions, is our ability to perform a continuity of observations from the low to high altitudes. From the examples presented, it can be noted that the evidence for coupling between the regions involves observing

ionospheric/thermospheric perturbations and noting their correlation to low altitude phenomena while eliminating other possible forcing sources. To begin to tackle this complex system problem coordinated and detailed investigations need to be undertaken that links low altitude observations through the middle atmosphere to the observations in the thermosphere and ionosphere. Specific objectives for improving our understanding of the coupling include:

- Determine the extent and significance that tropical storms in the troposphere play in producing localized electric fields and neutral gas motions in the ionosphere.
 - Measurements are needed of the electric fields at altitudes directly above large tropical storms (currently missing).
- Understand how large-scale neutral winds, gravity waves, and ion drifts are coupled in the low latitude upper atmosphere and how they respond to solar variations and to variable influences from the lower and middle atmosphere.
 - Perform low altitude observations of neutral winds and gravity waves with simultaneous observations of the mesosphere and lower thermosphere wind fields.
- Understand the surface-to-space energy pathway for ULF emissions.
- Discover the cause of planetary-scale plasma upwellings at low latitudes and investigate the effect of unstable flux tubes on inner plasmasphere dynamics.
- Determine the effect on mesosphere and lower thermosphere chemistry due to volcanic discharge, ionization due to sprites and jets, mixing due to gravity wave breaking and turbulence.
- Explain the chaotic and explosive transfer of energy between irregularity scale lengths ranging from 1000 km to 1 cm that occur in the low latitude ionosphere at night.

Recommendations:

Ground-Based Observations:

- Position permanent TIDBIT radars along with ground-based GPS receivers, CCD cameras (as appropriate) and ionosondes at locations likely to experience significant weather disturbances or gravity wave generation such as: East of the Rocky Mountains, among the Caribbean Islands, west of the Andes, and among the Indonesian Islands.
- Utilize low latitude ISRs to study affects from nearby storms

Space-Based Investigations:

- Dedicate a tropical satellite mission dedicated to exploring the tropospheric/ionospheric/thermospheric connection (see white paper on Tropical Coupler).
- Extend C/NOFS satellite operations to allow continued investigation of low latitude ionospheric irregularities that can then be combined with low altitude observations.
- Utilize the ISS for remote sensing experiments focused on the mesosphere and lower thermosphere to observe gravity waves and composition. The ISS is ideally located to observe the low latitudes that are most likely to see large tropical storms.
- Provide international collaboration to the ASIM (Atmosphere-Space Interactions Monitor) European Mission to be launched in 2013. ASIM is an instrument suite focused on the study of severe thunderstorms and their effects on the atmosphere and ionosphere both electrically and chemically.
- Utilize high altitude aircraft and balloons to determine electric fields above storms, and ULF emissions from urban areas.

Coordinated Campaigns:

- Utilize ionosonde, ISR, and Global Lightning Network with appropriate space-based observations.

Modeling Efforts:

- Continue efforts to extend models ionospheric/thermospheric models such as TIEGCM to include lower altitude forcing.
- Extend low altitude tropospheric models into the upper atmosphere.
- Include low altitude forcing in other assimilative models such as SAMI2, GAIM, IDA4D by using low-altitude model output as lower boundary conditions.

References:

Bishop, R. L., N Aponte, G. D. Earle, M. Sulzer, M. F. Larsen, G. S. Peng, Arecibo observations of ionospheric perturbations associated with the passage of Tropical Storm Odette, *J. Geophys. Res.*, 111, doi:10.1029/2006JA011668, 2006.

Earle, G. D., A. M. Musumba, S. L. Vadas, Satellite-based measurements of gravity wave-induced midlatitudes plasma density perturbations, *J. Geophys. Res.*, 113, doi10.1029/2007JA012766, 2008.

Farell, W. M, M. D. Desch, J. G. Houser, Modification of the upper atmosphere over power lines: A geological effect, *J. Geophys. Res.*, *103*, 11,573-11,578, 1998.

Laštovička, J., Lower ionosphere response to external forcing: A brief review, *Adv. Space Res.*, *43*, doi:10.1016/j.asr.2008.10.00, 2009.

Lieberman, R. S., D. M. Riggin, D. A. Ortland, S. W. Nesbitt, R. A. Vincent, Variability of mesospheric diurnal tides and tropospheric diurnal heating during 1997-1998, *J. Geophys. Res.*, *112*, doi:10.1029/2007JD008578, 2007

The International Space Station: Platform for Future Upper Atmospheric Investigations

**Author: Dr. Rebecca Bishop, The Aerospace Corporation,
Dr. James Roeder, The Aerospace Corporation,**

With the addition in 2009 of the Japanese Experiment Module (JEM) Exposed Facility on the Kibo module, the International Space Station can now be considered a viable experiment platform for upper atmospheric investigations. The ISS orbit altitude averages near 350 km with an inclination of 51.6° and a period of 91 minutes. There are many advantages to placing experiments on the ISS:

- Provides increased access to LEO orbit and whose utilization is being strongly encouraged by NASA.
- Provides flexibility for experimenters by allowing on-orbit, near real-time commanding.

The ISS Kibo module consists of a pressurized module and an exposed facility. The pressurized module performs biological and metallurgical experiments. Of interest to our community is the exposed facility, which has ten separate experiment ports and has become a host for several missions observing the space and earth environment:

- 1) HREP (Hyperspectral Imager for the Coastal Ocean (HICO) & Remote Atmospheric & Ionospheric Detection System (RAIDS): RAIDS observes lower thermosphere and ionospheric emission to obtain neutral temperature profiles, composition information, and ionospheric density. HICO provides high-resolution images of changes of the Earth's coastal areas.
- 2) MAXI: An X-ray astronomy mission.
- 3) SMILES: Stratospheric mission to observe sub-millimeter wave emission lines of trace gases.
- 4) SEDA-AP (Space Environment Data Acquisition equipment-Attached Payload): Particle experiment to measure neutrons, plasma, heavy ions, and high-energy light particles in the local environment.

In addition to Kibo, experiments may be placed along the station's truss assemblies.

Potential Ionospheric and Atmospheric Investigations:

- Ionospheric density variations, depletions, and small-scale turbulent structures.
 - Scintillation: Ionospheric bubbles at equator caused by natural instabilities scatter radio waves causing communication and navigation link outages
 - Geomagnetic storm variations in ionosphere cause increased error in navigation systems such as GPS
- Neutral density and wind variations
 - Drag Rate Variation: may result in unpredicted orbit changes for satellites, poor determination of LEO reentry, and uncertainty in satellite and debris trajectories leading to collisions

- Improve ionospheric/thermospheric models
 - Empirical models enhanced by assimilating new comprehensive measurements in a region (lower thermosphere) with relatively sparse observational data
 - Validate physics-based model output in order to develop prediction and forecasting capabilities
- Global climate change
 - Sourcing and tracking of greenhouse gas emissions for climate science. These emissions influence stratospheric/mesospheric and lower thermospheric chemistry and composition
 - Thermal IR (TIR) enables measurement throughout diurnal cycle. Thermal imaging provides information on cloud temperature and thickness that can be used to support tropospheric/upper atmospheric coupling investigations.
 - Habitat and ecosystem change. Cloud cover, water vapor content, weather patterns can all be observed using sensors that can be placed on the ISS. Observations on tropospheric changes are needed to aid in investigations of low altitude forcing on the upper atmospheric regions.

Examples of Space Sensors Suitable to the ISS Platform

- **Rotating Electric Field Ion Mass Spectrometer (REFIMS)** measures local composition of ions – previously flown on sounding rockets
- **Ionization Gauge** gives accurate local pressure and density – flown on Streak mission and sounding rockets
- **GPS Occultation Receiver (GPSRO)** remotely senses ionosphere using GPS, Galileo, Glonass, COMPASS signals over occultation paths near Earth's limb – GPS occultation receiver on AFRL C/NOFS mission and others
- **UV Limb-Scanning Instrument** provides altitude profiles of ions and neutrals (ex: RAIDS now flying on ISS JEP)
- **UV Down-Looking Imaging Spectrometer** provides regional/global maps of ionospheric and atmospheric parameters – GUVI instrument flown on NASA TIMED satellite
- **IR Down-Looking Camera** measures airglow to sense variations and waves in upper atmosphere – similar to existing ground-based imagers
- **Infrared Hyperspectral Imaging Sensors** capture high-spatial resolution infrared image with detailed spectral information. They can improve identification of chemical composition on the ground and in the atmosphere.
- **Plasma Impedance Probe/Langmuir Probe** will provide measurements of the plasma density in the near environment. Coupled with GPS occultation receivers, it allows the GPSRO to provide absolute TEC.

Potential Operations Impact

The science investigations performed from the unique platform of the ISS will not only benefit the science community, but will also be of interest to several government organizations.

- Lower thermosphere and ionospheric characterization and routine monitoring would provide additional data for improving space weather models created and maintained by NASA, NOAA, and National Science Foundation (NSF).
- The ability to investigate and improve our understanding of thermospheric density and temperature is required to improve atmospheric drag and radio propagation models used for operational purposes by commercial and military organizations.
- The use of a GPS occultation receiver to observe scintillation from the ISS and assist with our understanding of density structures of variations can be used by the FAA to help improve the size and reliability of navigation errors for aircraft and suborbital flights
- Continuous observations from highly sensitive imaging sensors on the ISS that monitor tropospheric and land-use changes are of interest to the Environmental Protection Agency (EPA), US Geological Service (USGS), and the National Weather Service.

MMAP: A Magnetic-field mapping mission concept

Dr. Jacob Bortnik¹

Dr. Yukitoshi Nishimura^{1,2}

[1] Dept. Atmospheric & Oceanic Sciences, University of California at Los Angeles.

[2] Solar-Terrestrial Environment Laboratory, Nagoya University, Nagoya, Japan.

Summary

This white paper describes a novel mission concept that is designed to provide near real-time mapping of the geomagnetic field. Our idea is based on the recent discovery that a sequence of whistler-mode chorus elements measured *in situ* by a spacecraft, can be uniquely linked to a single pulsating auroral (PA) patch, observed by a ground-based all-sky imager [Nishimura *et al.*, 2010]. We propose a mission in which several spacecraft are placed in, or near, Geosynchronous Earth Orbit (GEO) suitably instrumented to measure whistler-mode chorus waves at high resolution (<1 sec). Near the average magnetic footprint of each spacecraft, we propose installing an imaging instrument that will be able to detect the pulsating auroral patch that is linked to the spacecraft, and hence uniquely tie the location in the ionosphere and magnetosphere together. The design of this mission is modular and very scalable: nominally, the imaging instrument could be an all-sky white light imager (as was the case in our example, discussed further in Sec. 3), but preferably it should include multiple wavelengths to estimate the spectral content of the precipitation. To avoid contamination by scattered light on the dayside, we suggest deploying an array of imaging riometers, placed in Antarctica in order to complement the existing array of THEMIS all-sky imagers in the northern hemisphere, and also to avoid less accessible regions in the northern hemisphere for instrument placement. The GEO spacecraft fleet could be dedicated entirely to the MMAP mission, or alternatively we could include a plasma wave instrument suite onto existing platforms that are scheduled for launch, such as the LANL or GOES satellites. The number of spacecraft-imager pairs is also scalable: we nominally recommend 6 spacecraft, distributed as uniformly as possible in Magnetic Local Time (MLT) for a spatial sampling of ~4 hours of MLT, but the mission could be as small as a single spacecraft-imager pair or it could be many more than 6. Obviously the larger the number of spacecraft-imager pairs, the better the MLT resolution we will attain in magnetic field mapping.

Relevance

The ability to observationally map the Earth's magnetic field (albeit at a single L-shell) is crucially important for many areas in space physics, including radiation belt and ring current modeling (where GEO acts as an outer boundary), plasmasheet and/or substorm studies (GEO is roughly an inner boundary), and related fields. Understanding the Earth's varying magnetic field has been mentioned directly or indirectly in nearly all guiding documents, including the 2002 NRC decadal survey (challenges 3, 4, and 5), NASA strategic plan for 2006-2016, NASA Science plan for SMD, 2007-2016, NASA heliophysics roadmap for science and technology 2009-2030, and others. This proposal includes cross-cutting science themes that are relevant to Atmosphere-Ionosphere-Magnetosphere interactions and solar wind-magnetosphere interactions, as well as the broader area of space weather and its effective prediction.

1. Introduction

The configuration of the Earth's magnetosphere is determined by the internal field of the Earth and an external current system that forms through interaction of the solar wind impinging on the Earth's main field. The complicated chain of physical processes that sets up this current system involves a coupling of the magnetosphere to the ionosphere (MI coupling), and is fundamental in many areas of magnetospheric physics. For example, knowing the magnetic field structure is essential for radiation belt and ring current particles that follow magnetic drift shells. These drift shells can change significantly and are often difficult to predict during disturbed conditions, making a comparison of fluxes in different regions unreliable and hard to compare with simulations. In spite of a large amount of observations and modeling studies of the magnetospheric magnetic field, accurate mapping of the Earth's magnetic field lines has been a problem that is as difficult to solve as it is important. This difficulty arises from the highly time-varying nature of the solar wind and magnetospheric current system, and makes it difficult to construct an accurate field model with a high precision of ~ 100 km in the horizontal direction at the ionospheric altitude.

In this white paper, we propose a mission based on a recently discovered method of linking a distant point in space to a point in the ionosphere. This method relies on the scattering of energetic electrons by whistler-mode chorus waves that exhibits a unique temporal pattern that can be detected and accurately located in the ionosphere using ground-based imagers. Although chorus is not always present in space (which might at first appear detrimental to our proposed mission), it is nevertheless controlled by geomagnetic activity, such that it is 'switched on' and intensifies precisely when geomagnetic activity intensifies, i.e., the magnetic field model becomes the least reliable, and when we need to know it the most! Thus, the chorus PA link is well established in moderate and active times, but additionally, it has been shown that chorus at moderate levels (10 pT to 30 pT) occurs with a surprisingly large probability (30%-40%) even in quiet times [Li et al., 2009]. In the remainder of this white paper, we illustrate the technique of linking a point in space to a point in the ionosphere, and sketch out the mission concept in Sec. 3.

2. Scientific background

The conceptual model that relates chorus emissions and PA is illustrated in Figure 1. This figure shows a group of electrons (Figure 1, blue arrows) that were initially trapped by the Earth's magnetic field encountering a chorus element propagating away from the equator (red arrows). The electrons are scattered in pitch-angle by the chorus, causing them to precipitate into the upper atmosphere, and to produce photo-emissions, that are subsequently observed by the ASI.

Figure 2 shows results for a specific event recorded on February 15, 2009. Images from the Narsarsuaq, Greenland ASI (Figure 2A) at the four selected times, display a series of PA patches with a scale size of ~ 100 km, at $\sim 64^\circ$ - 67° magnetic latitude (MLAT) embedded in a weak background of diffuse aurora. The images in the second and fourth panels in Figure 2A were obtained simultaneously with intense lower-band chorus (Figure 2B), observed by the THEMIS-A spacecraft in the particle burst (FFP) mode, while it was located slightly to the south of the magnetic equator (ideal for detecting plasma waves). The

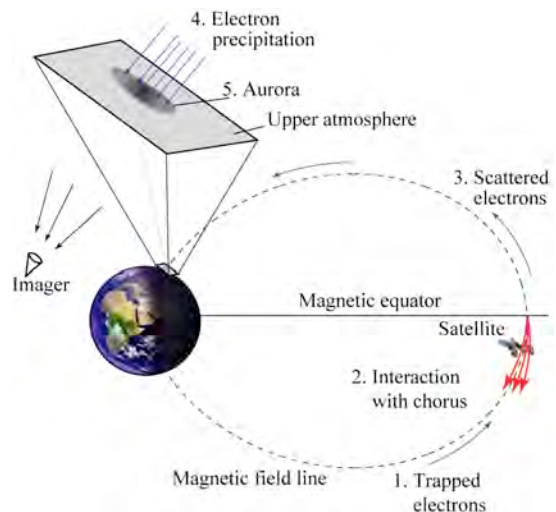


Figure 1. Schematic showing the geometry of chorus wave propagation (red arrows), electron precipitation (blue arrows) and PA.

chorus occurs as a series of repetitive discrete bursts at frequencies between 0.05 and 0.3 f_c , characteristic of lower-band emissions. Intense lower band chorus was present during this time period, while both upper band chorus (0.5-0.8 f_c) and ECH (>1.0 f_c) waves were absent, even though the spacecraft was located close to the equator.

A bright auroral patch (red arrow in Figure 2A) was identified to the west of the T96 footprint of the spacecraft, which pulsed in phase with the chorus intensity modulation, whilst adjacent auroral patches pulsed out of phase. This is the first clue of a relation between the chorus emission measured at the spacecraft and the pulsating auroral patch. The correlation between chorus intensity modulation and the PA was then investigated quantitatively by using the entire period of the wave observation shown in Figure 2B. Cross-correlation coefficients between the lower-band chorus intensity and auroral luminosity in each imager pixel were calculated to find the

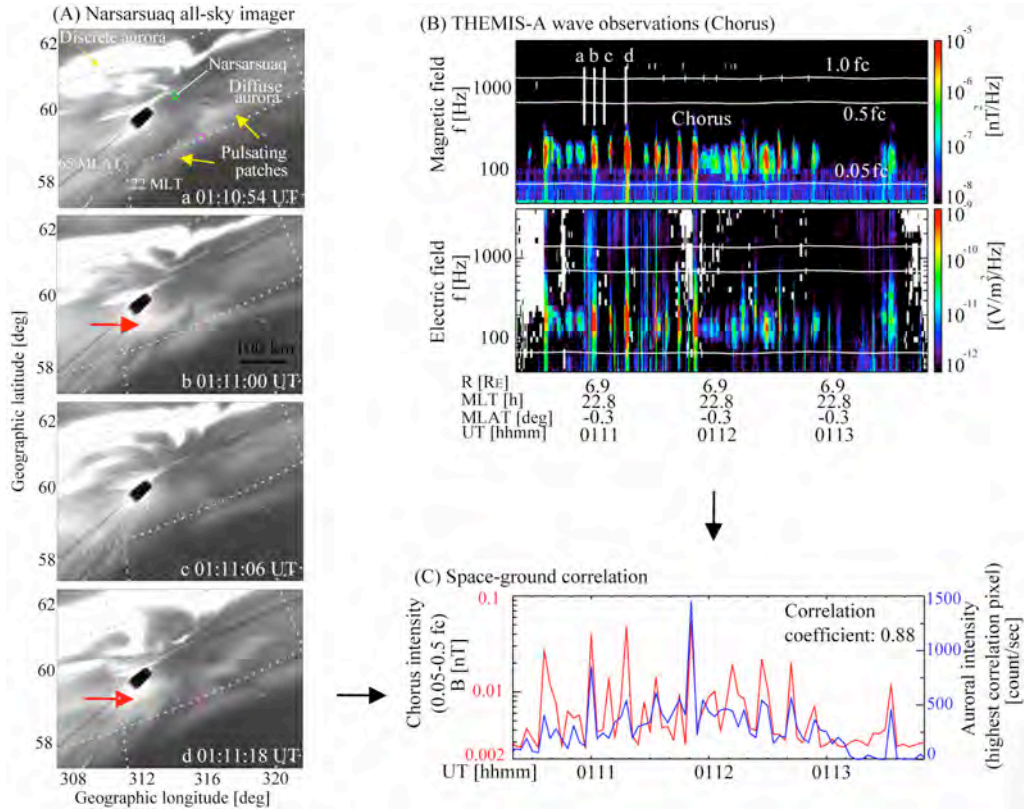


Figure 2. Coordinated observation of PA by the Narsarsuaq ASI and THEMIS-A spacecraft during 01:10:20-01:13:50 UT on 15 Feb 2009. (A) Snapshots of imager data projected onto the geographic coordinates at 110 km altitude. The pulsating patch which strongly correlates with chorus is indicated by the red arrows. ASI snapshot times are also marked in Panel (B) by white vertical lines. The pink square shows the magnetic footprint of the THEMIS-A spacecraft using the T96 magnetic field model. The center of the imager FOV is given by the green square in Panel a. Dashed lines give magnetic coordinates every 3° in latitude and 1 hour in local time. The black spot near the center of each image is an artificial object. (B) THEMIS-A observation of bursts of lower-band chorus shown in electromagnetic field spectra. The white lines indicate 0.05, 0.5 and 1.0 f_c using the measured magnetic field. (C) Correlation of lower band chorus integrated magnetic field intensity over 0.05-0.5 f_c (red) and auroral intensity (blue) at the highest cross-correlation pixel.

highest correlation pixel. Although the time lag between these two waveforms was considered, we found that the simultaneous correlation was much higher. This is consistent with the short travel times (<1 sec) of >10 keV precipitating equatorial electrons down the field line causing PA, and the short lifetime of the excited atmospheric atoms (~ 1 sec). Both of these timescales fall within the time resolution of the imager (3 sec). Figure 2C shows the integrated chorus intensity (red) and the auroral intensity at the pixel with the highest cross-correlation (blue). The correlation coefficient is remarkably high (0.88), and the auroral pulsations have an almost one-to-one correspondence with each burst of chorus. The high correlation supports our inference that intensity-modulated lower band chorus plays a major role in driving the PA.

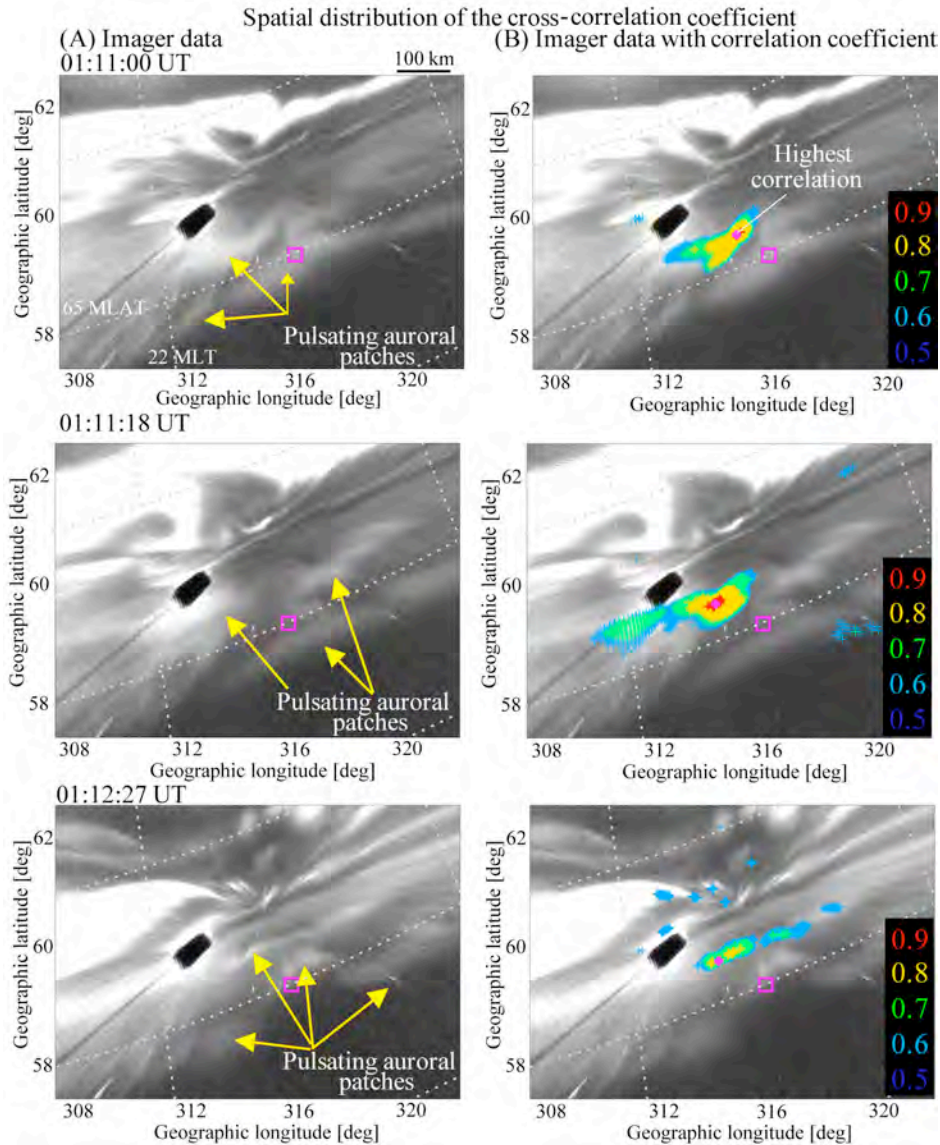


Figure 3. Temporal evolution of the spatial distribution of cross-correlation coefficients during the period of chorus observation shown in Figure 2. (A) Imager data showing PA and (B) cross-correlation coefficient superimposed onto the imager data during a 1 min time interval including each snapshot time in Panel (A). Pixels with correlation coefficient below 0.5 are not color-coded.

The spatial distribution of the cross-correlation coefficient calculated for every imager pixel is overlaid on the auroral imager data in Figure 3B. The original images are shown in Figure 3A and display pulsating auroral patches as marked by yellow arrows. The first image of Figure 3B shows a well-grouped patch of high correlation. The correlation coefficient outside this region diminishes rapidly with distance, despite several pulsating auroral patches located within the imager field of view. The high-correlation region demarcates the auroral patch marked in the second panel (b) of Figure 2A, highlighting that this is the only auroral patch that is correlated with the chorus intensity variation. The highest cross-correlation (0.91) is located 0.83° MLAT north and 0.07 MLT (80 km) west of the T96 prediction of the magnetic footprint.

The “high valued” (>0.5) cross-correlation coefficients calculated for the images obtained at later times (second and third rows) are spatially grouped on a pulsating patch, similar to the top row, underscoring the fact that only a single pulsating patch is stably correlated with the chorus intensity. While the patch shape slowly evolves with time, the location of highest correlation stays essentially fixed, which strongly suggests that the actual spacecraft footprint has been uniquely identified using the cross-correlation between the chorus and PA.

An important implication of the above analysis is the ability to uniquely link a point in space, to a patch in the ionosphere. Magnetic field line mapping has been a problematic issue in magnetosphere-ionosphere coupling studies, but the PA provides a unique opportunity to identify the footprint of the magnetic field line threading the spacecraft to the precision of an auroral patch (<100 km) and possibly even down to a few pixels. These results consequently provide unprecedented opportunities for magnetic field modeling, using the capabilities of wide-coverage auroral observations in conjunction with wave observations in space.

3. Mission concept

The key idea behind the MMAP mission is the establishment of Spacecraft-Imager (S-I) pairs, where the spacecraft is placed in, or near, a Geostationary Earth Orbit (GEO) and a ground-based imaging instrument is placed as close as possible to the average location of the magnetic footprint of the spacecraft, such that the S-I pair is always magnetically linked (Figure 4). In the course of a magnetic disturbance, the magnetic field configuration in the magnetosphere will change in response to the enhancement of a variety of MI current systems, and the point in the ionosphere that links the GEO spacecraft will change location, which will be observed using the technique described in Sec. 3. Through this conjugate observation, we will be able to track the time-varying ionospheric footprint of the magnetic field line threading the spacecraft. With multiple S-I pairs in different longitudes, we may determine MLT dependence of the magnetic field configuration.

3.1. Spacecraft requirements

The MMAP spacecraft will be placed in, or near GEO, such that their average magnetic footprint is always within the typical field of view of the ground-based imager. By construction an orbit that is slightly eccentric about the mean location GEO, it is possible to map over a broader range of L-shells and thus extend the range of the S-I pair. For a fleet of spacecraft, it is possible to design sub-or super-rotating orbits that do not remain fixed over a specific imager, but such that an imager always has at least one spacecraft within its field of view. Each of the spacecraft can be either dedicated missions, or simply consist of a plasma wave suite that is added onto future GEO spacecraft that are scheduled for launch, such the LANL or GOES satellites. The minimum requirement is a multi-component wave instrument (3 components of the wave B-field and at least 1 component of E-field) that can observe wave frequencies in the range of $\sim 0.05f_c$ - $3f_c$ (roughly 0.1-10 kHz at GEO).

To reduce the data stream, on board processing can be employed to reduce the wave data to dynamic spectrograms, with a time resolution sufficient to resolve individual chorus elements (~ 0.1 sec or less). While the magnetic field configuration that we can highlight is limited to the field lines threading the GEO altitude, it has an advantage of having a continuous availability of conjunctions with ground-based instruments described in Section 3.2. Other spacecraft moving across L-shells, such as THEMIS and RBSP, will occasionally provide data for other L-shells.

3.2. Imager requirements

The ground-based imager needs to satisfy several requirements: the total field-of-view should be large enough to capture the magnetic footprint of the GEO spacecraft, even during highly disturbed periods when the footprint can deviate significantly from its average locations (roughly, $\text{FOV} > 1000 \times 1000$ km). The resolution of the imager, however, should be small enough to resolve individual PA patches, or smaller ($\sim$ few km). Nominally, the imagers can be white light, all-sky imagers as was used in the example above (Section 2), but a significant improvement can be

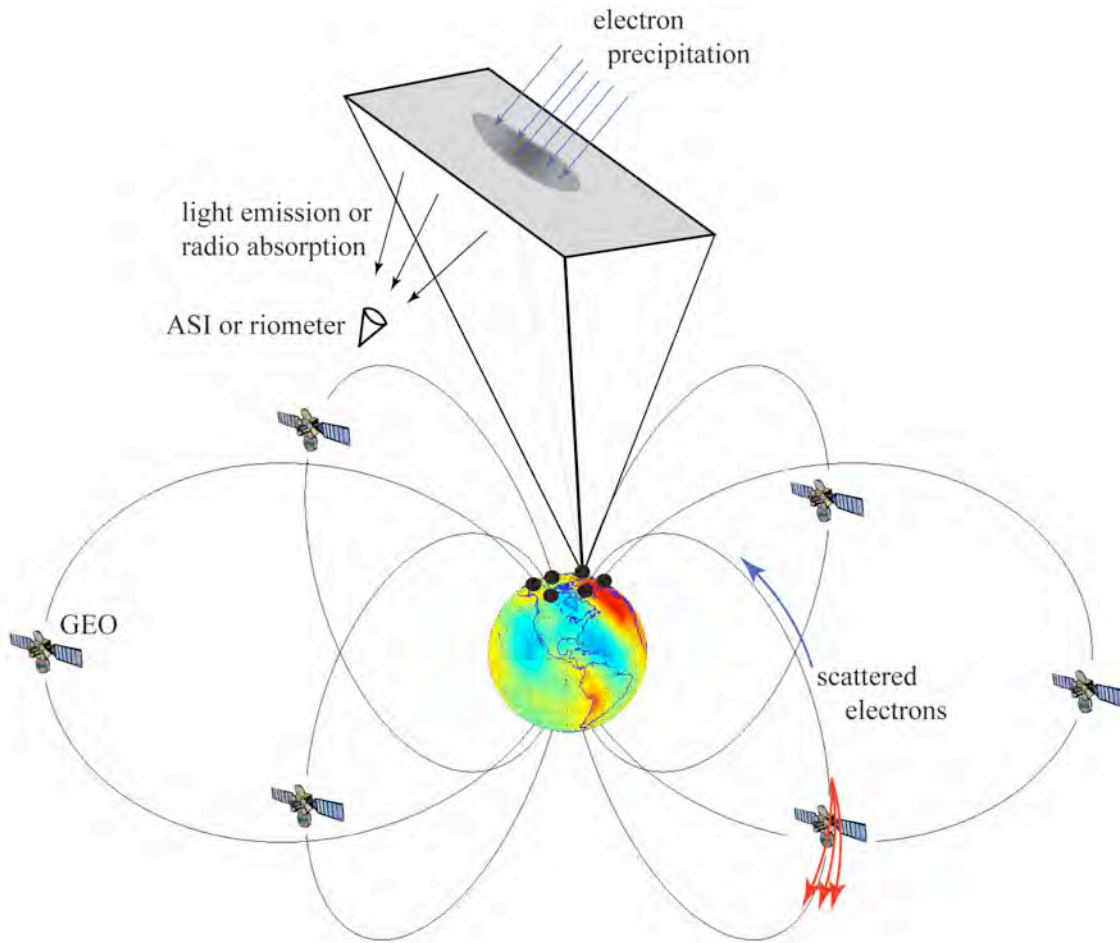


Figure 4. Schematic illustration of the MMAP mission concept: 6 GEO satellites are shown, whose magnetic footprints each thread an imaging instrument. The imaging instrument can be either an all-sky white-light (or multi-wavelength) imager, which detects light emissions, or an imaging riometer, which detects increased damping, both due to precipitating electrons.

attained in using multi-wavelength imagers, which can give an indication of the energy spectrum of the precipitated particles. Unfortunately, the imagers suffer from contamination by scattered light and are thus largely unusable on the dayside. As a remedy, we propose installing imaging riometers which are effective at all MLT, and are sensitive to the key range of energies in the PA (roughly >25 keV). We propose installing this set of riometers in Antarctica, to (i) complement the existing set of THEMIS ASI's in the northern hemisphere, (ii) to overcome the difficulties of installing scientific instruments in various regions outside of the US in the northern hemisphere, and (iii) to provide true conjugate field-mapping in the N and S hemispheres simultaneously, when possible. Although magnetic conjugate points where both northern and southern hemispheres have ground are somewhat limited, the footprint map of GEO in Figure 5 shows that Iceland to eastern Canada in the northern hemisphere have magnetic conjugacy at ~ 230 - 040 deg GLON of Antarctica. These are proposed locations to install the ground imagers.

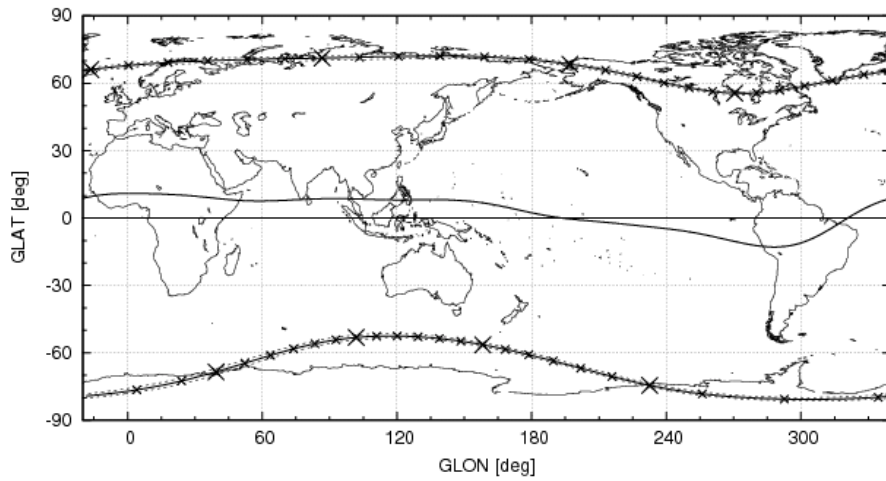


Figure 5. Geographic locations of the ionospheric footprint of the field lines threading GEO altitude (solid line). The dashed lines are for 6.1 and 7.1 RE. The small and large cross symbols show 1 and 6 MLT intervals. The Tsyganenko 01 [Tsyganenko, 2002a, b] magnetic field model under a quiet condition ($IMF\ B_y=B_z=0$, $P_{dyn}=1$ nPa and $Dst=0$) at 6 UT (UT is in the nightside) were used.

4. References

- Nishimura, Y. et al. (2010), Identifying the Driver of Pulsating Aurora, *Science*, 330(6000), 81-84, doi:10.1126/science.1193186.
- Tsyganenko, N. A. (2002a), A model of the near magnetosphere with a dawn-dusk asymmetry. 1. Mathematical structure, *J. Geophys. Res.*, 107, 1179.
- Tsyganenko, N. A. (2002b), A model of the near magnetosphere with a dawn-dusk asymmetry. 2. Parametrization and fitting to observations, *J. Geophys. Res.*, 107, 1176.

Evolved Tiny Ionospheric Photometer (ETIP): A Sensor for Ionospheric Specification

Scott Budzien¹, Damien Chua¹, Clayton Coker¹, Patrick Dandenault¹, Kenneth Dymond¹, Andrew Nicholas¹, Richard Doe², Geoff Crowley³

¹*Space Science Division, Naval Research Laboratory, Washington, DC*

²*SRI International, Menlo Park, CA*

³*Atmospheric and Space Technology Research Associates, San Antonio, TX*

THEME: Atmosphere-Ionosphere-Magnetosphere Interactions

EXECUTIVE SUMMARY

We describe a nadir-viewing far-ultraviolet (FUV) photometer concept as a complementary sensor to GPS radio occultation (GPSRO) techniques to meet space weather monitoring objectives for total electron content (TEC). The Evolved Tiny Ionospheric Photometer (ETIP) addresses the need for compact spaceflight sensors that contribute to ionospheric density profile specification and data assimilation. Photometric measurements of oxygen nightglow at 135.6 nm are operationally equivalent to TEC measurements for the purpose of ionospheric specification and forecasting. FUV nadir photometry is a powerful technique to accurately measure ionospheric structures in the presence of large gradients—a problematic measurement scenario for GPSRO methods.

The ETIP sensor design enhances reliability, performance, lifetime, and manufacturability of the flight-proven Tiny Ionospheric Photometer (TIP) sensor, in operation for over four years on the COSMIC mission. We incorporate innovative design elements from a new low size, weight, and power ionospheric photometer, developed by SRI International (SRI) for NRL's Space Weather CubeSat mission, to achieve an evolved design which provides improved duty cycle, signal-to-noise, sensitivity, and redundancy. The ETIP measures the naturally-occurring 135.6 nm radiant emission produced by electron-ion recombination processes to characterize horizontal ionospheric gradients that can compromise GPS occultation retrievals, to identify the presence of ionospheric depletions, and to improve the fidelity of global ionospheric specification. ETIP collects ionospheric measurements on the nightside (where horizontal gradients are strongest and spread-F depletions occur), and accumulates noise characterization and diagnostic data on the dayside.

The innovative ETIP sensor design consists of identical paired nadir-viewing far-ultraviolet photometers, which together occupy less volume and footprint than the heritage TIP photometer design and mitigates a long-wavelength contamination issue affecting previous TIP sensors. ETIP operates with a fast off-axis parabolic telescope feeding a heated SrF₂ optical blue-cut filter to eliminate undesirable short wavelength emission; a compact CsI photocathode photomultiplier tube, selected for insensitivity to unwanted long wavelength emissions, collects the 135.6 nm emission with extremely high sensitivity (500 counts s⁻¹ Rayleigh⁻¹). By using twin photometers, one unit can monitor ionospheric emission with 100% duty cycle while the other monitors dark counts and/or any residual long-wavelength contaminating (red-leak) emission. When red-leak is negligible, both photometers can collect 135.6 nm emissions, doubling the effective sensitivity to an unprecedented 1000 counts s⁻¹ Rayleigh⁻¹. This sensor design provides uniquely high sensitivity, operational redundancy, and graceful degradation at sensor end-of-life.

SENSOR DESCRIPTION

Though the need for space environment monitoring is growing and operational space weather models are becoming increasingly sophisticated and data-hungry, DoD and other agencies face the prospect of a shortfall in our space-based assets capable of monitoring the ionosphere-thermosphere system as the DMSP, C/NOFS, and COSMIC missions wind down. In this resource-limited environment, providing data to current and future (full-physics) versions of the global assimilative ionospheric models (GAIM)

[e.g. *Schunk et al.*, 2004; *Komjathy et al.* 2010] requires not simply generating a greater volume of homogeneous measurements (e.g. TEC from ground-based GPS and GPSRO), but providing GAIM with heterogeneous data sets which provide complementary gauges for global ionospheric specification. UV photometers measure ionospheric structure without regard to a preferred transmitter-receiver geometry and can provide early detection of ESF bubbles at nadir well before scintillation effects manifest at spacecraft altitudes. Nadir photometry integrates vertically to characterize horizontal ionospheric morphology and local gradients, while GPSRO integrates horizontally to reveal vertical structure. GPS-based TEC and photometer-based 135.6 nm UV radiance thus represent a heterogeneous data pair, and their mutual assimilation into GAIM significantly enhances the quantitative fidelity of the convergent ionospheric solution.

UV Remote Sensing and Nadir Photometry Method

Operational Utility of 135.6 nm Radiance

Naturally occurring ultraviolet and visible emissions from the night sky have been used since the late 1960's as a means of determining the number densities of electrons and O^+ ions in the F-region ionosphere. The emissions are produced by radiative recombination of O^+ ions and electrons to produce atomic oxygen in an excited state which subsequently decays by emitting a photon. The use of the nighttime O emissions to remotely sense the F-region electron density is described by *Tinsley and Bittencourt* [1975], *Chandra et al.* [1975], *Dymond et al.*, [1996, 1997], and *Meier* [1991]. The far-ultraviolet (FUV) emissions are of the most interest for space-based ionospheric sensing, because the Earth's lower atmosphere strongly absorbs FUV radiation. Thus, there is no FUV background. The workhorse among the UV emissions is the atomic oxygen emission line at 135.6 nm. In the nighttime F-region ionosphere the optically thin 135.6 nm emission doublet arises from the recombination of atomic oxygen ions,



with a photon volume emission rate proportional to $7.3 \times 10^{-13} \text{ cm}^3 \text{ s}^{-1} (\alpha_{135.6})$ times the square of ionospheric density ($n_e * n_{O^+}$) in CGS units [*Melendez-Alvira et al.*, 1999]. If height-integrated to yield observed nightglow brightness ($4\pi I$), we see that the result is proportional to the path integral of ionospheric density squared (due to charge neutrality with $n_e = n_{O^+}$):

$$4\pi I = 10^{-6} \int_0^{\infty} \gamma \alpha_{135.6} n_e(z) n_{O^+}(z) dz \quad (2)$$

Because this 135.6 nm radiance integral is similar to the TEC integral (line of sight path integral of n_e) both share assimilative value to ionospheric models such as GAIM and IDA4D [*Bust et al.*, 2004]. The 135.6 nm radiance, with an integrand proportional to the *square* of ionospheric density, is much more sensitive to ionospheric gradients than is TEC, making it more useful in assimilation programs to accurately reproduce ionospheric morphology.

Current State of the Art: COSMIC/TIP

The six TIP sensors [*Kalmanson et al.*, 2004] on the COSMIC satellites [*Rocken et al.*, 2000] were designed in large measure to address this critical role of measuring ionospheric gradients. The COSMIC/TIP sensors have been operating on-orbit continuously since mid-May 2006. The sensors were designed to be simple to operate, low maintenance, inexpensive, and of very high sensitivity. To make measurements with high enough precision to complement the GPSRO measurements, the signal-to-noise ratio of the TIP measurements had to be as high as possible within the size and spatial resolution constraints. The photometer sensor heads of the TIPs were designed to fit into $15 \times 15 \times 7.5 \text{ cm}^3$ volumes. The spatial resolution was desired to be better than 50 km. The TIP fields-of-view were 3.8° diameter, leading to a $\sim 30 \text{ km}$ footprint at an ionospheric altitude of 350 km when viewed from 800 km.

The heart of the TIP sensor is a heated SrF_2 window, which serves as a blue-cut filter (Figure 1). By heating the window to 100°C , the cutoff wavelength shifts longward to eliminate an unwanted oxygen emission line at 130.4 nm, yet passes 135.6 nm emission nearly unattenuated. This allows TIP to maintain high sensitivity by avoiding low-throughput interference filters and inefficient diffraction gratings. The TIP sensitivities ranged from $\sim 400\text{--}700\text{ counts s}^{-1}\text{ Rayleigh}^{-1}$ [Budzien et al. 2009; Dymond et al. 2009a]; the exposure times are nominally 1.14 s. For comparison, the GUVI instrument on NASA's TIMED satellite had a sensitivity of $28\text{ counts s}^{-1}\text{ Rayleigh}^{-1}$, but over a $12^\circ \times 0.4^\circ$ field-of-view and with an exposure time of 0.03 s. When the instruments are compared at the same spatial resolution ($\sim 30\text{ km}$), the GUVI sensitivity-dwell time product is $2.7\text{ counts Rayleigh}^{-1}$, while the TIPs are $\sim 450\text{--}800\text{ counts Rayleigh}^{-1}$; the TIPs are effectively 150–270 times the sensitivity of the GUVI at the same spatial resolution. This sensitivity difference reflects how TIP is highly-optimized for nighttime sensitivity in a narrow track along the spacecraft nadir, whereas GUVI was designed for mapping broad swaths of bright atmosphere.

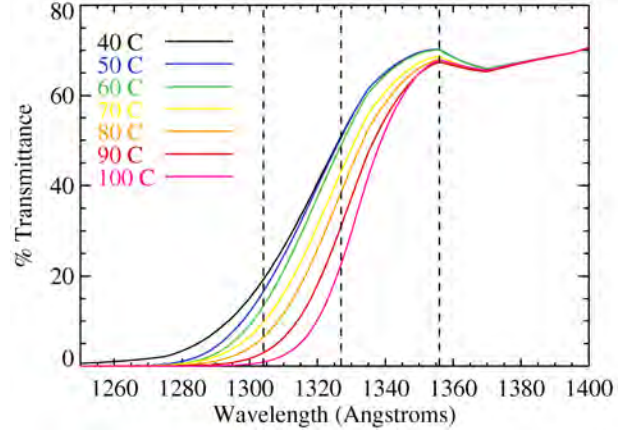


Figure 1: The transmittance of a heated SrF_2 window eliminates 130.4 nm at 100°C but passes 135.6 nm virtually unattenuated [Budzien et al., 2009]

A significant problem of photometer “red-leak” (unwanted visible and near-UV sensitivity) became evident upon initial TIP operations. The CsI photocathode of the TIP photomultiplier tube (PMT) detectors is normally “solar-blind”, exhibiting sensitivity which sharply drops by many orders of magnitude at wavelengths longward of 200 nm. A rapid flight unit development cycle (6 months!) and a constricted budget limited the time available for full detector photometric characterization. The detectors were not screened for sensitivity at wavelengths outside the FUV expected passband. Consequently, TIP is sensitive to lights from bright urban areas, scattered moonlight, and terminator glints from COSMIC beacon antennas, which collectively reduce the volume of the highest-quality ionospheric photometry data.

Therefore, to advance the state-of-the-art a primary design goal for ETIP is to characterize TIP-era red leak in the laboratory, to procure enhanced red roll-off PMTs to reduce red-sensitivity, and to monitor any residual red-leak on-orbit using a dedicated red-leak background filter.

Vertical TEC from Optical Remote Sensing

The vertical total electron content can be directly computed from ETIP measurements of the oxygen 135.6 nm radiance. The radiance integral (Equation 2) can be evaluated directly when a Chapman layer approximation [Chamberlain and Hunten, 1987] is assumed for the electron (and ion) density distribution. The Chapman layer closely approximates the O^+ ion distribution at night and upon integration radiance in Rayleighs becomes:

$$4\pi I = 0.212 (NmF2/10^6)^2 H_{km} \quad (5)$$

where $NmF2$ is the peak electron density in units of cm^{-3} and H_{km} is the scale height in kilometers. The comparable total electron content integral yields:

$$VTEC = 0.33 NmF2/10^6 H_{km} \quad (6)$$

where $VTEC$ indicates the vertically integrated electron density (vertical total electron content, VTEC) in TECU (10^{12} cm^{-2}). Because 135.6 nm radiance is proportional to the product of the vertical total electron content and peak electron density, it provides a more sensitive detection of ESF bubbles and other small-scale ionospheric structures than does TEC. Equations 5 and 6 allow VTEC to be derived from radiance:

$$VTEC = 0.717 \sqrt{4\pi I H_{km}} \quad (7)$$

where the value of H_{km} can be extracted from an assimilative model or derived from analysis of GPS occultations. Table 1 casts the ETIP sensitivity in VTEC units with governing assumptions on SNR and layer scale height.

Table 1. ETIP Measurement Sensitivity				
Measurement Output	Output Range	Accuracy	Cadence	Comment
Vertical TEC observable	0.5—250 TECU	0.5—0.25 TECU rel. sensitivity; 15% abs. accuracy	1 Hz	Horizontal Resolution 30 km @ 800km orbit 16 km @ 500km orbit
Electron Density Peak	NmF2 of 2×10^4 — 1×10^7 cm ⁻³			Scale height from GPS occultation or GAIM

Relevance to Space Weather Sensing Objectives

The ETIP concept is motivated by DoD space environment monitoring requirements and flow-down from those requirements to provide a large quantity of high quality measurements of ionospheric gradients, especially at night where the gradients are larger and more problematic. *Dymond et al.* [2009b] used the TIP instruments on the COSMIC constellation to show that high sensitivity photometric measurements could be used in a tomographic inversion scheme to complement the very high precision GPSRO measurements made by those satellites and provide a more accurate and complete picture of the electron density distribution in the night-side ionosphere. The TIP sensors significantly “raised the bar” for photometric measurements of the nighttime ionosphere. The combination of high spatial resolution and high sensitivity produced high precision ionospheric measurements even at mid-latitudes, during the early morning hours, at solar minimum conditions where the ionospheric densities are known to be extremely small.

Improving GAIM Results with ETIP

The ETIP UV radiance and TEC measurements will significantly improve the quality of electron density profile (EDP) specifications produced by ionospheric assimilation models like GAIM. The top panel of Figure 2 shows electron density profiles produced by GAIM at 2100 local time (LT) as a function of magnetic latitude and altitude when ingesting only ground-based GPS measurements. The middle panel shows measurements of the 135.6 nm radiance at this local time by two different COSMIC/TIP sensors (red traces) as compared to the 135.6 nm radiance that GAIM would produce using the EDP shown above (blue trace). The GAIM output clearly does not match the TIP observations as the relative intensities of the northern and southern ionization crests are opposite of what is actually measured. When the TIP observations are ingested into GAIM, the electron density profile produced by the model more accurately specifies the anomaly crests and the trough (Figure 3, bottom panel). This demonstrates how nadir-viewing UV photometer data such as those provided by ETIP can drive the GAIM model toward an electron density profile specification that is much more consistent with the real ionosphere.

Moreover, ETIP provides data with a constant, simple viewing geometry and uniformly high quality

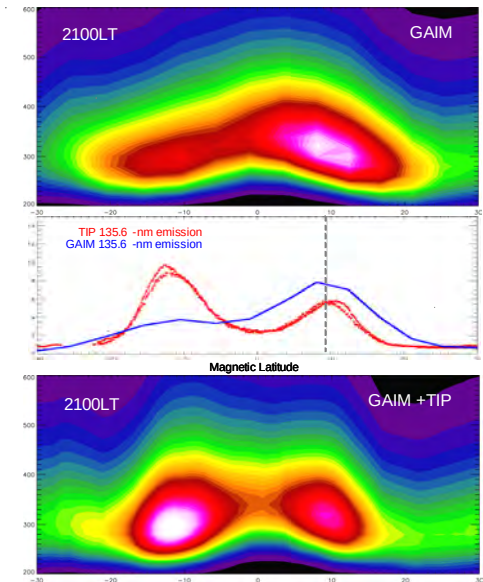


Figure 2: GAIM EDP specification without (top) and with TIP data (bottom). The middle panel show two TIP passes (red) and the initial GAIM predicted brightness (blue).

over land and ocean, which facilitates both ingestion and error correlation analysis, which are critical aspects of assimilative modeling. The dependence of the UV signal upon the square of electron density places a very strong constraint upon the electron density and TEC predicted by GAIM. Assimilation of nadir photometer data can even improve vertical accuracy of ionospheric specifications, despite the fact that ETIP itself provides no vertical information. This surprising capability arises from the aliasing of horizontal and vertical signal along the many long slant TEC lines-of-sight from other data sources feeding GAIM. By improving GAIM's horizontal ionospheric specification, ETIP data allow GAIM to more accurately render TEC, NmF2, and even hmF2 derived largely from complementary assimilated datasets. The improvements to ionospheric specification from such heterogeneous data are particularly important in the highly structured nightside ionosphere.

ETIP Design

Design Philosophy

The core ETIP design philosophy embraces two main principles: 1) reliance on TIP design heritage to the greatest extent possible (while addressing performance limitations and reliability) and 2) incorporation of innovative CubeSat Tiny Ionospheric Photometer (CTIP) [Noto *et al.*, 2009] subsystems to drive down size, weight, and power (SwAP) and enhance performance. The TIP sensors on all six COSMIC satellites have operated in orbit for four years and four weeks with no operational failures. The new design addresses red-leak issues by procuring new PMTs specially-selected for enhanced roll-off longward of 135.6 nm, and providing ETIP with a dedicated full-aperture BaF₂ filter capable of providing red-leak background monitoring.

Optical Design

The notional ETIP is envisioned as a dual modular photometer with two identical optical chains. The twin channels provide one channel dedicated to 100% duty cycle nightside monitoring of 135.6 nm ionospheric airglow to improve duty cycle over TIP, while the alternate channel provides simultaneous measurement of red-leak or dark counts (Figure 3). Each optical chain incorporates scattered light baffles, a first-surface F/1 off-axis parabolic (OAP) mirror, a heated SrF₂ blue-cut filter, solenoid actuated red-pass BaF₂ filter, solenoid shutter, and up-rated Hamamatsu PMT—all elements designed for the volume and power constrained CTIP program. The combination of a 50.8 mm focal length OAP mirror and 3.4 mm PMT photocathode illumination was selected to yield an instrumental FOV of 3.8° to match the specification of TIP. In order for these optics to comply with the CubeSat standard, significant light-weighting and power optimization was required. For example, the third generation heated SrF₂ filter subassembly, which removes strong OI emission at 130.4 nm, maintains the crystal at 100° C with 1.5 W despite -30° to +50° C thermal forcing. The photometric response of PMTs, BaF₂ and SrF₂ crystals in each of the two optical chains will be optimally matched during ETIP assembly so that each path serves a redundant function. The ETIP thus provides operational features heretofore

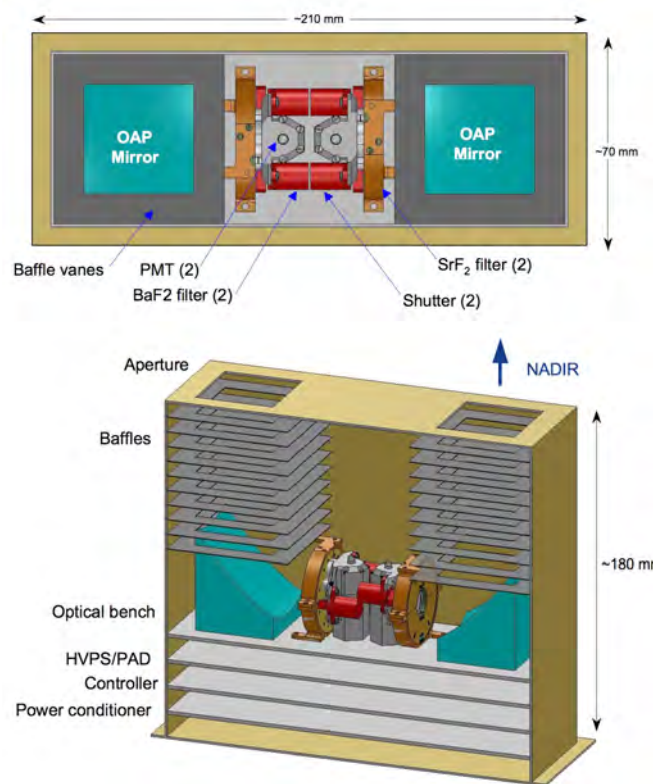


Figure 3: ETIP design - top and interior views.

unavailable in the TIP and CTIP designs. In addition to allowing uninterrupted 135.6 nm coverage during acquisition of red-leak and dark samples, the ETIP can also command both chains to acquire 135.6 nm photocounts so as to double on-band sensitivity.

Mechanical Design

Figure 3 outlines a notional ETIP layout with a vertical stack of electronic boards underneath an optical bench. The ETIP housing, supporting all internal subsystems and providing a light-tight barrier, will be designed along the lines of the TIP-era TICE electronics enclosure to allow individual access to EMI-shielded electronic boards during integration and testing. Because the ETIP optical paths are essentially clones, judicious layout of the electronic boards will allow an optional housing to be envisioned with a square mounting footprint as opposed to the rectangular footprint shown in Figure 3.

Optional Improvement: Additional Off-Nadir Pointing

The modularity and small footprint of the ETIP enables additional operational capabilities including off-nadir pointing. In addition to the baseline nadir oriented ETIP, this scenario envisions a second ETIP unit configured with a base wedge or front-end periscope so that its optical chain is oriented at a fixed off-nadir angle (ONA) in-track. Such in-track ONA measurements, in conjunction with nadir measurements, yields a measurement set with intersecting ray paths suitable for ionospheric tomographic imaging. ETIP configurations using mounting wedge versus periscope implementation allow sensor packaging options customizing vertical form factor, horizontal footprint, and photometric throughput.

Table 3. TIP and ETIP Comparison

Parameter	TIP	ETIP
Sensitivity (cnts/R-s)	500	500*
Red-leak Background Samples	None	Continuous
Max 135.6 nm Duty Cycle	93%	100%
Field of View	3.8°	3.8°
Volume	3000 cm ³	2700 cm ³
Mounting Footprint	300 cm ²	150 cm ²
Weight	2300 g	2000 g
Average Power	7.6 W	10 W
Filter Cycles	1,600,000	24,000,000

*Sensitivity doubling to 1000 possible at the expense of background monitoring and 3 W additional power

SUMMARY

ETIP is part of a new generation of space weather sensors with low size, weight, and power suitable for deployment on buses ranging from small to large satellites. To produce the initial two sensors, including development, risk reduction, operationalization, and extensive testing costs approximately \$4M. After the design program is complete, additional flight sensors could be procured for missions at a rough order of magnitude estimated cost of \$950k each, when produced singly. The ETIP design addresses the requirements for space weather sensors, and includes adequate flexibility for accommodation on a range of future flight opportunities, including microsatellite constellations. Such modularity and configurability are crucial to satisfy NRL's mission to transition innovative sensor concepts out of the laboratory and into the military, civilian, and commercial sectors.

REFERENCES

- Budzien, S., K. Dymond, C. Coker, and D. Chua, Tiny Ionospheric Photometers on FORMOSAT-3/COSMIC: On-orbit performance, *Solar Physics and Space Weather Instrumentation III, Proc. SPIE* **7438**, Silvano Fineschi, Judy A Fennely, eds., 13-1-13-11, 2009.
- Bust, G. S., T. W. Garner, and T. L. Gaussiran II, Ionospheric Data Assimilation Three Dimensional (IDA4D): A Global, Multi-Sensor, Electron Density Specification Algorithm, *J. Geophys. Res.*, **109**, A11312, doi:10.1029/2003JA010234, 2004.

- Chamberlain, J. W. and D. M. Hunten, *Theory of Planetary Atmospheres: An Introduction to Their Physics and Chemistry*, Academic Press, Inc., Orlando, FL., 1987.
- Chandra, S., E. I. Reed, R. R. Meier, C. B. Opal, and G. T. Hicks, Remote sensing of the ionospheric F layer by use of O I 6300-Å and O I 1356-Å observations, *J. Geophys. Res.*, **80**, 2327, 1975.
- Dymond, K. F., S. E. Thonnard, R. P. McCoy, and R. J. Thomas, A Technique for determining F region electron densities using optical measurements of recombination radiation, *Proceedings of the 1996 Ionospheric Effects Symposium*, 7-9 May 1996, Alexandria, VA., 346, 1996.
- Dymond, K. F., S. E. Thonnard, R. P. McCoy, and R. J. Thomas, An optical remote sensing technique for determining nighttime F region electron density, *Radio Science*, **32**, 1985, 1997.
- Dymond, K. F., S. Budzien, C. Coker, and D. Chua, On-orbit calibration of the Tiny Ionospheric Photometer on the COSMIC/FORMOSAT-3 satellites, *Solar Physics and Space Weather Instrumentation III, Proc. SPIE* **7438**, Silvano Fineschi, Judy A Fennely, eds., 14-1-14-11, 2009a.
- Dymond, K. F., S. A. Budzien, D. H. Chua, C. Coker, and J.-Y. Liu, Tomographic reconstruction of the low-latitude nighttime electron density using COSMIC/FORMOSAT-3 radio occultation and UV photometer data, *Terrestrial, Atmospheric, & Oceanic Sciences*, **20**, 215, 2009b.
- Kalmanson, P. C., S. Budzien, C. Coker, and Dymond, The Tiny Ionospheric Photometer instrument design and operation, *Proc. of SPIE*, **5660**, 259, doi: 10.1117/12.578341, 2004.
- Komjathy, A., B. Wilson, X. Pi, V. Akopian, M. Dumett, B. Iijima, O. Verkhoglyadova, and A. J. Mannucci (2010), JPL/USC GAIM: On the impact of using COSMIC and ground-based GPS measurements to estimate ionospheric parameters, *J. Geophys. Res.*, **115**, A02307, doi:10.1029/2009JA014420.
- Meier, R. R., Ultraviolet spectroscopy and remote sensing of the upper atmosphere. *Space Sci. Rev.*, **58**, 1991.
- Melendez-Alvira, D. J., R. R. Meier, J. M. Picone, P. D. Feldman, and B. M. McLaughlin, Analysis of the oxygen nightglow measured by the Hopkins Ultraviolet Telescope: Implications for ionospheric partial radiative recombination rate coefficients, *J. Geophys. Res.*, **104**, 14901, 1999.
- Noto, J., R. Doe, G. Crowley, R. Kerr, K. van Dyk, K. Leveque, Miniaturized Vacuum Ultraviolet Photometer, Presented at the Cubesat Developer's Conference, San Luis Obispo, CA, 2009.
- Rocken, C., Y. H. Kuo, W. Schreiner, D. Hunt, and S. Sokolovskiy, COSMIC system description, *Terr. Atmos. Oceanic Sci.*, **11**, 21, 2000.
- Schunk, R.W., L. Scherliess, J.J. Sojka, and D. Thompson, Global Assimilation of Ionospheric Measurements (GAIM), *Radio Sci.*, **39**, RS1S02, doi:10.1029/2002RS002794, 2004.
- Tinsley, B. A., and J. A. Bittencourt, Determination of F region height and peak electron density at night using airglow emissions from atomic oxygen, *J. Geophys. Res.*, **80**, 2333, 1975.

The Volumetric Imaging System for the Ionosphere (VISION)

S. A. Budzien¹, K. F. Dymond¹, D. Chua¹, C. Coker¹, A. C. Nicholas¹, and S. E. Thonnard²

¹Space Science Division, Naval Research Laboratory, Washington, DC

²Office of Naval Research, Arlington, VA

THEME: Atmosphere-Ionosphere-Magnetosphere Interactions

SUMMARY

The Volumetric Imaging System for the Ionosphere (VISION) utilizes limb and nadir images to reconstruct the three-dimensional distribution of electrons over a 1000 km wide by 500 km high slab beneath the satellite with $10 \text{ km} \times 10 \text{ km} \times 10 \text{ km}$ voxels. The primary goal of the VISION is to map and monitor global and mesoscale ($> 10 \text{ km}$) electron density structures, such as the Appleton anomalies and field-aligned irregularity structures.

The VISION consists of three UV limb imagers, two UV nadir imagers, a dual frequency Global Positioning System (GPS) receiver, and a coherently emitting three frequency radio beacon. The limb imagers observe the O II 83.4 nm line (daytime electron density), O I 135.6 nm line (nighttime electron density and daytime O density), and the N₂ Lyman-Birge-Hopfield (LBH) bands near 143.0 nm (daytime N₂ density). The nadir imagers observe the O I 135.6 nm line (nighttime electron density and daytime O density) and the N₂ LBH bands near 143.0 nm (daytime N₂ density). The GPS receiver monitors the total electron content between the satellite carrying VISION and the GPS constellation and detects scintillation. The three frequency radio beacon will be used with ground-based receiver chains to perform computerized radio tomography below the satellite. Measurements made using the two radio frequency instruments may also be used for on-orbit calibration and validation of the VISION UV measurements.

1. INTRODUCTION

The Space Science Division at the Naval Research Laboratory has developed a program for characterizing the Earth's thermosphere and ionosphere based upon observing the naturally occurring airglow using ultraviolet sensors on space-based platforms. These sensors are passive in nature, so that there are no frequency management issues as is the case with topside sounders and radio beacons. In addition, with modern advances in multilayer filter design these sensors can be relatively compact.

We describe a new instrument concept developed at the NRL that will perform volumetric imaging of the Earth's ionosphere and thermosphere from a low-Earth orbiting satellite platform. This experiment is called the Volumetric Imaging System for the Ionosphere or VISION. The VISION is designed to map and monitor the three-dimensional distribution of mesoscale ($> 10 \text{ km}$) structures in the thermosphere and ionosphere. VISION will observe the Earth's limb and nadir and combine these measurements using tomographic techniques to infer the electron and neutral density over a 1000 km wide by 500 km thick slab of atmosphere beneath the satellite.

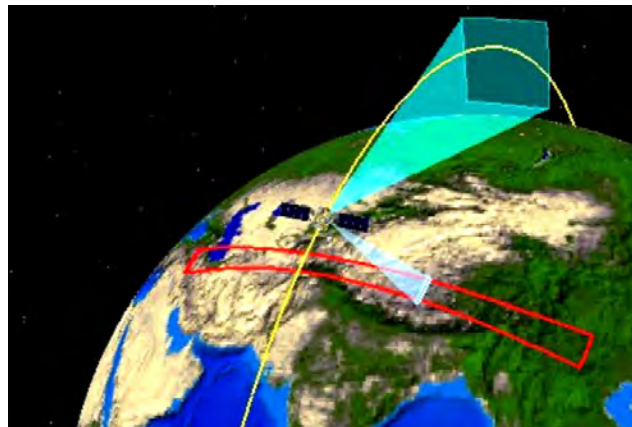


Figure 1: This shows the VISION concept of operation. The region that has been imaged by the cross-track scanner will be imaged by the limb imagers approximately 500 seconds later.

It can also observe the northern and southern auroral ovals to better understand the energy deposition in these important atmospheric regions.

2. SCIENCE OVERVIEW

The VISION will attempt to answer several key science questions to help improve our understanding of the coupled ionosphere thermosphere system by providing both high spatial resolution (~10 km) and high sensitivity (~500 ct/s/Rayleigh). These characteristics in combination with its vantage point in low-Earth orbit (~1000 km altitude) make it an ideal platform for addressing the key science questions below.

Key Science Questions

What is the global response of the ionosphere and thermosphere to solar and geomagnetic forcing? The VISION will produce high spatial resolution limb and nadir imaging data that will be used to reconstruct the three dimensional electron distribution in the day and nighttime ionosphere. Knowledge of the volumetric electron density distribution can be used to study the complex coupling between the winds, electric fields, the Earth's magnetic field, and neutral and ion chemistry that determine the state of the ionosphere. The VISION data are gathered along the orbit track, as indicated in Figure 1.

How are ionospheric composition changes related to the underlying neutral thermosphere? The neutral atmosphere is the ultimate source of the ionosphere, so measurements of the volumetric distribution of the principal species in the thermosphere is also necessary to better understand the ionosphere. The F-region ionosphere is produced primarily by photoionization of O at altitudes below 150 km. This production is balanced by recombination loss. One of the principal recombination loss processes is charge exchange of the O^+ with either N_2 or O_2 to produce N_2^+ or O_2^+ , which rapidly recombine with ambient electrons. This charge exchange process is strongly dependent on the N_2 and O_2 densities, as well as the O^+ density. The ratio of O to N_2 column densities has been observed to change during high geomagnetic activity. As this ratio decreases, the rate of photoionization of O decreases and the rate of recombination increases. Thus, the ionospheric electron density is expected to decrease because the production rate decreases and the loss rate increases. This has been observed using the *Dynamics Explorer* satellite (Strickland et al., 2001). The DE data, however, do not have very high spatial resolution. The VISION will be able to perform studies like the DE studies but with simultaneous volumetric sampling of both the neutral and electron densities instead of the two-dimensional spatial mapping of DE. This will provide better understanding of how the changes in the O to N_2 species densities are reflected in the ionospheric electron density.

What are the characteristics, behavior, and origin of ionospheric bubbles and mid-latitude depletions? The mid- and low-latitude nighttime ionosphere sometimes develops strong mid- to small-scale ionospheric irregularities. These irregularities are caused by plasma

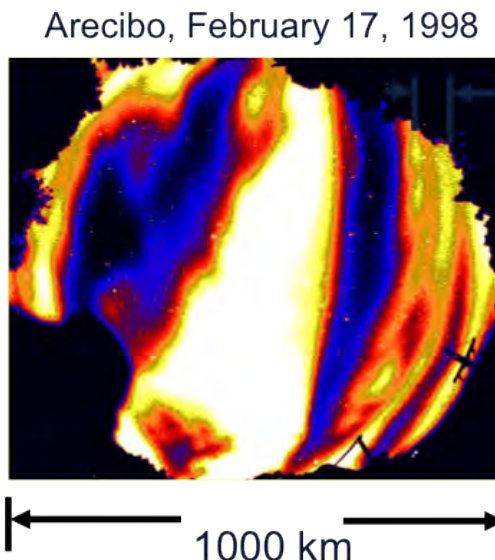


Figure 2: All sky camera image at 630 nm of a mid-latitude irregularity, acquired at the Arecibo Incoherent Scatter Radar Facility in Puerto Rico. The black irregular border is from near-field objects, such as trees and telephone poles. The bright bands running across the image from top to bottom are the undisturbed ionosphere. The dark bands running from top to bottom are ionospheric depletions or irregularities. These irregularities propagate in a southwesterly direction at 10-100 m/s. The electron density in these regions is nearly zero.

Data from 7/12/99 - Double Bubble (multiple vertical ionospheric depletions)
 GIMI Image: 12:47:12 UT, 30s
 LORAAS Data: 12:45:16 UT, 90s (15s missing)
 At 12:46:00, lat = -11°, lon = -162°, alt = 832 km

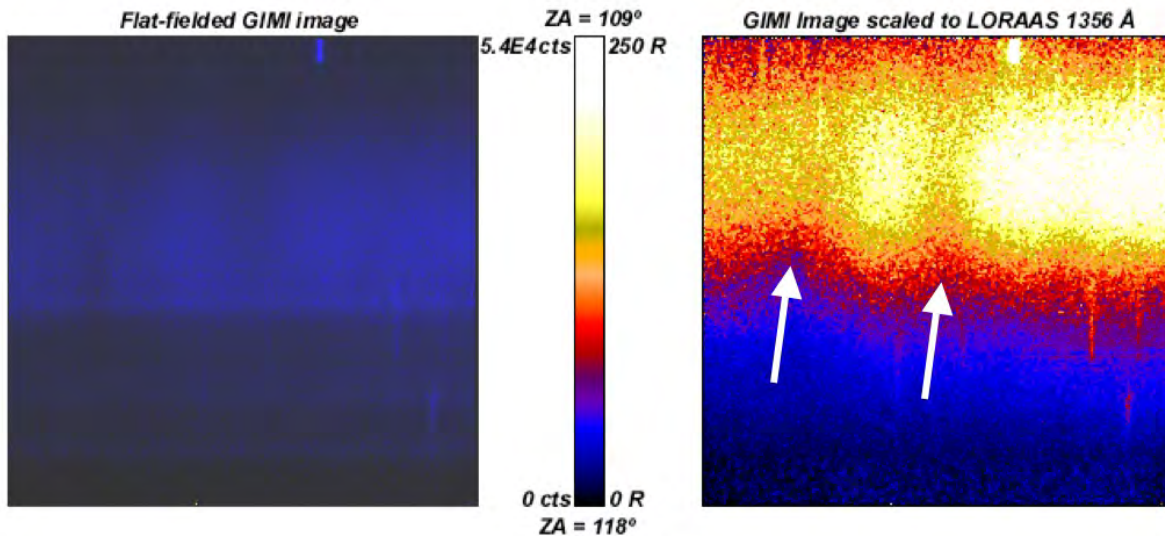


Figure 3: The left panel shows flat-field corrected data from the Global Imager of the Ionosphere (GIMI) experiment on the ARGOS satellite. The right hand panel shows the image after being converted to radiance units. The irregularities are indicated by the arrows in the image.

instabilities related to winds and electric fields. Figure 2 shows an all-sky camera map of a mid-latitude irregularity; this irregularity contains structures with 100 km dimensions. Figure 3 shows images of irregularities observed from space by the Global Imaging Monitor of the Ionosphere on the *Advanced Research and Global Observing Satellite (ARGOS)*. This image is of the Earth's limb and shows the vertical extent of two low-latitude irregularities. Using VISION, we hope to combine data sets like the two sets of images in Figure 2 and 3 to produce a three-dimensional reconstruction of the electron density. By knowing the local time of the bubble or irregularity, the scintillation properties can be predicted. VISION contains a tri-band (150, 400, and 1066.7 MHz) coherently radiating radio beacon with which to determine the electron density below the satellite using Computerized Ionospheric Tomography. In addition, the beacon can be used with the ground-based CIT receivers to study the scintillations caused by the irregularity.

How do the ionosphere and thermosphere respond to dynamical influences? Recent theoretical calculations have indicated that the troposphere and stratosphere can perturb the ionosphere and thermosphere. Ionospheric irregularity structures require a “seed” perturbation to initiate the instability. Winds, gravity waves, or electric fields may cause the necessary perturbations. Gravity waves may be induced by tropospheric dynamics. The effects of large tropospheric storms on the ionosphere and thermosphere cannot be observed from the ground, as the storm precludes gathering data. This is not a problem from space. This experiment will be able to look for ionospheric and thermospheric perturbations caused by tropospheric weather.

Daytime Observables

The principal ionospheric observable during the daytime will be the O II 83.4 nm lines. The O II 83.4 nm emission is produced principally by photoionization excitation of O at altitudes below 150 km. The 83.4 nm photons emitted downward are absorbed by the atmosphere. Those photons emitted upward enter the F-region ionosphere (altitudes 200–600 km) where they undergo multiple resonant scattering by

F-region O^+ ions. During the scattering process, the photons are entrapped and radiance field picks up a signature that is dependent on the altitude distribution of the O^+ ions. The VISION will primarily use the 83.4 nm emission to characterize the ionosphere at the Earth's limb.

The VISION will observe the O I 135.6 nm emission from neutral oxygen to infer the O density. On the dayside, this emission is primarily excited by photoelectron impact. However, there are contributions from radiative recombination and $O^+ - O^-$ mutual neutralization.

The VISION will measure the N_2 Lyman–Birge–Hopfield (LBH) band radiance near 143.0 nm to infer the N_2 density. This emission is excited solely by photoelectron impact. The 1403.0 nm bands are also near the peak of the O_2 Shumann–Runge continuum (Meier, 1991), which will allow the retrieval of the O_2 density by inversion of the limb measurements.

Nighttime Observables

The VISION will observe the O I 135.6 nm emission at night. This emission is excited primarily by radiative recombination of O^+ and electrons with a weaker contribution from $O^+ - O^-$ neutralization. The O^+ and therefore the electron density (by charge neutrality) can be inferred by inversions of the nighttime radiance. Tomographic inversion of the nighttime limb and nadir radiances will be used to produce the volumetric density field.

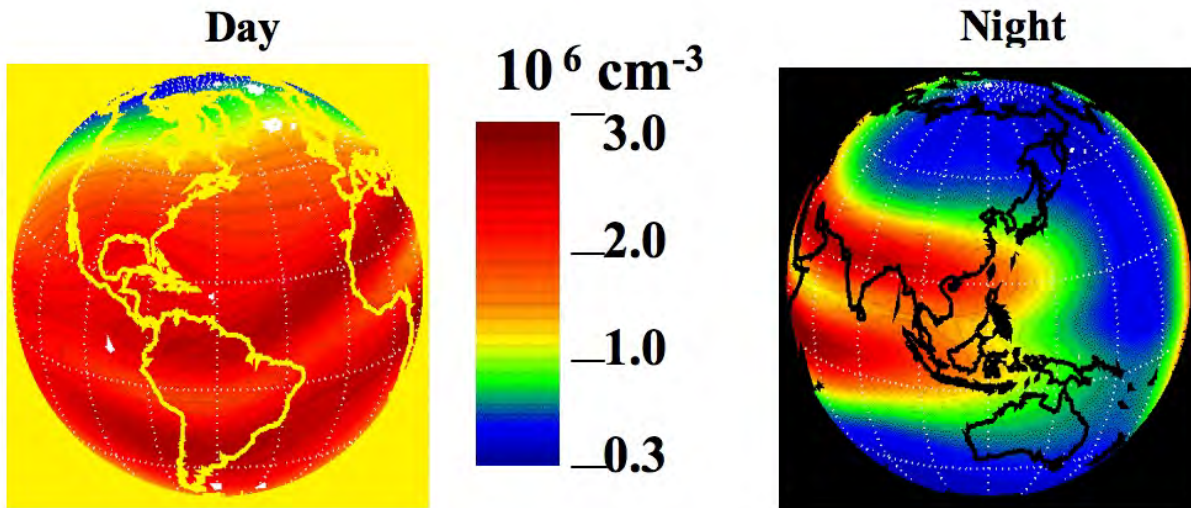


Figure 4: Map of the peak F-region electron density for mid-solar cycle conditions (10.7 cm flux of 140 solar flux units) at vernal equinox for day (left) and night (right). The peak density varies by an order of magnitude over the globe.

Radiances

Nighttime: The peak radiances were estimated by running the *International Reference Ionosphere* (IRI-90: Bilitza, 1990) for solar minimum and solar maximum conditions. Figure 4 shows a global map of the mid-solar cycle (10.7 cm solar flux of 140 solar flux units) peak electron density from the IRI-90. The radiances were then calculated by squaring the electron densities from IRI-90 and performing the line-of-sight integrations. The result was then scaled by the radiative recombination rate coefficients taken from Melendez-Alvira *et al.* [1999] and converted to units of Rayleighs. The limb radiances were derived from the nadir radiances by scaling them upward by a factor of 20 to account for limb brightening. The expected nadir radiances are: O II 83.4 nm less than 10 R, O I 135.6 nm between 0.1 R and 300 R (peak electron density $1 \times 10^5 - 5 \times 10^6 \text{ cm}^{-3}$), and N_2 LBH 143 nm approximately 0 R. The expected limb radiances are: O II 83.4 nm less than 10 R, O I 135.6 nm between 0.1 R and 3 kR, and N_2

LBH 143 nm approximately 0 R. Although the N₂ LBH and O II 83.4 nm emissions are excited by sunlight so the nighttime radiances are expected to be nearly zero, some O II 83.4 nm emission has been observed at night (*Chakrabarti et al.*, 1984). This emission is thought to be scattered over from the dayside by O⁺ ions in the magnetosphere.

Daytime: The nadir radiances were taken from *Meier* [1990]. The limb radiances were derived from these nadir radiances by scaling them upward by approximately a factor of 20 to account for limb brightening of the optically thin O I 135.6 and N₂ LBH 143.0 nm emissions. The O II 83.4 nm emission is optically thick and was instead scaled upward by approximately a factor of 2 to account for limb brightening. The expected limb radiances are: O II 83.4 nm between 1 R and 2 kR, O I 135.6 nm between 1 R and 20 kR, and N₂ LBH 143 nm between 1 R and 1 kR. The expected nadir radiances are: O II 83.4 nm between 1 R and 1 kR, O I 135.6 nm between 1 R and 1 kR, and N₂ LBH 143 nm between 1R and 100 R.

3. INSTRUMENT DESCRIPTION

The instrument design was driven by three basic requirements. First, the integration time to gather an image was selected to be 100 seconds or shorter for the limb imagers. The equatorial irregularities propagate Eastward at approximately 10–100 meters per second. The mid-latitude irregularities propagate southwesterly at 10–100 meters per second. The 100 second integration time ensures small smearing due to the irregularity's motion during an integration. Secondly, the limb imagers must image over 15° × 15° fields-of-view and the nadir imager must image along its 15° long slit at 0.5° spatial resolution. The 15° degree fields-of-view correspond to approximately 1000 km × 1000 km coverage at the limb. Thirdly, the experiment must have 10 km spatial resolution. The 10 km × 10 km resolution requires at least 100 resolution elements across the field-of-view. The spatial resolution requirement is driven by the requirement to spatially map irregularities that show spatial structures from the sub-kilometer level upward to the several thousands of kilometers. The electron densities vary from densities typical of the unperturbed ionosphere to nearly zero inside the depleted regions; this implies that the radiances vary from typical nighttime ionospheric radiances in the unperturbed regions to essentially zero in the depleted regions.

The VISION consists of three ultraviolet limb imagers operating at 83.4, 135.6, and 143.0 nm and a nadir cross-track scanner operating at 135.6 nm. During the daytime, the intensity gradients are much smaller than they are at night. Thus, the primary data sources during the daytime are the limb imagers. At night, the 83.4 nm and 143.0 nm signals are essentially zero, as the O II 83.4 nm and N₂ LBH emissions are excited by sunlight either directly by photoionization or indirectly by photoelectron impact. The primary nighttime data will be provided by the 135.6 nm cross-track scanner and the 135.6 nm limb imager.

Far-Ultraviolet Limb Imagers: The far-ultraviolet imagers consist of an off-axis parabolic telescope, which is fed by a three reflection filter assembly. The light is imaged onto an imaging microchannel plate detector. The detectors will be of the sealed tube type with a cesium-iodide photocathodes. The readout for the detector will be either a wedge-and-strip anode or a codacon. The telescope mirror is an 80 mm focal length, 2.5 cm diameter, first surface mirror with magnesium fluoride over aluminum coating on a zerodur blank. The filters are the three reflection bandpass filters developed for and flown on the Ultraviolet Imager (UVI) on the NASA POLAR satellite (*Zuckic et al.*, 1993). These filters have approximately 2.5 nm passbands (FWHM) and peak reflectances of 30–50%.

Extreme-Ultraviolet Limb Imager: The extreme-ultraviolet imager consists of a 80 mm focal length three mirror telescope with integral passband filters feeding an imaging microchannel plate detector. To further reduce the scattered light due to the bright Lyman-alpha geocorona, an indium/tin thin film filter will be used just prior to the detector face. The passband filters will be similar to those described in *Zuckic et al.* (1992).

Far-Ultraviolet Cross-Track Scanner: The far-ultraviolet cross-track scanner will consist of an off-axis parabolic telescope, which is fed by a three reflection filter assembly. The light is imaged onto an imaging microchannel plate detector. The detector will be of the sealed tube type with a cesium-iodide photocathode. The readout for the detector will be either a wedge-and-strip anode or a codacon. The telescope mirror is an 80 mm focal length, 50 mm diameter, first surface mirror with magnesium fluoride over aluminum coating on a zerodur blank. The filters are the three reflection bandpass filters developed for and flown on the Ultraviolet Imager (UVI) on the NASA *POLAR* satellite (Zuckic *et al.*, 1993). These filters have approximately 2.5 nm passbands (FWHM) and peak reflectances of 30–50%. The field-of-view will be scanned across the Earth's disk using a scanning mirror, which will be coated with magnesium fluoride over aluminum. The scan mirror must scan across 60° in 25 seconds. The scanner's slit subtends approximately 200 km when projected from the satellite at 1000 km altitude to 250 km. The satellite traverses this distance in approximately 26 seconds. The integration time of the cross-track scanner is then 0.2 seconds if the 10 km spatial resolution cross-track is to be maintained.

Sensitivity: We estimated the far-ultraviolet limb imager's sensitivity for the baseline design described above. The following formula was used to estimate the sensitivity of the FUV imager:

$$S = \frac{10^6}{4\pi} RTQA\Omega$$

where the reflectance of the mirror, R , is 0.8; the quantum efficiency of the detector, Q , is 0.1 (Siegmond *et al.*, 1987); the filter transmittance, T , is 0.5; the area of the primary mirror, A , is 5 cm²; the solid angle, Ω , is 0.063 steradians; and the constant $10^6/4\pi$ converts the results to Rayleighs. The resulting sensitivity at 135.6 nm is 1000 counts per second per Rayleigh. This sensitivity is approximately the same for the N₂ LBH 143 passband. During the daytime, this sensitivity will produce count rates that can damage the detector. The FUV limb imagers will be stopped down to reduce the entrance aperture area by an order of magnitude for the N₂ imager and by two orders of magnitude for the 135.6 nm imager.

We estimated the cross-track scanner's sensitivity for the baseline design described above. The following formula was used to estimate the sensitivity:

$$S = \frac{10^6}{4\pi} R^2 TQA\Omega$$

where the reflectance of the mirror, R , is 0.8; the quantum efficiency of the detector, Q , is 0.1 (Siegmond *et al.*, 1987); the filter transmittance, T , is 0.5; the area of the primary mirror, A , is 20 cm²; the solid angle, Ω , is 2.3×10^{-3} steradians; and the constant $10^6/4\pi$ converts the results to Rayleighs. The resulting sensitivity at 135.6 nm is 117 counts per second per Rayleigh. During the daytime, this sensitivity will produce count rates that can damage the detector. The instrument will be stopped down to reduce the entrance aperture area by an order of magnitude.

The sensitivity of the EUV imager was estimated using the following formula:

$$S = \frac{10^6}{4\pi} RTQA\Omega$$

where the combined reflectance of the mirrors/filters, R , is 0.1 (reflectance from Zuckic *et al.*, 1992 scaled by a factor of 2); the quantum efficiency of the detector, Q , is 0.1 (Siegmond *et al.*, 1987); the indium/tin filter transmittance, T , is 0.1; the area of the primary mirror, A , is 5 cm²; the solid angle, Ω , is 0.063 steradians; and the constant $10^6/4\pi$ converts the results to Rayleighs. The resulting sensitivity at 83.4 nm is 25 counts per second per Rayleigh.

Signal-to-Noise Ratio: We estimated the signal-to-noise ratio for the sensitivities of the instruments from above combined with the minimum radiances from above. We assumed that there was negligible

dark count (reasonable for microchannel plate detectors with cesium iodide photocathodes) assuming 10×10 km pixels (dividing the total count rates by 10^4 pixels) and an integration time of 100 seconds. The following formula was used:

$$N = S \Delta t I / (10^4)$$

where S is the sensitivity; Δt is the integration time (100 s for the limb imagers and 0.2 s for the cross-track scanner); I is the minimum radiance; and 10^4 is the number of pixels in the image (30 for the long slit image of the cross-track scanner). The resulting numbers of counts for the limb imagers are: 0.25 counts at O II 83.4 nm, 10 counts at O I 135.6 nm, and 10 counts at N₂ LBH. Assuming the data are shot noise limited – the noise is the square-root of the number of counts, the signal-to-noise ratios for the limb imagers are: 0.5 at O II 83.4 nm, 3 at O I 135.6 nm, and 3 at N₂ LBH 143 nm. The resulting number of counts for the cross-track scanner is 0.07 counts for a signal-to-noise ratio of 0.26. Although the signal-to-noise ratio of the 83.4 nm limb imager and the cross-track scanner are less than the required signal-to-noise ratio of 3 set down as an observational requirement, they can be improved by accepting lower spatial resolution by binning pixels. This assessment is true only for the weakest signals, for the strongest signals there will be very high signal-to-noise.

4. SUMMARY

We presented an overview of Volumetric Imager of the Ionosphere (VISION) a new instrument being built by the NRL. This instrument is designed to infer the volumetric distribution of the ionospheric electron density and the thermospheric neutral density over a 1000 km wide by 500 km thick slab of atmosphere beneath the satellite. The VISION sensor characterizes the three-dimensional of the ionosphere to develop global climatology of irregularity occurrence, investigate the influence of the background ionosphere on irregularity formation, and study irregularity development and morphology. The VISION contains a dual-frequency GPS sensor and a tri-band coherently emitting radio beacon. These additional sensors provide validation of the VISION measurements. The tri-band frequency beacon can also be used to study the scintillations produced by the irregularities.

5. ACKNOWLEDGMENTS

We are indebted to G. R. Carruthers and P. Walker for providing the GIMI images and to J. J. Makela and M. C. Kelley of Cornell University for providing the ground based irregularity images. This work was supported by the Office of Naval Research through the Ionospheric Specification and Forecasting Accelerated Research Initiative.

6. REFERENCES

- Bilitza, D., *International Reference Ionosphere 1990*, National Space Science Data Center, NSSDC/WDC-A-R&S 90-20, Greenbelt, Maryland, 1990.
- Chakrabarti, S., Kimble, R., and Bowyer, S., "Spectroscopy of the EUV (350–1400 Å) Nightglow", *Journal of Geophysical Research*, **89**, 5660, 1984.
- Meier, R. R., "Ultraviolet Spectroscopy and Remote Sensing of the Upper Atmosphere", *Space Science Reviews*, **58**, No. 1 & 2, 1-186, 1991.
- Melendez-Alvira, D. J., Meier, R. R., Picone, J. M., Feldman, P. D., and McLaughlin, B. M., Analysis of the Oxygen Nightglow Measured by the Hopkins Ultraviolet Telescope: Implications for Ionospheric

- Partial Radiative Recombination Rate Coefficients, *Journal of Geophysical Research*, **104**, 14901, 1999.
- Siegmund, O. H. W., Everman, E., Vallerger, J. V., Sokolowski, J., and Lampton, M., "Ultraviolet Quantum Efficiency of Potassium Bromide as an Opaque Photocathode Applied to Microchannel Plates", *Appl. Optics*, **26**, No. 17, 1987.
- Strickland, D. J., Craven, J. D., and Daniell, R. E., "Global Enhancements in O/N₂ Recorded by the DE-1 FUV Imager During Geomagnetically Disturbed Periods", *Eos Trans. AGU*, **84** (20), Spring Meeting Suppl., Abstract SA62A-02, 2001.
- Zuckic, M., Torr, D., Kim, J., and Torr, M., "Extreme Ultraviolet Filters for 58.4 and 83.4 nm", *Proc. SPIE* **1744**, 178, 1992.
- Zuckic, M., Torr, D. G., Kim, J., Spann, J. F., and Torr, M. R., "Filters for the International Solar Terrestrial Physics Mission Far-Ultraviolet Imager", *Optical Engineering*, **32**, 3069, 1993.

Heterogeneous Measurements for Advances in Space Science and Space Weather Forecasting

Scott Budzien¹, Damien Chua¹, Clayton Coker¹, Patrick Dandenault¹, Kenneth Dymond¹, Andrew Nicholas¹, Andrew Stephan¹, Richard Doe², Geoff Crowley³

¹*Space Science Division, Naval Research Laboratory, Washington, DC*

²*SRI International, Menlo Park, CA*

³*Atmospheric and Space Technology Research Associates, San Antonio, TX*

THEME: Atmosphere-Ionosphere-Magnetosphere Interactions

Introduction: Need for Heterogeneous Data Sets

Global-scale information about the space environment and its dependence upon solar and geomagnetic influences is more important than ever to society and decision makers. The ionosphere is the most variable component of the atmosphere: electron and plasma density can vary by orders of magnitude with respect to local time, latitude, season, solar cycle, and geomagnetic conditions. Ionospheric variability affects numerous civilian systems and military missions, including high-frequency radio and satellite communications, Global Positioning System (GPS) and radio navigation, satellite tracking and reentry, surveillance, and precision geolocation. The formation of ionospheric depletions in the post-sunset low-latitude ionosphere (equatorial spread-F, or ESF) is particularly effective in creating the RF scintillations that impair communication and navigation systems.

Unfortunately, though the need for space environment monitoring is growing and operational space weather models are becoming increasingly sophisticated and data-hungry, government agencies face the prospect of a shortfall in our space-based assets capable of monitoring the ionosphere-thermosphere system as the TIMED, DMSP, C/NOFS, and COSMIC missions wind down. In recent years, NASA has placed strong focus upon investigating the physics of the Sun, but devoted few missions and resources to study the Earth's thermosphere and ionosphere. In order to push space science to the next level, in which “seamless” measurement and modeling is performed to characterize space weather events from their source on the Sun to their effects in the Earth's upper atmosphere, more global measurements will be required to study the state of the thermosphere and ionosphere and the influences of space weather events.

In this resource-limited environment, however, providing data to current and future (full-physics) versions of global assimilative ionospheric models [e.g. Schunk et al., 2004; Komjathy et al. 2010] requires not simply generating a greater volume of homogeneous measurements (e.g. TEC from ground-based GPS and GPS radio occultation), but providing advanced assimilative models with heterogeneous data sets which provide complementary gauges for global ionospheric specification. Additionally, thermospheric temperature, composition, density, and (especially) winds as a function of altitude will prove essential for accurate, physics-based forecasting of ionospheric responses to space weather. These heterogeneous globally-acquired dataset should include sources such as randomly sampled GPS occultation data, in-situ particle and field measurements and mass spectrometers, and optical remote sensing to provide vertical profiles of composition, density, temperatures, and winds.

Case Study: GPS Radio Occultation and UV Nadir Photometry

The COSMIC constellation [Rocken et al. 2000], which combines GPS radio occultation (GPSRO) limb sensing with ultraviolet nadir photometry, illustrates the necessity of complementary measurement techniques to accurately characterize the ionosphere. UV photometers measure ionospheric structure without regard to a preferred transmitter-receiver geometry and can provide early detection of equatorial

spread-F (ESF) bubbles at nadir well before scintillation effects manifest at spacecraft altitudes. Nadir photometry integrates vertically to characterize horizontal ionospheric morphology and local gradients, while GPSRO integrates horizontally to reveal vertical structure. GPS-based TEC and photometer-based 135.6 nm UV radiance thus represent a heterogeneous data pair, and their mutual assimilation into GAIM significantly enhances the quantitative fidelity of the convergent ionospheric solution.

Operational Utility of 135.6 nm Radiance

Naturally occurring ultraviolet and visible emissions from the night sky have been used since the late 1960's as a means of determining the number densities of electrons and O^+ ions in the F-region ionosphere. The emissions are produced by radiative recombination of O^+ ions and electrons to produce atomic oxygen in an excited state which subsequently decays by emitting a photon. The use of the nighttime O emissions to remotely sense the F-region electron density is described by *Tinsley and Bittencourt* [1975], *Chandra et al.* [1975], *Dymond et al.*, [1996, 1997], and *Meier* [1991]. The far-ultraviolet (FUV) emissions are of the most interest for space-based ionospheric sensing, because the Earth's lower atmosphere strongly absorbs FUV radiation. Thus, there is no FUV background. The workhorse among the UV emissions is the atomic oxygen emission line at 135.6 nm. In the nighttime F-region ionosphere the optically thin 135.6 nm emission doublet arises from the recombination of atomic oxygen ions,



with a photon volume emission rate proportional to $7.3/10^{-13} \text{ cm}^3 \text{ s}^{-1}$ (α_{1356}) times the square of ionospheric density ($n_e * n_{O^+}$) in CGS units [*Melendez-Alvira et al.*, 1999]. If height-integrated to yield observed nightglow brightness ($4\pi I$), we see that the result is proportional to the path integral of ionospheric density squared (due to charge neutrality with $n_e = n_{O^+}$):

$$4\pi I = 10^{-6} \int_0^{\infty} \gamma \alpha_{1356} n_e(z) n_{O^+}(z) dz \quad (2)$$

Because this 135.6 nm radiance integral is similar to the TEC integral (line of sight path integral of n_e) both share assimilative value to ionospheric models such as GAIM and IDA4D [*Bust et al.*, 2004]. The 135.6 nm radiance, with an integrand proportional to the *square* of ionospheric density, is much more sensitive to ionospheric gradients than is TEC, making it more useful in assimilation programs to accurately reproduce ionospheric morphology.

Vertical TEC from Optical Remote Sensing

The vertical total electron content can be directly computed from UV photometer measurements of the oxygen 135.6 nm radiance. The radiance integral (Equation 2) can be evaluated directly when a Chapman layer approximation [*Chamberlain and Hunten*, 1987] is assumed for the electron (and ion) density distribution. The Chapman layer closely approximates the O^+ ion distribution at night and upon integration radiance in Rayleighs becomes:

$$4\pi I = 0.212 (NmF2/10^6)^2 H_{km} \quad (5)$$

where $NmF2$ is the peak electron density in units of cm^{-3} and H_{km} is the scale height in kilometers. The comparable total electron content integral yields:

$$VTEC = 0.33 NmF2/10^6 H_{km} \quad (6)$$

where $VTEC$ indicates the vertically integrated electron density (vertical total electron content, VTEC) in TECU (10^{12} cm^{-2}). Because 135.6 nm radiance is proportional to the product of the vertical total electron content and peak electron density, it provides a more sensitive detection of ESF bubbles and other small-scale ionospheric structures than does TEC. Equations 5 and 6 allow VTEC to be derived from radiance:

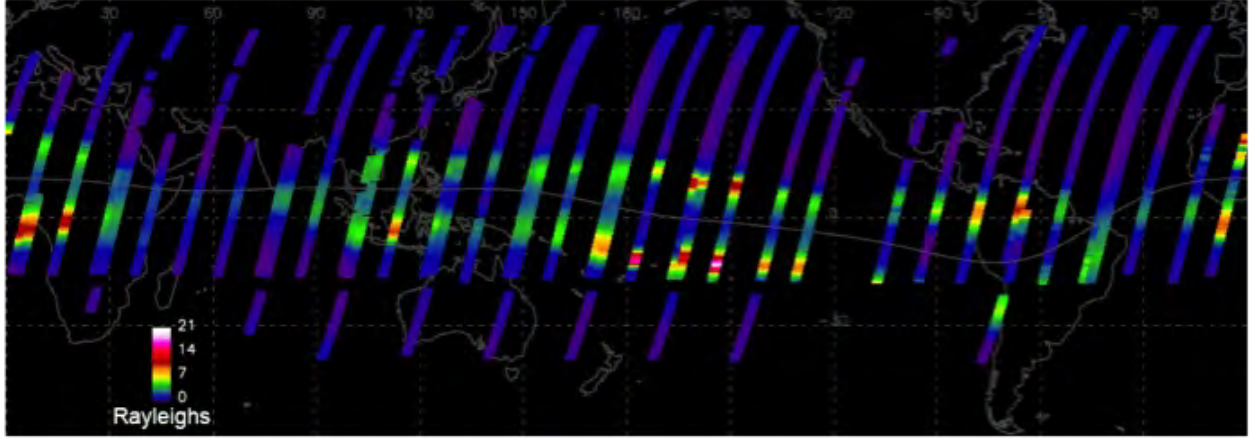


Figure 1: Nighttime ionosphere observed at 2100 local time over 24 hours by TIP on September 14, 2006 reveals strong longitudinal modulation of the equatorial anomaly arcs separation and brightness (electron density).

$$VTEC = 0.717 \sqrt{4\pi I H_{km}} \quad (7)$$

where the value of H_{km} can be extracted from an assimilative model or derived from analysis of GPS occultations.

UV Sensing of Ionospheric Gradients

Global scale dynamical processes (solar heating, waves, tides, electric fields, auroral precipitation) drive strong density gradients in the F-region ionosphere—gradients which can significantly distort transionospheric RF propagation and degrade communication and navigation assets. For example, the ionosphere is a major source of error for single frequency GPS receivers. Understanding the structure and evolution of such gradients guides an understanding for the strength of the dynamical drivers.

In particular, the low-latitude nighttime ionosphere is characterized by regions of large electron density enhancements north and south of the magnetic equator, known as the “equatorial anomaly” (Figure 1). An important process in the formation and evolution of the equatorial anomaly is the field-aligned plasma flow driven by the inter-hemispheric thermospheric winds. Winds transport the plasma up or down the field lines, producing asymmetries in the heights and densities of the anomaly crests. The magnitude, separation, height, and asymmetry of the equatorial anomaly crests reflect important dynamical driver parameters such as electric fields and winds. A nadir directed photometer is particularly sensitive to these horizontal ionospheric gradients, and Figure 2 demonstrates the fidelity of 135.6 nm measurements of the equatorial anomaly in direct comparison with a

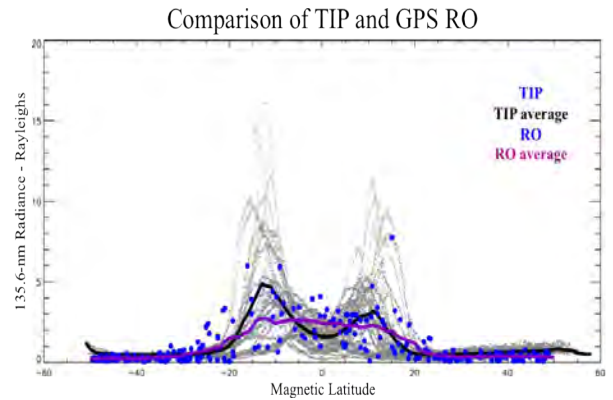


Figure 2: TIP radiance data for 24 hours (gray) on Sep 14, 2006 reflect the large variance of equatorial anomaly morphology at 2100 local time with longitude. The global average arc radiance is plotted in black. The predicted TIP radiances from corresponding COSMIC GPS occultation profiles (blue) exhibit lower spatial resolution, and the arcs are not resolved in their global average (violet).

TEC-based solution derived from GPSRO, which exhibit inherently lower horizontal resolution. TIP characterization of large-scale ionospheric features provides information about important ionospheric drivers.

The nighttime low-latitude ionosphere is also characterized by smaller-scale ESF bubbles or plumes which may result in rapid fluctuation or scintillation of satellite radio communication signals up to a few GHz in frequency. Both spread-F and scintillation events are particularly severe at low latitudes under geomagnetically quiet conditions. Figure 3 underscores the capability of 135.6 nm photometry to detect small-scale ESF plumes. TIP observations of these smaller features help identify and locate phenomena relevant to operational systems.

Relevance to Space Weather Sensing Objectives

The TIP concept is motivated by DoD space environment monitoring requirements and flow-down from those requirements to provide a large quantity of high quality measurements of ionospheric gradients, especially at night where the gradients are larger and more problematic. *Dymond et al.* [2009] used the TIP instruments on the COSMIC constellation to show that high sensitivity photometric measurements could be used in a tomographic inversion scheme to complement the very high precision GPSRO measurements made by those satellites and provide a more accurate and complete picture of the electron density distribution in the night-side ionosphere. The TIP sensors significantly “raised the bar” for photometric measurements of the nightside ionosphere. The combination of high spatial resolution and high sensitivity produced high precision ionospheric measurements even at mid-latitudes, during the early morning hours, at solar minimum conditions where the ionospheric densities are known to be extremely small.

UV Photometer and Day-to-day Morphology

Interest in the connection between lower atmospheric dynamics and influences upon ionospheric structure and irregularity formation has been increasing in recent years. Many researchers are cognizant of the longitudinal modulation of the nighttime equatorial arcs which has been attributed to atmospheric tides, and that these morphological patterns exhibit seasonal variations. These large-scale features have been observed in extraordinary detail by the TIP sensors over the years (Figure 1).

Less well-recognized perhaps, because many studies have focused upon multi-day averages of ionospheric morphology, is that these ionospheric morphological structures exhibit a high degree of day-to-day variability [Coker et al. 2010a]. Multiple TIP sensors were used to produce high resolution maps of the post-sunset equatorial anomaly for consecutive days (Figure 4). For the near solar-minimum conditions of September 2006 the maps clearly show 4-cell pattern associated with the diurnal eastward zonal tide (DE3) in the oscillatory separation of the crests. However, these patterns display significantly more dramatic variations than are indicated by earlier studies using multi-day averages of electron

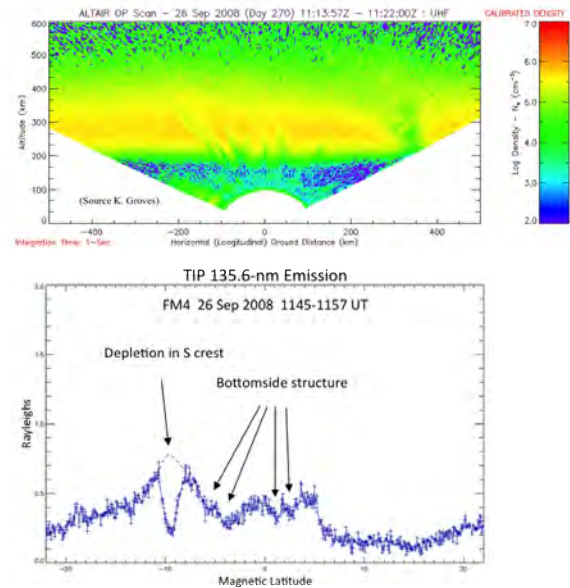


Figure 3: (a) Ionospheric depletion structures and gradients are present in ISR data from Kwajalein acquired during the Sep 26, 2006 COSMIC campaign [Coker et al., 2008]. (b) TIP data from a concurrent COSMIC overpass diagonal to the ISR scan exhibit signatures of depletion structures mapped along field lines. Against the background ionosphere TIP data reflect not only fully-formed depletions, but also non-penetrating bottomside structures associated with spread-F.

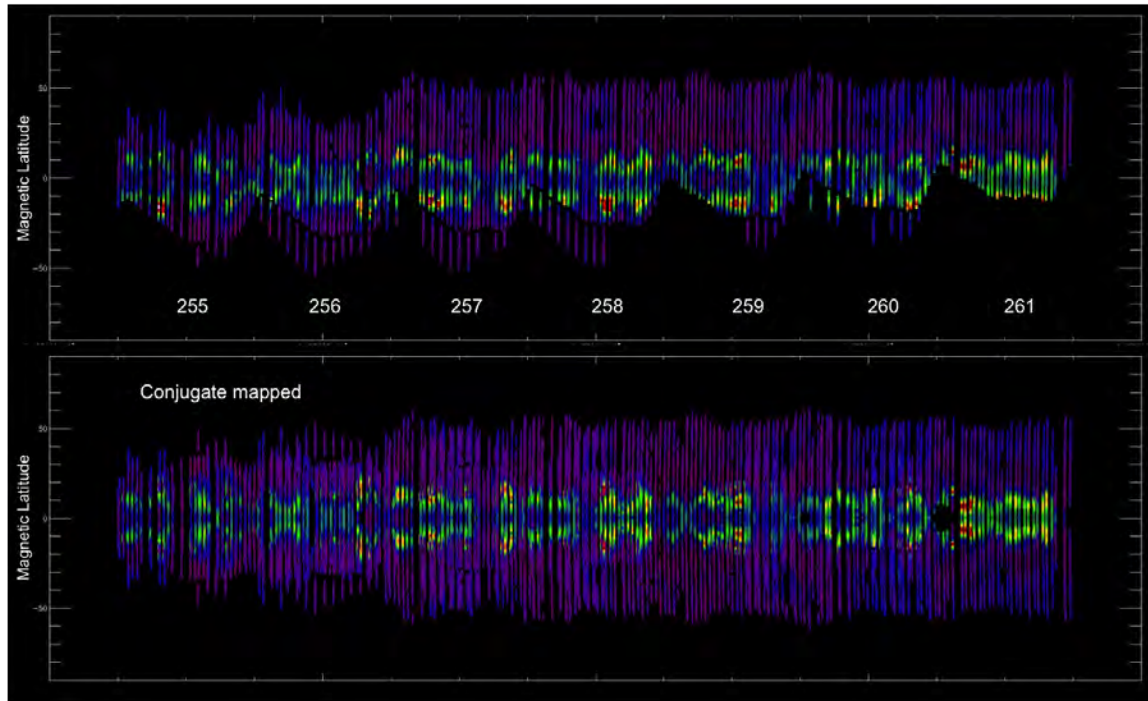


Figure 4: One week time sequence of post-sunset equatorial anomaly measurements from TIP. The upper panel shows actual measurements, while the lower panel shows averaged, conjugate mapped images. The periodicity of the 4-cell pattern is clearly seen in the variability of crest separation, and day-to-day changes in the pattern are visible as well. Comparison of the two panels reveals significant variability in N-S crest asymmetry in the measurements compared to conjugate averages.

density. In some longitude sectors the anomaly disappears completely on several days, then reappears some days later.

Multiday averages of this longitudinal pattern significantly underestimate the variation in crest magnitude, crest separation, and depth of the trough, while overestimating the latitudinal width of the crest. Daily variations from the average 4-cell pattern are typically 30-60% in the low latitudes, with 60-90% and even >90% excursions near the anomaly crests. Additionally, daily variations in magnitude of individual cells are observed and appear to occur on hemispheric scales, suggesting a large scale day-to-day variability in the global neutral wind pattern.

What is the implication for global ionospheric specification using data assimilation? Areas, such as oceans, where the equatorial arcs are not adequately sampled by ground GPS to provide horizontal resolution, revert to the underlying climatological ionosphere model. Climatology is inherently incapable of accurately reproducing these highly variable low-latitude features due to the averaging used to derive the model. Significant errors are likely in vertical TEC, crest separation, and crest asymmetry.

GAIM Modeling Challenges with GPSRO and Ground GPS

With the launch of the COSMIC constellation of GPS radio occultation sensors in 2006, research began on assimilating data from multiple GPS radio occultations into global assimilative ionospheric models. These space-based observations provide the advantage of global coverage compared to ground-based GPS stations but suffer from low horizontal resolution. Plans are on the table to deploy even larger constellations of GPS radio occultation sensors. Will these larger constellations with double or quadruple the number of observations of the space environment drive the assimilative models to the next level of performance in resolution and accuracy?

Early in the COSMIC mission multiple spacecraft were in the same orbit, providing a high density of

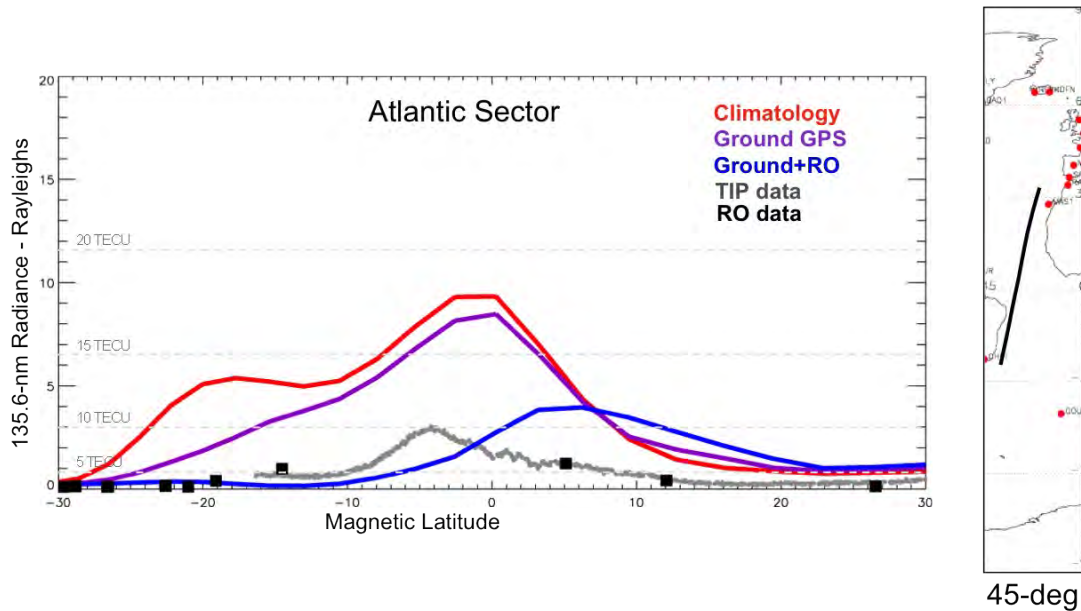


Figure 5: This example illustrates the difficulty modeling the anomaly using GPSRO where ground GPS coverage is sparse and cannot provide sufficient horizontal constraints on the anomaly crest morphology.

radio occultation data near the local time of that orbit – a condition analogous to proposed COSMIC-2 constellations with multiple sensors in each orbit. The resulting high density radio occultation data have been ingested into the USC global assimilative model, and the performance of the model are validated using high resolution UV photometer data from COSMIC [Coker et al. 2010b]. Multiple runs of the GAIM model were performed: first with climatology alone, then with ground-based GPS TEC values, and finally with ground-based TEC and satellite GPSRO data. Comparisons with TIP data reveal the horizontal performance of these retrievals and demonstrates the improvement obtained from ingesting the high density radio occultation data. When dense GPSRO data was added to the GAIM model, the ionospheric RMS error was reduced by 15% compared to ground-based GPS alone, with improvement clearly observed in the trough of the ionospheric anomaly.

However, no improvement was observed in specifying the crests of the anomaly. For instance, when horizontal constraints are limited (sparse ground GPS stations) and the anomaly is weak, the model often mispositions and/or broadens the anomaly crest because the satellite GPSRO has intrinsically low horizontal resolution. In the Atlantic sector example (Figure 5) the error is not a subtle or a sub-grid scale phenomenon, but an error of 10 degrees in latitude (1000 km!) in the crest location. In a case for the Kwajalein sector, the presence of ground GPS stations in the southern hemisphere was insufficient to constrain the magnitude and peak of the northern anomaly crest. In regions near landmasses with good ground GPS coverage the ionospheric specification from GAIM using ground GPS and dense satellite GPSRO matched TIP reasonably well.

Improving GAIM Results with UV Photometry

The nadir UV radiance and TEC measurements will significantly improve the quality of electron density profile (EDP) specifications produced by ionospheric assimilation models like GAIM. The top panel of Figure 6 shows electron density profiles produced by GAIM at 2100 local time (LT) as a function of magnetic latitude and altitude when ingesting only ground-based GPS measurements. The middle panel shows measurements of the 135.6 nm radiance at this local time by two different COSMIC/TIP sensors (red traces) as compared to the 135.6 nm radiance that GAIM would produce using the EDP shown above (blue trace). The GAIM output clearly does not match the TIP observations as the relative intensities of the northern and southern ionization crests are opposite of what is actually measured. When the TIP

observations are ingested into GAIM, the electron density profile produced by the model more accurately specifies the anomaly crests and the trough (Figure 6, bottom panel). This demonstrates how nadir-viewing UV photometer data such as those provided by TIP can drive the GAIM model toward an electron density profile specification that is much more consistent with the real ionosphere.

Moreover, nadir photometry provides data with a constant, simple viewing geometry and uniformly high quality over land and ocean, which facilitates both ingestion and error correlation analysis, which are critical aspects of assimilative modeling. The dependence of the UV signal upon the square of electron density places a very strong constraint upon the electron density and TEC predicted by GAIM. Assimilation of nadir photometer data can even improve vertical accuracy of ionospheric specifications, despite the fact that nadir photometry itself provides no vertical information. This surprising capability arises from the aliasing of horizontal and vertical signal along the many long slant TEC lines-of-sight from other data sources feeding GAIM. By improving GAIM's horizontal ionospheric specification, UV nadir radiance data allow GAIM to more accurately render TEC, NmF2, and even hmF2 derived largely from complementary assimilated datasets. The improvements to ionospheric specification from such heterogeneous data are particularly important in the highly structured nightside ionosphere.

Conclusion

The purpose of this case study is not to highlight the strength or weakness of any particular sensor system or model implementation, but to make a larger point. The upper atmosphere and ionosphere comprise a complex coupled system which is significantly undersampled. As space science progresses to model and understand more complex interactions among atmospheric regions and the global atmospheric response to space weather events, no single measurement technique will suffice to adequately characterize the system. Heterogeneous measurements with complementary strengths and limitations will be required for assimilation into physics-based models. Examples of heterogeneous data might include GPS radio occultation limb data and ultraviolet nadir photometry; GUVI/SSUSI cross-track O/N₂ maps coupled with SSULI in-track composition and density vertical profiles; global ionospheric images from geocentric orbit plus a constellation of in situ measurements within the atmosphere; or randomly-distributed occultations coupled with regular, tomographic slices of the atmosphere.

The various measurement techniques provide differing efficacy in constraining and improving global models for atmospheric specification and forecasting. Skill scores for model accuracy improvement can provide a metric for evaluating the efficacy of different measurements and provide guidance for prioritizing space weather sensor impact. Ultimately, however, no single measurement technique will be the “silver bullet” for characterizing and investigating the response of the atmosphere to space weather events: the measurements needed for mission in the the next decade will necessarily be heterogeneous datasets.

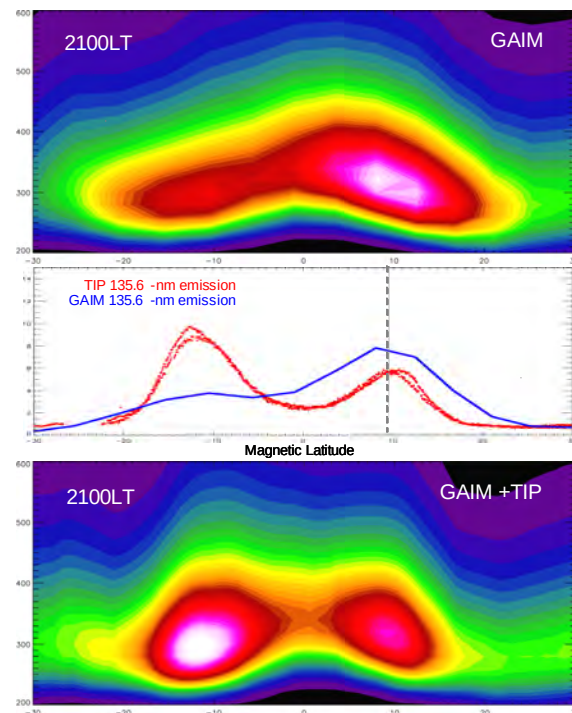


Figure 6: GAIM EDP specification without (top) and with TIP data (bottom). The middle panel show two TIP passes (red) and the initial GAIM predicted brightness (blue).

References

- Budzien, S., K. Dymond, C. Coker, and D. Chua, Tiny Ionospheric Photometers on FORMOSAT-3/COSMIC: On-orbit performance, *Solar Physics and Space Weather Instrumentation III, Proc. SPIE* **7438**, Silvano Fineschi, Judy A Fennely, eds., 13-1–13-11, 2009.
- Bust, G. S., T. W. Garner, and T. L. Gaussiran II, Ionospheric Data Assimilation Three Dimensional (IDA4D): A Global, Multi-Sensor, Electron Density Specification Algorithm, *J. Geophys. Res.*, **109**, A11312, doi:10.1029/2003JA010234, 2004.
- Chamberlain, J. W. and D. M. Hunten, *Theory of Planetary Atmospheres: An Introduction to Their Physics and Chemistry*, Academic Press, Inc., Orlando, FL., 1987.
- Chandra, S., E. I. Reed, R. R. Meier, C. B. Opal, and G. T. Hicks, Remote sensing of the ionospheric F layer by use of O I 6300-Å and O I 1356-Å observations, *J. Geophys. Res.*, **80**, 2327, 1975.
- Coker, C., K. F. Dymond, S. A. Budzien, D. H. Chua, S. Basu, R. Caton, K. M. Groves, and T. R. Pedersen, “Multisensor observations of low latitude irregularity development”, *EOS Trans. AGU*, **89**, Fall Meet. Suppl., Abstract SA24B-04, 2008.
- Coker, C., K. Dymond, S. Budzien, and D. Chua, “Day-to-day variability of the global post-sunset equatorial ionization anomaly”, *COSPAR*, Bremen, Germany, 2010a.
- Coker, C., P. B. Dandenault, K. F. Dymond, S. A. Budzien, D. H. Chua, X. Pi, A. Komjathy, and G. S. Bust, “Global Assimilative Ionospheric Modeling and COSMIC-2”, *Beacon Satellite Symposium 2010*, Barcelona, Spain, 2010b.
- Dymond, K. F., S. E. Thonnard, R. P. McCoy, and R. J. Thomas, “A Technique for determining F region electron densities using optical measurements of recombination radiation”, *Proceedings of the 1996 Ionospheric Effects Symposium*, 7-9 May 1996, Alexandria, VA., 346, 1996.
- Dymond, K. F., S. E. Thonnard, R. P. McCoy, and R. J. Thomas, An optical remote sensing technique for determining nighttime F region electron density, *Radio Science*, **32**, 1985, 1997.
- Dymond, K. F., S. A. Budzien, D. H. Chua, C. Coker, and J.-Y. Liu, Tomographic reconstruction of the low-latitude nighttime electron density using COMSICCOSMIC/FORMOSAT-3 radio occultation and UV photometer data, *Terrestrial, Atmospheric, & Oceanic Sciences*, **20**, 215, 2009.
- Komjathy, A., B. Wilson, X. Pi, V. Akopian, M. Dumett, B. Iijima, O. Verkhoglyadova, and A. J. Mannucci (2010), JPL/USC GAIM: On the impact of using COSMIC and ground-based GPS measurements to estimate ionospheric parameters, *J. Geophys. Res.*, **115**, A02307, doi:10.1029/2009JA014420.
- Meier, R. R., Ultraviolet spectroscopy and remote sensing of the upper atmosphere. *Space Sci. Rev.*, **58**, 1991.
- Melendez-Alvira, D. J., R. R. Meier, J. M. Picone, P. D. Feldman, and B. M. McLaughlin, Analysis of the oxygen nightglow measured by the Hopkins Ultraviolet Telescope: Implications for ionospheric partial radiative recombination rate coefficients, *J. Geophys. Res.*, **104**, 14901, 1999.
- Rocken, C., Y. H. Kuo, W. Schreiner, D. Hunt, and S. Sokolovskiy, COSMIC system description, *Terr. Atmos. Oceanic Sci.*, **11**, 21, 2000.
- Schunk, R.W., L. Scherliess, J.J. Sojka, and D. Thompson, Global Assimilation of Ionospheric Measurements (GAIM), *Radio Sci.*, **39**, RS1S02, doi:10.1029/2002RS002794, 2004.
- Tinsley, B. A., and J. A. Bittencourt, Determination of F region height and peak electron density at night using airglow emissions from atomic oxygen, *J. Geophys. Res.*, **80**, 2333, 1975.

Domestication of Scientific Satellites

Supriya Chakrabarti, Timothy Cook, Theodore Fritz (Boston University), Ron Polidan (Northrop Grumman)

Contact: Supriya Chakrabarti (617 353 5990; supc@bu.edu)

Summary

To fill the wide gap between CUBESATs and Small Explorer missions, and make space accessible to the new generation of explorers, a flexible and scalable satellite system design is necessary. Such a system will be ideal for many heliophysics studies.

Background

We describe the development of open standards, technologies and processes necessary experiments that require observations from small orbiting platforms, thereby creating a new class of science missions. In five years we will demonstrate key technologies and validate their readiness for heliophysics missions as well as enable space-based studies in other disciplines such as, planetary studies, remote sensing and astrophysics. To ensure a high probability of success, we will also articulate a process for developing high reliability and low-cost space-flight missions that will be based on successful past experiences. It will incorporate the best practices used in the modern networked community during the design, development as well as test and integration phases. In the end, it will enable a long list of scientific discoveries, flight-test key technologies, and train generations of experimentalists who will lead major missions of the future.

A number of sample science experiments that could immediately take advantage of such missions have been identified. By enabling a low-cost and high-frequency option for space-based missions, we will revitalize space research.

Overview

Early in the 20th century, automobiles appeared as extraordinary vehicles – and now they are part of life everywhere. Late in the 20th century, internet and portable phones appeared as innovations – and now are omnipresent requirements. At mid-century, the first satellites were launched into space – and now some 50 years later – “making a satellite” remains in the domain of highly infrequent events. Why do all universities, government laboratories and similar organizations not have their own research satellites? Why is the work force capable of doing so remarkably small? Why do highly focused science objectives that require just a glimpse from space never get a chance to fly? We will attempt nothing short of making mini satellites for

research (following the Moretto and Robinson, 2008 classification) as available as a small airplane. Flowing from such a technological innovation will be an array of science achievements spanning the full scope of heliophysics and by extension to other disciplines needing space-based observations.

The engineering and scientific infrastructure for an aerospace society is still in its adolescence due mainly to independently set goals that result in case-by-case unique solutions. This did not occur with computers, cell phones and GPS; they can all connect to each other. This initiative will do the same for satellite technologists, scientists, educators, students, and American society. Once successful, a possible proliferation of these satellites will become the (happy) problem to solve, not their rarity as we have today.

Historically, there have been two primary impediments to place a science experiment in orbit – high launch costs (Space Studies Board, 2000) and the high cost of spacecraft systems and related processes. The first problem appears to have been addressed through the availability of several low-cost (< \$10M) launch opportunities. This initiative will address the second.

Lowest-cost orbital space science missions – NSF sponsored CUBESATs – are approximately 1Kg in mass and cost approximately \$100K+ for the *science payload* and the *spacecraft* (which together forms a *satellite* or *observatory*). Our initiative will bridge the expensive missions which are “too large to fail” and the CUBESATs and suborbital programs (Figure 1). We will do so while deliberately involving students in meaningful roles and creating an environment of spaceflight mission development that encourages experimentation and innovation and together will provide a roadmap to “domesticate” scientific satellites.

This program will achieve this feat by replacing the current one-of-a-kind satellites (*craft-class*) into a customizable *commodity-class* (like the DELL™ computers). We will demonstrate that our efforts will shorten the development time of these missions so that students could be involved from experiment design to data analysis. It would therefore become the national model for hands-on training for these participants and will leave an important legacy in developing a scientifically and technically competent workforce.

America’s space enterprise is facing alarming problems as the graying workforce of the 1960’s moves rapidly towards retirement. Our national security, economic health, and environmental sustainability increasingly rely on knowledge collected by on-orbit platforms, but the nation has failed to develop a new pool of scientists and engineers who can serve this need. It seems likely that in addition to the countries in Europe, India, China and Japan will soon close our historical lead in space. At the same time, the competition for talented young people is growing. Today’s *Gen Net* students and their faculty from the *Baby Boom* are increasingly engaged in new methods of learning in a digital age. To meet these challenges, the current classroom practices must change, and the value of “out-of-class” experiences must be acknowledged. We will aggressively seek new ways to engage students in critical scientific and engineering knowledge while persuasively guiding their education to keep them in the field.

The science, technology and work-force development benefits of small satellite missions have recently been articulated (Gruntman, 2007; Moretto and Robinson, 2008; Baker and Worden,

2008). It has been convincingly argued that these missions are a cost-effective way of conducting science from space in many disciplines and, with their rapid turn-around time, can incorporate state-of-the-art technology, which current SMEX or larger programs cannot. When one considers the combined value of science return, technology maturing and training of the next generation of space scientists, technologists and managers, the program described here could rightfully be viewed as, with apologies to the overused commercial, priceless.

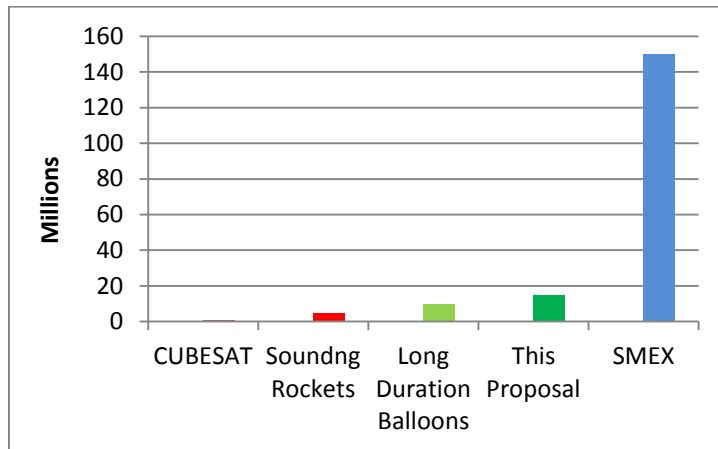


Figure 1: A preliminary estimate for the total cost of a satellite mission (including launch) following the approach outlined here. It assumed that the instrument complexity is similar to those designed, developed and flown on sounding rockets, and the integration and test activities and mission assurance program are similar to what is used by commercial, launch providers. This exercise revealed that our proposed plans are well-placed between the CUBESATs, the suborbital program and the smallest of the NASA orbital science missions – the Small Explorers (SMEX).

At the end of five years, we envision the following important products:

- Articulation of standards for spacecraft bus of this class, their major subsystems and their interfaces to each other and to the science payload
- Maturation of technology that are beneficial to heliophysics missions through flight tests
- Compilation of a listing of the best practices and processes that lead to a high quality flight missions

Objectives

Many questions from a wide range of science disciplines, from heliophysics to space physics, from microgravity to remote sensing, could be answered by observations from small satellites in Low Earth Orbit (LEO). Moretto and Robinson (2008) have listed 18 possible missions concerning Space Weather issues alone that could be conducted from small satellites at LEO (see their Table 1) (They defined small satellites as those that cost \$100M, take 2 – 3 years to develop and have 250 – 750 Kg mass). These focused experiments could be developed by university researchers at a cost that is significantly lower than the present norm.

For the heliophysics missions envisaged here, a mass between 100 and 250 kg is contemplated. Moretto and Robinson would characterize these as *Mini* satellites with price tag of \$50 – 75 M and a development time of approximately two years. *We can develop standards, technology and processes to lower the total cost of these missions to \$15 M or less.*

Historically, a major obstacle to low-cost space-based science missions has been launch cost. However, in the past decade, tremendous efforts by the space industry and government support have lowered the launch cost tremendously (Table 1). With such affordable launch costs, it is not difficult to imagine a total recurring mission cost of about \$15 M for future heliophysics mini satellite missions.

Table 1: Several Low-cost launch options are now available for small science satellite missions (data from Moretto and Robinson, 2008 and Baker and Worden, 2008).

Organization / launcher	Capability / cost	Comments
Sandia Natl. Lab. / Super Strypi	300 kg to LEO/\$9M	First launch: 2009
Orbital / Minotaur 1	500 kg to LEO/\$20M	First launch: 2006
Air Launch / QuickReach	635 kg to LEO / \$5M	Under development
SpaceX / falcon 1	670 kg to LEO / \$7M	First launch: 2008
Ride share using ESPA-ring / various	Up to six 180 kg to LEO / "few million dollars"	First launch: 2007
Ride share aboard ISC Kosmotras / Dnepr	Total capability 4500 kg to LEO / Variable	First launch: 1999

Recently conducted science missions of the Mini satellite class were called Student Explorer Demonstration Initiative (STEDI) and University-class explorers (UNEX). They demonstrated that this class of missions is capable of delivering excellent science at a fraction of the cost of SMEX class missions with somewhat higher risk. Three such missions have been launched – CHIPS (<http://chips.ssl.berkeley.edu/>), SNOE (<http://lasp.colorado.edu/snoe/>), and TERRIERS(<http://www.bu.edu/satellite/>). Of those, TERRIERS suffered a bus failure shortly after launch; SNOE conducted over 5 years of scientific measurements resulting in over 15 peer-reviewed and many more conference proceedings and abstracts totaling over 70 publications (see: <http://lasp.colorado.edu/snoe/publications/journals.html>). Finally, the CHIPS observations have demonstrated that the local area of the galaxy is not emitting radiation in the expected manner. To date these observation defy a coherent explanation (see Hurwitz et al. 2005); they will likely reshape our view of the structure and dynamics of the galaxy.

The three UNEX/STEDI missions have been conducted for a total cost (for all three missions) of around \$100M. This has resulted in two highly successful missions and one failure – a 66% success rate. Contrast this with the SMEX program with a cost of \$150M per mission and achieves a success rate of about 90% (nine successful SMEX-es and one failure – WIRE; SPIDR was cancelled). These data could be arguably interpreted as a cost of about \$167M per successful SMEX mission vs. \$33M per successful UNEX mission. Furthermore, one might expect the UNEX success rate to rise as our experience with such systems increases and more vendors gear up to meet this market.

Strawman Science mission

To demonstrate the diversity of science topics that could be studied from mini satellites, we describe below one viable mission. It is easy to envision other missions, some of which will yield classic discovery-mode science while others will need observation of unexpected phenomena that require quick response (the target of opportunity).

Demystifying the Ignorosphere

As an example of another capability afforded by the initiative, we consider a small satellite mission to study the atmospheric region between 130 km and 200 km. This is the region around the Earth where the magnetosphere connects to the lower region of the ionosphere, which overlaps the base of the thermosphere. Experimentally, this region of the upper atmosphere cannot be reached by airplanes or balloons; sounding rockets cannot probe it for any extended period of time. Satellites operating in this altitude regime return to the Earth only after a few days due to the strong atmospheric drag. As a result, this near-space is poorly understood and is called the *ignorosphere*. A mini-satellite-class mission can contribute much needed in-situ data, such as neutral winds and electric fields in addition to allowing a space platform for remote sensing instruments that could operate under fully sunlit condition. While the lifetime of a satellite without any orbit maintenance in this regime may be only a few days, 35-years of advancement in technology and global coverage from only one satellite will return sufficient high quality of data to merit *several* Ph. D. and masters level dissertations. While a few-days observations may not be cost-effective for explorer-class satellite missions, with its unique combination of cutting-edge science, technology and educational objectives, a low-cost approach such as that described here, is the ideal vehicle to address this important space physics question.

It is possible to list many other missions of similar scope in heliophysics as well as in other disciplines that require a short observation period, multiple satellites or some combination. While the value of missions like these is generally acknowledged, the current rate of about one \$100–300 M mission every couple of years and occasional \$1–5 B ones (primarily in astrophysics of solar systems exploration areas), cannot support such a broad range of experiments or accept the risk inherent in the more speculative – and thus more interesting – explorations. This initiative will make a new tool available to the science community to make discoveries, develop new observational approaches and provide a training ground for the next generation of space explorers. This will be its lasting legacy.

Implementation Approach

Unlike the risk-averse large science missions, our approach will include the infusion of advanced spacecraft technology such as power or communication systems or the use of new technology in structures and components that are *designed for demise* (D4D – see for example, ses.gsfc.nasa.gov/ses_data/2003/031007_Hull_presentation.ppt). Working as a university-industry-government team, we will validate this new approach through prototyping, vigorous testing (including suborbital flights) and summarize the findings as standards. To make this transition smooth, our team includes industry partners with direct experience with standards and standard spacecraft bus.

In most satellite mission the science instruments have the least flight heritage. Many of the spacecraft bus components are well developed and can be procured commercially. Thus, the task of satellite design and building often is that of system integration around a central processor with all subsystems (e.g., attitude control or command and data handling); the science instruments are input/output devices in this model. We will use conformal design tools (see below), to articulate a new methodology for the development and implementation of the proposed missions.

Technology infusion

We plan to investigate how the capabilities of an instrument could be significantly improved with selected infusion of technology. Boston University (BU) has revolutionized ground-based planetary studies which started with a simple rhetorical question: *If Galileo had owned a CCD* (Baumgardner and Mendillo1993). By replacing the eye with a sensitive CCD in the back of a 10-cm dia telescope (which is only slightly larger than the telescope Galileo had used and smaller than most amateur telescopes), the team discovered the largest object in the solar system, sodium tails of Mercury, Moon and a third tail on a comet! None of these discoveries could have been made had we stayed on the path of increasingly larger diameter telescopes.

On-board decision making software tools have been used to plan, execute tasks based on predetermined goals, to detect, analyze, and respond to science events, and to downlink only the highest value science data (see for example, <http://ai.jpl.nasa.gov/public/projects/ase/>). The development, refinement and application of tools for optimal flight and mission operations will be investigated.

Development of Open Standards for Science Missions

To meet the needs of the science and technical investigations described above, we will first develop an open standard for this class of satellite missions. They will include specifications (mechanical, electrical and data interfaces) for key spacecraft subsystems such as power, telemetry, and attitude control, so that an optimized mission could be developed.

We will create a catalog of subsystems similar to what is used in the sounding rocket community (http://fulcrum.gi.alaska.edu/hex/documents/Wallops_Info%5CSounding_Rocket_Handbook_July2001.pdf), so the prospective science missions will have ready choices. Finally, we will produce a standard of practices and process similar to the Goddard Open Learning Design (called GOLD rules, see <http://snebulos.mit.edu/projects/reference/NASA-Generic/GSFC-STD-1000.pdf>) describing such things as the number, timing and types of reviews, as well as allowable components consistent with the level of acceptable risk. There will also be provisions for timely adjustments in the future.

It is reasonable to ask how our efforts compare to other standards that are being developed by NASA or the Department of Defense's *Operationally Responsive Space* (ORS). NASA's lower priority for a small satellite program gave birth to NSF's CUBESAT initiative described above. The *Integrated Systems Engineering Team* (ISET) has published a comprehensive set of standards (<http://projects.nrl.navy.mil/standardbus/>). Most, if not all, mini-satellite-class missions could be accommodated by the requirements shown in Table 2 and will demonstrate the creativity of the space physics community once such a program become routine.

Table 2: Key Requirements for the system retain the necessary functions of expected science missions while eliminating complex, expensive and unnecessary features

ISSET Implementation *	CREST Implementation	Rationale
LEO and HEO orbits	LEO orbits only	System simplification; low mission cost
Availability: >99%;		
Reliability: >85% over mission lifetime (> 1 year)	Reliability: > 85% over mission life ($\geq$ a few weeks)	Consistent with the historic success rate of the sounding rocket program
Mission life: > 12 months	Mission life: > a few weeks	Lowers mission operation cost
Spacecraft mass: < 250 kg	Spacecraft mass: $\leq$ 200 kg	Design parameter
Orbit Average Power to payload: 400 W (LEO)	Orbit Average Power: $\leq$ 300 W (LEO)	Adequate for most astrophysics and other NASA SMD missions
Spacecraft bus-payload interface: defined	Interfaces to subsystems: defined	The observatories will be designed using conformal design principles
Satellites to be launched powered off	Satellites to be launched powered off	Simplifies integration and test activities and shortens development time
0.05° attitude control	0.01° attitude control	Similar to the sounding rocket capabilities
Slew rate of 2° per sec	Slew rate of 0.5° per sec	Adequate for tracking astronomical targets
Expected recurring cost for the Spacecraft Bus: \$5M - \$25M	Expected recurring cost for complete mission: $\leq$ \$15M	Competitive with current Suborbital Program (see Figure 1)

* ISET data from <http://projects.nrl.navy.mil/standardbus/>

Summary

A Technology Center to transform small science satellite missions is suggested. It will incorporate open standards, infuse state-of-the-art technology, use a virtual organization and develop a blue print for low-cost satellite missions for science exploration while involving students and professionals from a diverse background. This mini satellite initiative will revitalize the enthusiasm in experimental space research that has suffered significantly in recent decades from rare “big satellite missions” and make space accessible to a broad community.

References

- Baker, D** and **Worden S**, EOS, *Trans. Am. Geophys. Union*, **89**, 33, 2008.
- Baumgardner J.** and **Mendillo, M.**, *Sky and Telescope*, **85**, 19, 1993.
- Gruntman M**, AIAA-2007-6042, Space-2003, Long Beach, Calif. 2007.
- Higgins M** and **Kram K**, *Acad. Mgt. Rev.*, **26**, 264-288, 2001.
- Hurwitz M**, **Sasseen T** and **Sirk M**, *Astrophys.J.* **623** 911-916, 2005.
- Kanter R**, *Men and Women of the Corporation*, New York: Basic, 1977.
- McKnight S**, **Tadmor G**, **Cole W**, **Everbach C**, **Ruane M**, *ASEE Annual Conference*, Albuquerque, NM, June 2001.
- Mill J**, **O'Neil R**, **Price S**, **Romick G**, **Uyp O** Manuel, **Gaposchkin E**, *J. Spacecraft and Rockets*, **31**, 900, 1994.
- Morales C**, **Gomez J**, **Trapero J**, **Bowyer S**, **Edelstein J**, **Korpela E**, *Astrophys. and Space Sci.*, **263**, 393-396, 1998.
- Moretto T** and **Robinson, R. M.**, *Space Weather*, **6**, S05007, doi:10.1029/2008sW000392.
- Space Studies Board**, *Assessment of Mission Size Trade-offs for NASA's Earth and Space Science Missions*, Natl. Acad. Press, Washington, D. C., 2000.

An Ionospheric Modification Facility for the Magnetic Equator

J. L. Chau, Jicamarca Radio Observatory, Peru (jorge.chau@jro.igp.gob.pe)

S. A. Gonzalez, Arecibo Observatory, Arecibo, Puerto Rico, USA (sixto@naic.edu)

M. P. Sulzer, Arecibo Observatory, Arecibo, Puerto Rico, USA (msulzer@naic.edu)

E. Nossa, Cornell University, Ithaca, New York, USA (en45@cornell.edu)

B. Isham, Interamerican University, Bayamón, Puerto Rico, USA (bisham@email.bc.inter.edu)

Summary

This is a conceptual proposal for the deployment of an ionospheric modification facility, also called and HF facility or ionospheric heater, near the geomagnetic equator. A heater is a powerful high-frequency (HF), or shortwave, radio transmitter that can be used to induce a number of phenomena in the ionospheric plasma. Some of these phenomena provide insights into complicated plasma physics processes that may occur elsewhere in nature but that are difficult or impossible to explore in the laboratory or numerically. Other HF-induced phenomena provide diagnostics of naturally-occurring ionospheric phenomena.

HF ionospheric modification may affect the propagation of radio signals passing through the modified region, which is how the phenomenon was first discovered (i.e. the Radio Luxembourg effect [Bailey and Martyn, 1934]). HF transmissions may modify plasma density, increase electron and ion temperatures (which is why they are called heaters), create field-aligned plasma density irregularities that may be interrogated by small coherent scatter radars, and generate VLF and ELF radiation, which has practical societal utility (e.g. in underwater communications and underground remote sensing). By inducing turbulence, and through other processes, they add energy to electrostatic ion-acoustic and Langmuir modes that may be observed by incoherent scatter. These processes also induce optical and radio emissions that may be observed from the ground, and cause electron acceleration, although the detailed mechanism is unclear.

Heaters have been deployed in Plateville, Colorado, Fairbanks and Gakona, Alaska, Islote and Esperanza, Puerto Rico, Tromsø and Longyearbyen, Norway, Kharkov, Ukraine, and at Nizhniy Novgorod and other locations in Russia. These are all high- and middle-latitude sites. However, it has been suggested that the most spectacular heater-induced phenomena may be waiting to be discovered at low latitudes, under the geomagnetic equator.

Because of the close connection between heating effects and electrostatic waves, heaters are most productive when located close to an incoherent scatter radar. We propose that an equatorial HF facility should consequently be installed close to the Jicamarca Radio Observatory near Lima, Peru. Some initial work toward this end is already being performed.

Advancing solar and space physics

Ionospheric modification provides a powerful tool for exploring both artificially-driven plasma physics and natural aeronomomic processes in the upper atmosphere, and permits controlled, repeatable experiments, as opposed to making indirect inferences from large sets of disconnected observations. Consequently, hypotheses can be posed and tested, and experiments refined and repeated, so as to achieve rapid closure of existing science problems as well as the discovery of new ones (see reviews by Duncan and Gordon [1982], Frolov *et al.* [1997], and Gurevich [2007]).

In the ionosphere, a pump electromagnetic wave emitted from a high-power HF transmitter decays parametrically into a high-frequency byproduct wave (Langmuir, electron Bernstein/upper hybrid) and a low frequency byproduct wave (lower hybrid, ion acoustic, ion Bernstein, purely growing). Some of these byproducts can be observed directly using incoherent scatter radar (ISR), which is

the most incisive diagnostic available for studying heater-induced phenomena. Others result in the stimulated emission of electromagnetic waves (SEE), which can then be observed on the ground with simple HF receivers. While linear theory predicts the onset and early behavior of the waves produced in heating experiments, turbulence paradigms involving wave trapping and caviton production and collapse are required to account for the complex electrostatic (ISR) and electromagnetic (SEE) spectra that are observed.

Another feature of ionospheric modification experiments is electron heating and suprathermal electron production and electron acceleration. Instabilities driven by the pump wave heat the electrons up to about 3500 K, which can result in ion upwelling, and ion temperatures may increase by up to a few hundred K [Rietveld *et al.*, 2003]. Although the specific mechanisms involved are still unresolved, heater-induced plasma instabilities are also known to accelerate electrons up to 60 eV or more [Gustavsson *et al.*, 2005]. These electrons collide with oxygen atoms and nitrogen molecules and produce optical emissions much as natural electron impact from precipitation does in the aurora. Both thermal and suprathermal processes are believed to be at work, with 630-nm (red line) excitation being accessible through direct electron heating and 557.7-nm (green line) excitation through electron acceleration. Suprathermal electrons can be diagnosed both through incoherent scatter and optical techniques. Coordinated experiments provide the most direct path towards resolving the processes involved in electron acceleration.

A signature feature of ionospheric modification experiments is the production of small-scale artificial field-aligned plasma density irregularities (AFAIs) just below the HF reflection height [Fialer, 1974; Minkoff *et al.*, 1974]. The mechanism responsible is the thermal parametric instability (linear regime) [e.g., Fejer, 1979] followed by the resonance instability (nonlinear regime) [e.g., Vas'kov and Gurevich, 1977]. Using spaced-receiver radar imaging techniques, it is possible to generate three-dimensional radar images of the heater-induced irregularities with excellent spatial and temporal resolution (of the order of a second and a few hundred meters, respectively) [e.g. Nossa *et al.*, 2009]. AFAIs offer an incisive diagnostic of both ionospheric modifications and naturally occurring processes. Ionospheric heating effectively “shines a light” on existing ionospheric density layers, which can then be monitored and analyzed in much greater detail than conventional remote sensing techniques generally afford (e.g. Figure 1).

Another method for monitoring the background ionosphere using ionospheric modification was introduced by Bernhardt *et al.* [2003] in experiments on sporadic-E ionization layers at Arecibo. The method, termed “radio-induced aurora” or RIA, combines ionospheric modification with optical imaging. The idea involves emitting pump-mode radiation at a frequency below the F region peak frequency. Where there are no sporadic E layer patches, the radiation propagates into the F region and produces red line emissions at the F region interaction height. Where there are sporadic E layer patches the pump-mode radiation interacts in the E layer, producing gaps or “shadows” in the red line emissions and green line emission at E region altitudes. Airglow imagers can detect both phenomena, leading to spatial maps of the E layer structure in two horizontal dimensions.

Among the more practical applications of ionospheric modification are the generation of radio signals in the VLF (3 – 30 kHz), ULF (300 – 3000 Hz), and ELF (3 – 30 Hz) bands [Ferraro *et al.*, 1982]. Such signals are useful for undersea communications (i.e. with submarines) and subsurface imaging, both having strategic significance for the nation. Conventional antennas used in these bands are electrically short by necessity (compared to a wavelength) and therefore inefficient. Using HF transmissions we may attempt to alter ionospheric currents (periodically in time and perhaps also in space) and thereby cause them to radiate. Most efforts to date have involved the auroral electrojet current, although emissions can also be caused by driving diamagnetic currents or currents associated with the pondermotive force.

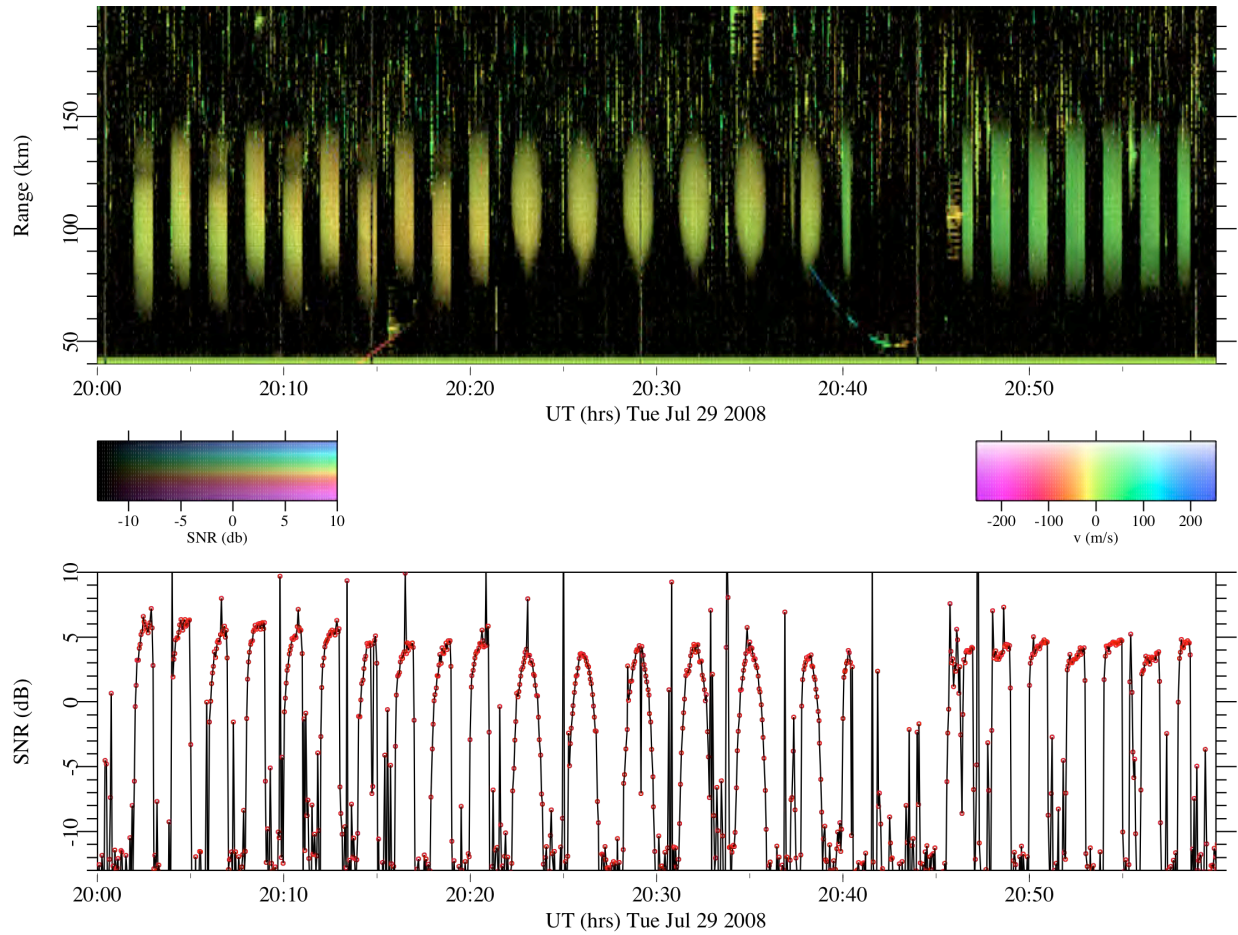


Figure 1. UHF radar range-time Doppler and SNR plots for three different experiments at HAARP during the summer of 2008. The morphology of the generated echoes changes depending on the parameters used. The experiments are repeatable, creating unique conditions for studying plasma physical and natural phenomena (from Nossa *et al.* [2009]).

Science investigations along many of the lines outlined above could also be conducted with a heating facility near the geomagnetic equator. More importantly, a number of new and potentially spectacular phenomena could also be explored. Some of these were first described in a review paper by *Erukhimov et al.* [1997], while others are derived from years of experience at HAARP, Jicamarca, and elsewhere. The most obvious objectives of an equatorial heater are outlined below.

Artificial periodic irregularities

A standing wave pattern is formed where the transmitted HF wave reflects back to Earth, creating nonuniform heating and quasiperiodic density irregularities. These irregularities can be probed using low-power HF waves at and close to the pump frequency. A modern ionosonde can be used for this purpose. By timing the irregularity onset and decay times, relaxation processes, state parameters, and transport coefficients can be determined for the natural atmosphere.

At middle and high latitudes, a given geomagnetic field line may connect, or short-circuit, the artificial irregularities, subjecting them to rapid dissipation by parallel ambipolar diffusion, particularly in the F region. This limits the utility of the technique as a practical diagnostic. At the magnetic equator, however, the artificial irregularities will form orthogonally to the geomagnetic field. Dissipation in the F region will then be limited by the electron cooling rate, which is very slow.

Erukhimov et al. [1997] estimated that the artificial irregularity amplitude will then be three to four orders of magnitude larger than for comparable high-latitude experiments, approaching something like 10% RMS for a heater with an effective radiated power of 100 MW. This would be a dramatic effect and would offer a versatile, cost-effective, exploratory tool for equatorial aeronomy.

Triggering equatorial spread F

Equatorial spread F (ESF) has been the cause celebre of equatorial aeronomy since its discovery over 70 years ago. The term refers to ionospheric interchange instabilities that occur frequently after sunset at latitudes between the Appleton ionization anomalies, generating ionospheric irregularities which evidence themselves through radar scatter and clutter, optical signatures, and radio scintillation and refraction. ESF disturbs operational communication and navigation systems, including GPS, and interferes with SAR imagery. Forecasting ESF is a priority of the National Space Weather Program and the Air Force C/NOFS satellite program. Forecasting efforts have proceeded slowly, however, and undetected triggers or inhibitors are often suspected to be at work.

Ionospheric modification offers a means of experimenting with ESF triggering processes and testing the sensitivity of the stability of the ionosphere to different perturbations. Electron heating in F-region modification experiments is due mainly to anomalous absorption associated with the excitation of plasma instabilities. The subsequent effect on the plasma number density in the lower F region is mainly controlled by photochemistry and depends on the altitude and the temperature dependence of the recombination coefficient. Since the recombination rate is a decreasing function of temperature, the tendency is for electron density to actually increase during heating experiments there, whereas density decreases during heating at higher altitudes.

Allen et al. [1974] reported being able to induce spread F reproducibly in their mid-latitude heating experiments. During heating experiments at Arecibo *Djuth et al.* [1987] measured large temperature and density perturbations that they attributed to the pre-existence of marginal dynamical instability in the ionosphere. At equatorial latitudes instability triggering is likely to have the most dramatic and consequential results, which could then be systematically investigated.

Parametric decay instabilities

Two classes of plasma instabilities occur during ionospheric modifications: (a) parametric decay instabilities, which occur where the HF pump frequency matches the plasma frequency, and (b) thermal parametric instabilities, which occur at somewhat lower altitudes where the pump frequency matches the upper hybrid frequency. The former drives Langmuir and ion acoustic waves and gives rise to Langmuir turbulence, electron acceleration, and related processes. The latter produces stimulated radio emissions and artificial field-aligned plasma density irregularities (AFAIs). One effect of AFAIs is to scatter away the incident pump radiation, causing anomalous absorption and inhibiting the growth of the parametric decay instabilities above, which as a result become difficult to study after the irregularities emerge.

At the magnetic equator, O-mode HF pump waves emitted strictly vertically will propagate in the quasitransverse mode, and the wave electric field will consequently be parallel to the geomagnetic field. Without a transverse electric field component, the pump wave will not be able to excite thermal parametric instabilities, and AFAIs will not be created. For HF rays having slightly off-vertical incidence, the HF waves will be quasi-longitudinal, and irregularity excitation will occur as it does at middle and high latitudes. However, there will be a narrow cone directly over the heating facility through which HF radiation will be able to propagate directly through to the reflection height. Parametric decay instabilities will be driven continuously within that cone, offering a new means of studying their long-term behavior and evolution.

Equatorial electrojet current modulation

Attempts to generate low-frequency radio waves at high latitudes with HF heating have suffered from the unpredictability of the auroral electrojet as well as from the adverse effects of particle precipitation and deviative absorption, which prevents the pump power from reaching the electrojet current. The equatorial electrojet, meanwhile, exhibits extremely reliable diurnal behavior, and low-latitude heating experiments do not suffer from the effects of precipitation. Preliminary tests carried out in the 1980's using the Jicamarca 50-MHz radar as an underdense heater demonstrated the plausibility of VLF generation using the equatorial electrojet. The installation of a heating facility would provide a full experimental and operational capability.

Other factors also suggest that the equatorial electrojet is more suitable for low-frequency wave emission than the auroral electrojet. *Papadopoulos* [2008] discuss how the radiation generated in the equatorial electrojet will couple much more efficiently into the Earth-ionosphere waveguide than radiation generated at high latitudes, which tends to couple into the magnetosphere. *Lehtinen and Inan* [2009] conducted full-wave numerical simulations of VLF/ELF emissions generated by a heater at the geomagnetic equator that show all of the radiated power contained within the Earth-ionosphere waveguide and a total generated energy flux 30 times greater than predicted for the auroral electrojet case under comparable conditions. The emitted radiation was furthermore directional, implying still higher effective radiated power by comparison. The equatorial electrojet contains a significant vertical current component which, when modulated, acts like a vertical dipole element, which is well suited for exciting the Earth-ionosphere waveguide.

The aforementioned predictions are remarkable and suggest that an important operational capability could emerge from an equatorial heater. From a space physics perspective, the scientific utility of the heater would come from testing theories and exploring the related plasma physics.

Modification of equatorial irregularities

A heater at the equator will be very useful for performing controlled experiments of equatorial mesospheric and 150-km irregularities. In the case of mesospheric echoes, for example, one can determine whether or not charged particles play an active role. Recent radar studies of 150-km echoes indicate the presence of both field-aligned irregularities and naturally-enhanced ion-acoustic waves [e.g., *Chau et al.*, 2009]. Moreover, the modulation of these echoes by extreme ultraviolet solar radiation indicates the presence of Langmuir turbulence, the controlled investigation of which would be another promising area of study for an equatorial heater.

Cost estimate

We have estimated the cost of a 100-MW ERP-class heating facility able to generate the heater-related phenomena described here, and comparable to the former HIPAS (Alaska), existing SPEAR (Norway), and new Arecibo systems. The cost estimates have been obtained with the help of colleagues and contractors constructing the new Arecibo heater, modified for local conditions in Peru. The main estimated costs are: (1) antenna and transmitter design and prototype construction and testing (~\$500,000), (2) final system construction and installation, including building and land preparation (between \$4 and \$8 million USD, depending on the transmitter option selected).

Preliminary studies indicate that a suitable array would consist of 6x6 cross-polarized log-periodic antennas. The design will utilize distributed, custom-made, solid-state, roughly 10-kW transmitters. Given the high efficiency expected from this transmitter system, and the total transmitter power, diesel or natural gas generators would be used. The expected operational cost, assuming 150 transmitting hours per year, would be around \$100,000/year.

Benefits

Below we describe some of the practical benefits to the scientific and broader communities.

Previous studies

The proposed concept has not been identified in previous NRC decadal surveys.

Significant contributions to panel themes

The proposed ground-based instrument provides significant contributions to the Atmosphere-Ionosphere-Magnetosphere Interactions panel theme as well as coupling to other themes.

Important science questions

This project addresses some of the most pressing basic science questions in AIM including the physics behind electron acceleration, wave-wave coupling, and ionospheric instabilities. It also has the potential of providing a new diagnostic of ionospheric parameters, rate constants, and transport coefficients, adding to and going beyond what incoherent scatter techniques currently provide.

Societal benefits and operations

Low-frequency electromagnetic waves can play an important role in submarine communication and subsurface remote sensing, and an equatorial heater represents a cost-effective means of achieving these capabilities. Forecasting of equatorial spread F is important to a number of space weather interests, and controlled, repeatable ionospheric experiments may accelerate progress in this area.

Other systems and programs complemented

This project complements existing ionospheric modification programs underway in Europe, Alaska, and Puerto Rico. A significant community of users surrounding the annual Radio Frequency Ionospheric Interactions (“Santa Fe”) Workshop and Polar Aeronomy and Radio Science (PARS) summer school already exists. Likewise, the project builds on existing infrastructure at Jicamarca and on the close ties between NSF, Cornell University, and the Peruvian Geophysical Institute (IGP).

Affordability

This project leverages existing infrastructure in Peru, and experience and design know-how developed through Arecibo’s recent heater deployment, which included the involvement of some IGP personnel. The cost will be very modest compared to previous heater projects.

Readiness

As described above, ionospheric heating is a well-known field. The novelty of the proposed idea relies on the location, i.e. under the magnetic equator. The main challenge would be to keep within a reasonable budget. Personnel and companies involved in the construction of the Arecibo heater would join the scientific and technical expertise at the Jicamarca Radio Observatory to carry out the construction and installation of the equatorial heater. Operation of the proposed heater could be handled in the same way as for the Jicamarca Radio Observatory, which has been jointly operated by the Peruvian Geophysical Institute, Cornell University, and NSF for over 30 years.

International plans and activities

The proposed heater would attract broad international participation. For example, both Western and Eastern Europe have long traditions in this field, and a correspondingly large community of interested and participating experimentalists and theorists.

References

- Allen, E. M., G. D. Thome, and P. B. Rao, HF phased array observations of heater-induced spread F, *Radio Sci.*, 9, 905–916, 1974.
- Bailey, V. A., and D. F. Martyn, Influence of electric waves in the ionosphere, *Phil. Mag.*, 22, 369, 1934.
- Bernhardt, P. A., N. A. Gondarenko, P. N. Guzdar, F. T. Djuth, C. A. Tepley, M. P. Sulzer, S. L. Ossakow, and D. L. Newman, Using radio-induced aurora to measure the horizontal structure of ion layers in the lower thermosphere, *J. Geophys. Res.*, 108(A9), 1336, doi:10.1029/2002JA009,712, 2003.
- Chau, J. L., R. F. Woodman, M. A. Milla, E. Kudeki, Naturally enhanced ion-line spectra around the equatorial 150-km region, *Annales Geophysicae* Vol. 27, 933–942., 2009.
- Djuth, F. T., B. Thidé, H. M. Ierick, and M. P. Sulzer, Large F-region electron temperature enhancements generated by high-power HF radio waves, *Geophys. Res. Lett.*, 14, 953–956, 1987.
- Duncan, L. M., and W. E. Gordon, Ionospheric modification by high power radio waves, *J. Atmos. Terr. Phys.*, 44(12), 1009–1017, 1982.
- Erukhimov, L. M., N. A. Mityakov, and B. Thidé, On the excitation of the ionosphere by high-power radio waves in the geomagnetic equator region, *Radiophys. Quantum Electron. (Engl. Trans.)*, 40, 165, 1997.
- Fejer, J. A., Ionospheric modification and parametric instabilities, *Rev. Geophys. Space Phys.*, 17, 135–153, 1979.
- Ferraro, A. J., H. S. Lee, R. Allshouse, K. Carroll, A. A. Tomko, F. J. Kelly, and R. G. Joiner, VLF/ELF radiation from the ionospheric dynamo current system modulated by powerful HF signals, *J. Atmos. Terr. Phys.*, 44(12), 1113–1122, 1982.
- Fialer, P. A., Field-aligned scattering from a heated region of the ionosphere – observations at HF and VHF, *Radio Sci.*, 9, 923–940, 1974.
- Frolov, V. L., L. M. Erukhimov, S. A. Metev, and E. N. Sergeev, Temporal behavior of artificial smallscale ionospheric irregularities: review of experimental results, *J. Atmos. Sol. Terr. Phys.*, 18, 2317–2333, 1997.
- Gurevich, A. V., Nonlinear effects in the ionosphere, *Uspekhi Fizicheskikh Nauk.*, 177(11), 1145–1177, 2007.
- Gustavsson, B., et al., The electron distribution during HF pumping, a picture painted with all colours, *Ann. Geophys.*, 23, 1747–1754, 2005.
- Lehtinen, N. G., and U. S. Inan, ELF/VLF wave radiation produced by an equatorial ionospheric heater, oral presentation at National Radio Science Meeting; Boulder, CO; Abstract HG2-9, January 5–8, 2009.
- Minkoff, J., P. Kugelman, and I. Weissman, Radio-frequency scattering from a heated ionospheric volume, 1, VHF/UHF field-aligned and plasma line backscatter measurements, *Radio Sci.*, 9, 941–955, 1974.
- Nossa, E., D. L. Hysell, C. T. Fallen, and B. J. Watkins, Radar observations of artificial E-region field-aligned irregularities, *Ann. Geophys.*, 27, 2699–2719, 2009.
- Papadopoulos, K., Pulsed artificial electrojet generation, AGU, FallMeet. Suppl., Abstract SA42A-01, 2008.
- Rietveld, M., M. Kosch, N. Blagoveshchenskaya, V. Kornienko, T. Leyser, and T. Yoeman, Ionospheric electron heating, optical emissions and striations induced by powerful HF radio waves at high latitudes: Aspect angle dependence, *J. Geophys. Res.*, 108, 1141, doi:10.1029/2002JA009,543, 2003.
- Vas'kov, V. V., and A. V. Gurevich, Resonance instability of small-scale plasma perturbations, *Sov. Phys. JETP Engl. trans.*, 46, 487–494, 1977.

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. HITDP 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.

A White Paper for National Research Council Heliophysics Decadal Study
Space Lidar Mission to Study Middle and Upper Atmosphere Dynamics and Chemistry

Xinzhao Chu

University of Colorado at Boulder

Elsayed R. Talaat

Applied Physics Laboratory, The Johns Hopkins University

Chester S. Gardner

University of Illinois at Urbana-Champaign

Adrian J. McDonald

University of Canterbury

Executive Summary

We discuss the scientific merits and technological feasibility of a new space mission to fly a Na Doppler lidar in near-Earth orbit to sense the Earth's mesosphere and lower thermosphere (MLT) globally. This instrument can make high-resolution temperature, wind, and Na density measurements from 75-115 km (MLT region) to enable scientists to delineate and understand the middle and upper atmosphere chemistry, structure and dynamics, especially the impact of gravity waves – the parameterization of which is a fundamental issue in current atmospheric modeling for climate and meteorology. The proposed space lidar is based upon mature resonance-fluorescence Doppler lidar technology, which has evolved from a robust heritage of technological advancements, many contributed by the middle and upper atmosphere community. The scientific community has been making Na lidar measurements for more than 40 years and airborne Na and Fe lidar measurements for 25 years.

Gravity waves are a key driver in the MLT but predicting their impact on atmospheric chemistry and dynamics is a major uncertainty in current atmospheric models. The proposed lidar mission is expected to bring closure to a large number of open questions by providing high-resolution measurements that can be used to characterize these waves and their effects in the MLT on a global basis. The proposed mission will also break new ground on measurements, as it will provide the best ever line-of-sight resolution of global temperature and wind profile measurements. At the same time, measurement of the global distribution gravity waves and their spectra will add to the understanding of key indicators of radiative cooling in the mesosphere, such as polar mesospheric clouds. Finally, the proposed mission will create a climatology of mesosphere and lower thermosphere winds and temperatures of unprecedented accuracy against which the data of future missions can be benchmarked.

Key scientific issues to be addressed by this lidar mission include:

- Global distribution of the gravity wave spectra and their impacts on the MLT region
- Relationship among gravity waves, zonal mean structures, planetary waves, and tides
- Global estimates of gravity wave forcing and transport which will constrain gravity wave drag (GWD) parameterizations in general circulation models in this region for the first time
- Global distribution and variability of meteoric Na, polar mesospheric clouds, and their heterogeneous interactions.

The Na Doppler lidar can be constructed and orbited in ~7 years on the order of NASA SMEX for nadir pointing to measure temperature, vertical wind and Na density, or on the order of NASA MIDEX for off-zenith pointing with two orthogonal beams to profile temperature, horizontal wind, and Na density globally with very high resolution and accuracy. This white paper is based on a two-year feasibility study conducted by two teams under the support of the European Space Agency.

1. Introduction

The acquisition of global temperature and wind measurements in the mesosphere and lower thermosphere (MLT) is given a high priority in the upper atmosphere community. The combination of wind and temperature measurements is required to resolve issues associated with the parameterization of gravity waves. These waves are central to a multitude of upper atmosphere phenomena and are most likely the dominant mechanism for energy and momentum flux transport from the lower atmosphere into the mesosphere, lower thermosphere, and ionosphere [Fritts and Alexander, 2003]. Gravity wave drag (GWD) parameterization remains a substantial weakness in even the latest generation of sophisticated numerical models of the atmosphere, this being associated with the difficulty in measuring these small-scale phenomena. While radar, passive optical, and satellite-borne instruments are crucial to the realization of the goals, modern lidar technologies have made particularly significant contributions to scientific studies, including gravity wave issues, that depend on high-resolution and high-accuracy measurements of winds and temperatures in the MLT (e.g., see a summary by Chu and Papen [2005] and recent papers by Gardner and Liu [2007, 2010], Yuan *et al.* [2008] and Liu and Gardner [2005]).

Recognizing lidar as one of the most powerful tools for observing the MLT, the European Space Agency (ESA) has supported two teams on a feasibility study of producing a spaceborne lidar for MLT studies [Sarris *et al.*, 2009]. Dr. Elsayed Talaat of Applied Physics Laboratory and Dr. Xinzhaoh Chu of the University of Colorado were involved in the feasibility studies, in collaboration with European and Canadian colleagues [Talaat *et al.*, 2008; Sarris *et al.*, 2010]. Through these two feasibility studies, it is clear that resonance-fluorescence Doppler lidar is the best choice for the proposed MLT science mission. This conclusion is derived from a distinct advantage of resonance fluorescence lidar over any other type of lidar. That is, the atomic absorption and fluorescence process provides the largest backscatter cross section in the MLT region by many orders of magnitude, leading to very high resolutions for temperature and wind measurements. This gift, given by our mother nature, makes the resonance Doppler lidar the only option for MLT science study with lidars.

Among the choices of resonance fluorescence lidars, we investigated two best candidate lidars: a Na Doppler lidar and an Fe/O₃ Doppler lidar. After extensive studies, we concluded that the Na-resonance Doppler lidar designed specifically for temperature and wind measurements is the best choice. This conclusion is based on the fact that Na has the largest product of effective backscatter cross section and constituent concentration which leads to the strongest signals for a given instrument power-aperture product. Furthermore, Na lidar has the most mature technology heritage and readiness because of extensive groundbased and airborne observations dating back to 1969 [Bowman *et al.*, 1969]. These advantages make the proposed Na Doppler lidar mission highly attainable and ready for flight within a short period. This mission will enable the science community to address the critical gravity wave and atmospheric modeling issues in a timely fashion. At the same time we also concluded that the second option of developing an alexandrite-laser-based Fe and O₃ lidar, operating at 248 and 372 nm, will represent a good mission once the diode-pumped alexandrite laser technology is further developed and reaches space quality, or should other UV laser sources emerge, as a mid- to long-term plan.

This white paper is built upon the feasibility studies and a proposal put forward to ESA's call for Earth Explorer Opportunity Missions EE-8, but highlights the future this platform might allow.

2. Scientific Motivation

Even with new advances from remote-sensing measurements from spacecraft missions dedicated to mesosphere and lower thermosphere studies, such as SME, UARS, TIMED and AIM, the MLT remains an under-sampled region with many remaining open questions. Being the "gateway" that connects Earth's environment and space, the MLT is a region of great importance in energy balance processes and a major link in vertical energy transfer. It is within these layers that great surges of energy meet: solar radiation and energetic particles contribute to downward energy transfer, whereas atmospheric tides and waves contribute to upward energy transfer from the stratosphere and below. Thus, this region is a boundary layer that determines the temperature and density characteristics of the surrounding layers.

Extensive modeling has shown, in response to increased greenhouse gas concentrations, temperatures in the mesosphere are expected to drop even further, in sharp contrast to the warming of the lower atmosphere. Thus, in a time of increased concern about global climate change, the fact that the mesosphere might act as the “canary in a coal mine”, being a sensitive indicator of global temperature change, makes its study an increasingly important matter. Finally, the continuous and ever-increasing presence of mankind in space, and the importance of the behavior of this region to multiple issues related to aerospace technology, such as orbital calculations, vehicle re-entry, space debris lifetime etc., make its extensive study a pressing concern for the space-science community.

In recent years, remote-sensing satellites have obtained the first global characterization of the basic structure of the MLT region in terms of large-scale temperature and wind climatologies. These measurements have also shown the high temporal variability of both the zonal mean state as well as large scale organized perturbations such as planetary waves and atmospheric tides. While we have a much richer picture of the structure and variability of the mesosphere than a decade ago, **what is missing is a fundamental characterization of how the basic state is established and maintained.** The underlying structure of the mesosphere is driven in large part by dynamical processes that produce major departures from radiative equilibrium. By far, the largest dynamical effects are believed to be from small-scale gravity waves. The cumulative impact of gravity waves is believed to be the dominant mechanism for energy and momentum flux transport from the lower atmosphere into the mesosphere, lower thermosphere, and ionosphere. Model studies have shown for many years that the dynamical forcing by gravity waves plays a substantial role in controlling the mesosphere thermal structure. Recent lidar observations have shown that these waves also play a significant role in vertical transport, which has a major effect on the constituent structure of the mesopause region (see Figure 1). The cold polar summer mesopause is thought to result ultimately from momentum deposition by gravity waves from the lower atmosphere, which reverse the zonal wind and drives upwelling and adiabatic cooling at high latitudes [Garcia and Solomon, 1985]; gravity waves are also responsible for the semi-annual oscillations at lower latitudes, as well as for the higher eddy mixing that is estimated through modeling in the mesosphere.

However, the measurements that could validate GWD parameterizations are still extremely limited. Global coverage of gravity wave variances has become available from satellite measurements [e.g., Wu *et al.*, 2006, and references therein] but these to date have significant limitations. First, these techniques are generally limited to below 80 km, so that the bulk of the mesospheric gravity wave activity is not measured. Second, these measurements usually have poor resolution either vertically or horizontally and therefore see just a small portion of the gravity wave spectrum. Third, because of the poor spatial resolution, these measurements are unsuitable to derive the spectral properties of gravity waves, and analyses usually calculate the residual broadband variance. Fourth, these measurements are of temperature only and therefore the gravity wave momentum flux is a derived quantity. So most of these missions focus on the lower atmosphere, as opposed to the mesosphere, where the gravity waves break, the “surf region” essentially for these waves.

Apparently, these key scientific issues must be addressed with high-resolution temperature and wind measurements, which are readily achievable with lidar, as demonstrated by ground-based lidar measurements [e.g., Chu and Papen, 2005; Gardner and Liu, 2007]. While these lidars provide very high temporal and vertical resolutions to resolve gravity waves, the ground-based sites are geographically sparse and cannot provide representative global climatologies of gravity waves. It is clear that high-resolution lidar measurements must be made from movable platform in order to provide good constraints

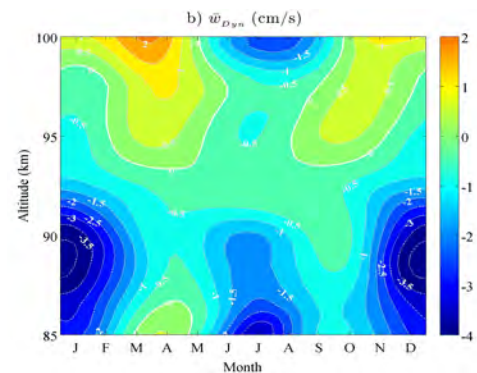


Figure 1. Seasonal variations of the gravity wave-induced, vertical dynamical transport velocity that was derived from wind, temperature and Na density measurements made with a Na Doppler lidar at Starfire Optical Range, NM [Gardner and Liu, 2010].

on gravity wave drag parameterization. Extensive airborne Na and Fe lidar observations have been conducted during the past 25 years which have helped the scientific community gain technical experience with deploying high-precision lidars on airborne platforms and analyzing the resulting data [e.g., *Gardner*, 1991; *Chu et al*, 2002]. Based on the feasibility studies we conducted with ESA, putting a resonance Na Doppler lidar into space will provide a solution to the issues mentioned above.

A zenith pointing Na Doppler lidar would provide measurements of vertical profiles of absolute temperature, absolute Na density and vertical wind perturbations. These data would enable the global characterization of the mesopause region temperature and Na density structure as well as gravity wave variances and spectra. In addition it would be possible to characterize wave-induced vertical transport of Na (see Figure 1) and other mesopause region species by computing the vertical Na flux profiles and perhaps the heat flux profiles [e.g. *Gardner and Liu*, 2007 and 2010]. Vertical constituent transport by gravity waves plays a crucial role in mesopause region chemistry which establishes the structure and seasonal variations of key species such as odd-oxygen. By restricting wind observations to the vertical wind perturbation profiles, rather than absolute winds, the nadir pointing accuracy and stability of the instrument can be relaxed. Ideally, the spaceborne Na Doppler lidar will have two beams pointing off-zenith in an orthogonal configuration and one beam pointing zenith so that temperature, vector wind and Na density can be measured simultaneously and globally with very high resolution and accuracy. This would further enhance the scientific benefits of the experiment, including the direct measurements of gravity wave momentum flux.

Besides active wave dynamics, the upper mesosphere is a region of abrupt chemical change, with very large mixing ratio gradients of species such as O, H and O₃. Ground-based measurements of these species in the upper mesosphere are not possible with sufficiently fine vertical resolution, and limb-scanning satellite observations are compromised by having to average over large horizontal distances. Thus, lidar observations provide the best possible test of models of vertical transport through the mesosphere and lower thermosphere. In addition, the space lidar proposed here will enable the meridional overturning circulation between about 50 and 90 km to be studied. This circulation appears to transport essentially all of the ablated meteoric metals to the winter polar vortex, followed by rapid descent into the lower stratosphere.

Another exciting application of the spaceborne lidar will be to explore the connections between polar mesospheric clouds (PMCs) and the meteoric metal layers. PMCs continue to be intensively studied by AIM and lidar [*Russell et al.*, 2009; *Chu et al.*, 2003, 2006; *Baumgarten et al.*, 2008] because they appear to be an indicator of climate change in the upper atmosphere. The complete removal of Fe atoms inside a PMC was observed by an Fe Boltzmann lidar [*Plane et al.*, 2004] at the South Pole. Similar heterogeneous removal was observed with Na and K [*Gardner et al.*, 2005; *Thayer and Pan*, 2006; *Raizada et al.*, 2007]. These studies showed that Fe, Na, and K were removed very efficiently by uptake of the metal onto the ice surface of the cloud particles. However, these were ground-based lidar observations at a fixed location, and thus were made in the Eulerian (fixed) frame of reference. The evolution of the PMC from nucleation to eventual sublimation, and the fate of the iron species after nucleation, could not be observed. These are important issues to be addressed with the spaceborne lidar.

3. Required Measurement Capabilities

In order to fulfill our stated science objectives of resolving gravity waves and Na density, we require high-resolution measurements in both space and time. The global distribution of gravity waves can be calculated from temperature and wind data observed with 2 km in vertical resolution and a horizontal footprint of 100 km. Significant improvement in the global climatology of Na density can be achieved with better than 0.5 km resolution in the vertical and 100 km in the horizontal. These resolutions can only be achieved with active measurements by lidars.

4. Space Lidar Principles and Technology Heritage

This mission concept is built upon the unique features of the MLT region, i.e., the neutral atomic sodium (Na), iron (Fe), potassium (K), etc. coming from meteor ablation [Plane, 2003]. These metal species provide excellent tracers of the atmospheric status like temperature and wind in the MLT region [Chu and Papen, 2005]. To illustrate the mission concept, an artist drawing (not to scale) is shown in Figure 2. Aboard a spacecraft orbiting the Earth in a noon/midnight orbit, a nadir-pointing Na Doppler lidar emits laser pulses that are precisely tuned to the D₂ absorption line of neutral atomic Na. As the pulses propagate downward through the Earth's atmosphere, the laser photons excite resonance fluorescence from Na atoms in the mesosphere and lower thermosphere, Rayleigh scattering from background atmospheric molecules, and Mie scattering from aerosols and clouds. The backscattered photons travel back to the space Na Doppler lidar to be collected by an optical telescope. The Na absorption process experiences Doppler broadening and Doppler shift due to the thermal (temperature) and dynamical (wind) motions of atoms and background molecules in this region. As a result, by probing the Na atomic absorption spectrum through tuning the laser frequency to several different frequencies – a method derived from the field of high-precision laser spectroscopy, we can accurately measure the atmospheric temperature and line-of-sight (LOS) wind in the MLT region where the Na atoms are in thermal equilibrium with the ambient atmosphere [Chu and Papen, 2005; Anorl and She, 2003].

The proposed lidar concept is an all-solid-state three-frequency Na Doppler lidar that is based on sum-frequency-generation (SFG) with two pulsed Nd:YAG lasers. Such a concept derives from both a strong heritage and current state-of-the-art lidar technologies. It integrates the cutting-edge technologies of solid-state lasers, laser spectroscopy, optics, electro-optics, sensors, and electronics into a single system to produce a powerful and robust tool with unmatched measurement capabilities. To achieve the highest resolution, a three-frequency ratio technique that has been applied in groundbased Na, K, and Fe Doppler lidars [She and Yu, 1994; Friedman et al., 2003; Chu et al., 2008a] is chosen as the operational mode of the proposed space Na lidar. The key component of the entire lidar is the generation of laser pulses that are in resonance with the Na D₂ line at 589.15826 nm. Our solution is to generate 589 nm laser with two diode-laser-pumped, pulsed Nd:YAG lasers operating at 1064.591 and 1319.250 nm mixed through a SFG process. This is possible because of the

coincidence of laser wavelengths: $\frac{1}{589.159\text{nm}} = \frac{1}{1064.591\text{nm}} + \frac{1}{1319.250\text{nm}}$. Such an approach has been successfully applied in a ground-based Na resonance Doppler lidar by Kawahara et al. [2008]. An earlier version of this Japanese lidar was successfully operated at Syowa Station in Antarctica for two years [Kawahara et al., 2003].

An overall end-to-end design of the proposed lidar system has been worked out through the ESA proposal process by Dr. Xinzhaoh Chu's group in collaboration with Canadian and European colleagues [Chu et al., 2010b]. It contains three subsystems: the lidar transmitter, the lidar receiver, and the data acquisition & control system (Figure 3). A significant advantage that resonance Doppler lidar has over

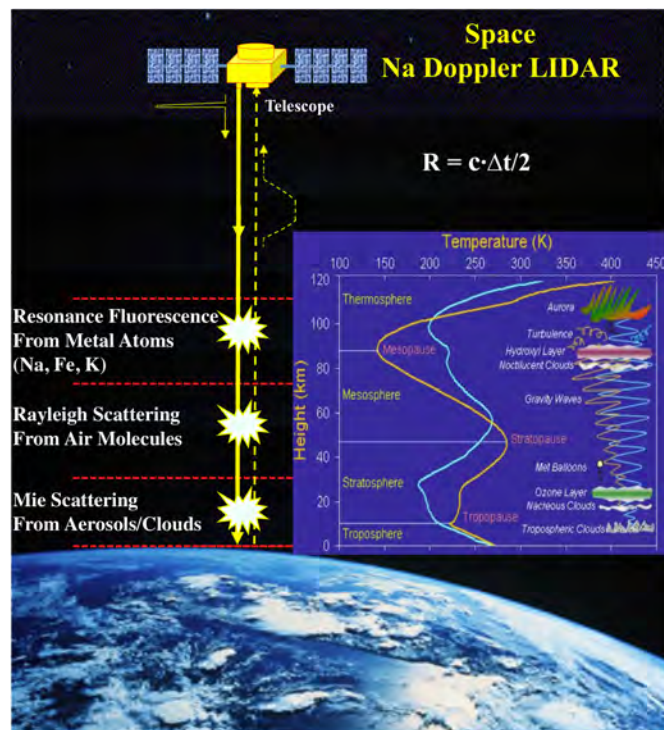


Figure 2. Illustration of Na lidar sensing of the Earth's atmosphere from space by Chu et al. [2010b].

other types of Doppler lidars is the fact that the spectral analyzer/discriminator required for spectral analysis of the Doppler broadened and shifted spectra (in order to derive temperature and LOS wind) is provided by mother nature in the atmosphere, i.e., the atomic absorption lines of Na, Fe, or K. Therefore, no spectral analyzer is needed in the receiver chain, so resonance fluorescence lidar employs a broadband receiver to collect all return photons. This provides further advantages for resonance lidar, as high-resolution spectral analysis in the receiver system greatly reduces throughput. The result is high signal-to-noise ratio (SNR) for MLT measurements.

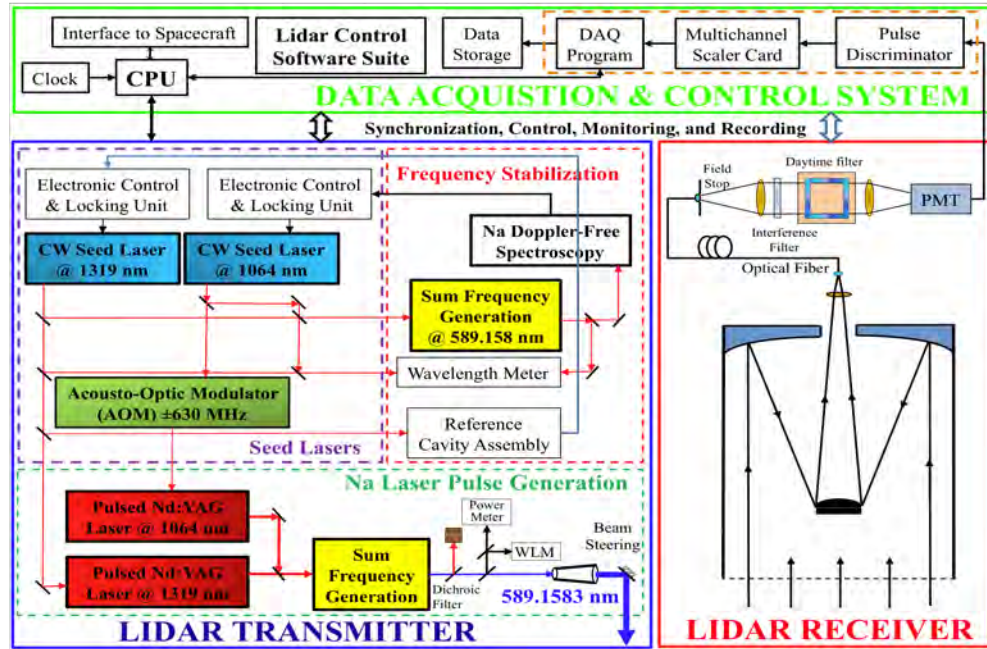


Figure 3. Architecture of the proposed spaceborne Na Doppler lidar by *Chu et al.* [2010b].

Although the feasibility studies were sponsored by ESA, the expertise of resonance-fluorescence Doppler lidar largely resides in USA middle and upper atmosphere community. The strong lidar heritage was built upon the pioneering work by Dr. Chiao-Yao She and Dr. Chester S. Gardner in the last two decades [*She et al.*, 1990; *Gardner*, 1991], and now being continued and advanced by the younger generation. Besides Na Doppler lidar, four other significant efforts on lidar technology development are the 770-nm K temperature lidar by German IAP and US Arecibo Observatory [*von Zahn and Höffner*, 1996; *Friedman and Chu*, 2007], the 386-nm Fe Doppler lidar by Dr. Josef Höffner's group at the German IAP [*Lautenbach and Höffner*, 2004], and the 372-nm Fe-resonance Doppler lidar by Dr. Xinzhaio Chu's group at the University of Colorado [*Chu et al.*, 2008]. All four lidars are based on injection-seeded pulsed alexandrite ring lasers. Although the flash-lamp-pumped regime precludes their application in an orbital vehicle due to space, weight and power requirements, the injection-seeding technology for pulsed lasers and the laser frequency diagnostic and control technologies developed from these lidars are very valuable to the proposed instrument concept [*Höffner and Lautenbach*, 2009; *Chu et al.*, 2010a; *Chu and Huang*, 2010]. Particularly, in seeking the high-accuracy (bias-free) wind and temperature measurements with the Major Research Instrumentation (MRI) Fe/Rayleigh/Mie Doppler lidar, innovations were made in high-resolution spectroscopy, optical heterodyne detection, ramp-fire injection-seeding pulsed laser technologies, and absolute frequency calibration of pulsed lidar [*Chu et al.*, 2010a; *Chu and Huang*, 2010]. These technologies have been borrowed by the space lidar mission to help achieve robust lidar system. Our proposed mission architecture integrates state-of-the-art technologies currently available and being developed and derives the instrument concept with the highest volume of technology heritage, so making the mission highly feasible and ready for flight within a short period.

5. Expected Performance and Impacts

All the key technologies in the proposed Na Doppler lidar have been successfully demonstrated in ground-based lidar measurements. The main issue facing the space mission is to achieve the required lidar specifications while meeting all the space-readiness requirements, in particular, within the space-allowed power and weight limits. An example is given in Table 1 where our studies show that it is highly feasible to achieve the listed performance within the ESA Earth Explorer limits.

Table 1. Space Na Doppler Lidar Performance at 380-km Orbit and Nadir Pointing

	Temperature Error	Radial Wind Error	Temporal Resolution	<i>Vertical Resolution</i>
Na @ Night (84-101 km) ¹	$\pm 1.0 - 2.1$ K	$\pm 1.0 - 1.7$ m/s	30 sec	2 km
Na @ Day (84-101 km)	$\pm 2.8 - 8.4$ K	$\pm 2.8 - 8.5$ m/s	30 sec	2 km

¹ First error is estimated for Na layer peak ~ 92 km, and second error is estimated for Na layer at 84 and 101 km.

The performance in Table 1 was estimated with conservative laser power (2.5 W) and receiver size (92 cm in diameter). The resolutions can certainly be improved with powerful laser and larger telescope. Better spatial and temporal resolutions can also be achieved for different requirements of the measurement precisions for different scientific products. A possible path to take is to complete an airborne or balloon-borne test of the lidar, and then to transfer the gained expertise to a satellite platform.

The payoff of the space lidar will be enormous, not only will it bring revolutions to the upper atmosphere science, but also will it push the envelope of laser and other related technologies in the United States and in the world. These advanced technologies will build a firm foundation for upper atmosphere science breakthrough.

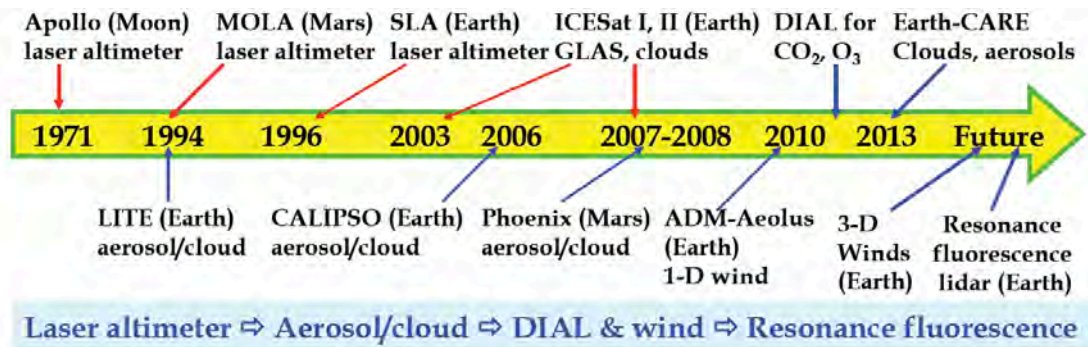


Figure 3. An illustration of the history of lidar in space

Tracking the footprints of spaceborne lidars (see Figure 3), it is clear that this technology is strongly linked and also makes significant contributions to the progress of laser, spectroscopy, and other related technologies. The space lidar goes from relatively simple laser altimeter through aerosol and cloud detection to more sophisticated Doppler and resonance fluorescence lidars. Certainly, the payoff increases dramatically as more advanced spectroscopy and technologies are utilized to sense the atmosphere. The proposed space laser is the best choice among all the current available options, including Raman shift of an alexandrite laser, frequency doubling of a fiber laser, and a Nd:YAG-based Raman laser. As solid-state laser technologies are progressing very fast, more options become available. A bright future is awaiting the space lidar for remote sensing of the middle and upper atmosphere.

References

- Arnold, K. S., and C.Y. She, Metal fluorescence lidar and the middle atmosphere, *Contemp. Phys.*, 44, 35-49, 2003.
- Baumgarten, G., J. Fiedler, F.-J. Lübken, F.-J. Lübken, Particle properties and water content of noctilucent clouds and their interannual variation, *Journal of Geophysical Research-Atmospheres*, 113(D6), 2008.
- Bowman, M. R., A. J. Gibson, and M. C. W. Sandford, "Atmospheric sodium measured by a tuned laser radar," *Nature* **221**, 456–457 (1969).
- Chu, X., W. Pan, G. Papen, C. S. Gardner, and J. Gelbwachs, "Fe Boltzmann temperature lidar: design, error analysis, and initial results at the North and South Poles", *Appl. Opt.*, 41, 4400-4410, 2002.
- Chu, X., G. S. Gardner, and R. G. Roble, Lidar studies of interannual, seasonal and diurnal variations of polar mesospheric clouds at the South Pole, *Journal of Geophysical Research*, 108 (D8), 8447, doi: 10.1029/2002JD002524, 2003.
- Chu, X., and G. Papen, Resonance fluorescence lidar for measurements of the middle and upper atmosphere, *Laser Remote Sensing*, ed. Fujii and Fukuchi, CRC Press, ISBN: 0-8247-4256-7, pp. 179-432, 2005.
- Chu, X., P. J. Espy, G. J. Nott, J. C. Dietrich, and C. S. Gardner, Polar mesospheric clouds observed by an iron Boltzmann lidar at Rothera (67.5°S, 68.0°W), Antarctica from 2002-2005: properties and implications, *Journal of Geophysical Research*, 111, D20213, doi: 10.1029/2006JD007086, 2006.
- Chu, X., W. Huang, J. S. Friedman, J. P. Thayer, MRI: Mobile Fe-resonance/Rayleigh/Mie Doppler lidar principle, design, and analysis, in *Proc. of 24th Int. Laser Radar Conf. (ILRC)*, Boulder, CO, June 23–27, 2008, pp. 801-804, 2008a.
- Chu, X., W. Huang, J. S. Friedman, and A. T. Brown, "CRRL/CTC: Doppler-free saturation-absorption and polarization spectroscopy for resonance fluorescence Doppler lidars," in *Proc. of 24th Int. Laser Radar Conf. (ILRC)*, Boulder, CO, June 23–27, 2008, pp. 809–812, 2008b.
- Chu, X., W. Huang, J. P. Thayer, Z. Wang, and J. A. Smith, Progress in MRI Fe-Resonance/Rayleigh/Mie Doppler Lidar, *Proceeding of the 25th International Laser Radar Conference*, pp. 947-950, St. Petersburg, Russia, 4-9 July 2010a.
- Chu, X, et al., Proposed Instrument Concept for Space Lidar Mission, in *ESA EE-8 proposal*, 2010b.
- Chu, X. and W. Huang, Fe Doppler-Free Spectroscopy and Optical Heterodyne Detection for Accurate Frequency Control of Fe-Resonance Doppler Lidar, *Proceeding of the 25th International Laser Radar Conference*, pp. 969-972, St. Petersburg, Russia, 4-9 July 2010.
- Friedman, J.S., C.A. Tepley, S. Raizada, Q.H. Zhou, J. Hedin, and R. Delgado, Potassium Doppler-resonance lidar for the study of the mesosphere and lower thermosphere at the Arecibo Observatory, *J. Atoms. Sol.-Terr. Phys.*, preprint, 2003.
- Friedman, J. S., and X. Chu, "Nocturnal temperature structure in the mesopause region over the Arecibo Observatory (18.35°N, 66.75°W): Seasonal variations", *Journal of Geophysical Research*, 112, D14107, doi:10.1029/2006JD008220, 2007.
- Fritts, D., and M.J. Alexander, Gravity wave dynamics and effects in the middle atmosphere, *Reviews of Geophysics*, 41(1), 1003, 64, 2003.
- Garcia, R.R. and S. Solomon, The Effect of Breaking Gravity Waves on the Dynamics and Chemical Composition of the Mesosphere and Lower Thermosphere, *Journal of Geophysical Research*, 90(D2), 3850-3868, 1985.
- Gardner, C. S. , Introduction to ALOHA-90: The Airborne Lidar and Observations of the Hawaiian Airglow Campaign, *Geophys. Res. Lett.*, *ALOHA-90 Special Issue*, vol. 18, no 7, pp. 1313-1316, July 1991.
- Gardner, C. S., J. M. C. Plane, W. Pan, T. Vondrak, B. J. Murray, and X. Chu, "Seasonal variations of the Na and Fe layers at the South Pole and their implications for the chemistry and general circulation of the polar mesosphere", *Journal of Geophysical Research*, 110, D10302, doi: 10.1029/2004JD005670, 2005.

- Gardner, C. S., and A. Z. Liu (2007), Seasonal variations of the vertical fluxes of heat and horizontal momentum in the mesopause region at Starfire Optical Range, New Mexico, *J. Geophys. Res.*, 112, D09113, doi:10.1029/2005JD006179.
- Gardner, C. S., and A. Z. Liu (2010), Wave-induced transport of atmospheric constituents and its effect on the mesospheric Na layer, *J. Geophys. Res.*, 115, D20302, doi:10.1029/2010JD014140.
- Höffner, J., and J. Lautenbach, Daylight measurements of mesopause temperature and vertical wind with the mobile scanning iron lidar. *Optics Letters*, Vol. 34, Issue 9, pp. 1351-1353, doi:10.1364/OL.34.001351, 2009.
- Kawahara, T. D., C. S. Gardner, and A. Nomura, Observed temperature structure of the atmosphere above Syowa Station, Antarctica (69°S, 39°E), *J. Geophys. Res.*, preprint, 2003.
- Kawahara, D. T., Nozawa, Satonori, et.al. A high power sodium temperature lidar based on injection seeded LD-pumped Nd:YAG lasers. 24th International Laser Radar Conference. 861-864, 2008.
- Lautenbach, J., and J. Höffner, Scanning iron temperature lidar for mesopause temperature observation, *Appl. Opt.*, 43, 4559-4563, 2004.
- Liu, A. Z., and C. S. Gardner (2005), Vertical heat and constituent transport in the mesopause region by dissipating gravity waves at Maui, Hawaii (20.7°N), and Starfire Optical Range, New Mexico (35°N), *J. Geophys. Res.*, 110, D09S13, doi:10.1029/2004JD004965.
- Plane, J. M. C., Atmospheric Chemistry of Meteoric Metals, *Chem. Rev.*, 103, 4963-4984, 2003.
- Plane, J. M. C., B. J. Murray, X. Chu, and C. S. Gardner, Removal of meteoric iron on polar mesospheric clouds, *Science*, 304, 426-428, 2004.
- Raizada, S., M. Rapp, F. J. Lubken, J. Hoffner, M. Zeche, J. M. C. Plane, Effect of ice particles on the mesospheric potassium layer at Spitsbergen (78 degrees N), *Journal of Geophysical Research-Atmospheres*, 112(D8), 2007.
- Russell, J. M. III et al., The Aeronomy of Ice in the Mesosphere (AIM) mission: Overview and early science results, *J. Atmos. Solar Terr. Phys.* (2009), 71, 289-299.
- Sarris, T. E., E. R. Talaat, A. Papayannis, X. Chu, P. Dietrich, J. Penson, and M. Koukouli, Project: "Study and Monitoring of Earth Mesosphere by Means of Lidar Techniques", ESA's General Studies Programme (GSP), Final Report (2009).
- Sarris, T. E., E. R. Talaat, A. Papayannis, X. Chu, M. Daly, P. Dietrich, J. Penson, V. Antakis, A. Vouldis, M. Koukouli, Ch. Kalaitzidis, E. Armandillo, Study Results for the Global Lidar Exploration of the Mesosphere, *Proceeding of the 25th International Laser Radar Conference*, pp.1355-1358, St. Petersburg, Russia, 4-9 July 2010.
- She, C. Y., H. Latifi, J. R. Yu, R. J. Alvarez II, R. E. Bills, and C. S. Gardner, Two-frequency lidar technique for mesospheric Na temperature measurements, *Geophys. Res. Lett.*, 17, 929-932, 1990.
- She, C. Y., and J. R. Yu, Simultaneous three-frequency Na lidar measurements of radial wind and temperature in the mesopause region, *Geophys. Res. Lett.*, 21, 1771-1774, 1994.
- E. R. Talaat, T. E. Sarris, A. Papayannis, E. Armandillo, X. Chu, M. Daly, P. Dietrich, and V. Antakis, "GLEME: Global Lidar Exploration of the Mesosphere," *Proc. 24th ILRC* 832-834 (2008).
- Thayer, J.P., Pan, W., 2006. Lidar observations of sodium density depletions in the presence of polar mesospheric clouds. *Journal of Atmospheric and Solar- Terrestrial Physics* 68, 85-92.
- von Zahn, U., and J. Hoffner, Mesopause temperature profiling by potassium lidar, *Geophys. Res. Lett.*, 23, 141-144, 1996.
- Wu, D. L., P. Preusse, S. D. Eckermann, J. H. Jiang, M. d. I. T. Juarez, L. Coy, and D. Y. Wang, Remote sounding of atmospheric gravity waves with satellite limb and nadir techniques, *Adv. Space Res.*, 37, 2269-2277, 2006.
- Yuan, T., C.-Y. She, D. A. Krueger, F. Sassi, R. Garcia, R. G. Roble, H.-L. Liu, and H. Schmidt (2008), Climatology of mesopause region temperature, zonal wind, and meridional wind over Fort Collins, Colorado (41°N, 105°W), and comparison with model simulations, *J. Geophys. Res.*, 113, D03105, doi:10.1029/2007JD008697.

Whole Atmosphere LIDAR for Whole Atmosphere Study
A Concept Paper for National Research Council Heliophysics Decadal Study

Xinzhao Chu and Jeff Thayer
University of Colorado at Boulder

After more than 50 years of development since laser was introduced to light detection and ranging (LIDAR), LIDAR has become a key instrument for middle and upper atmosphere as well as lower atmosphere study because of its high capabilities to measure wind, temperature, density, aerosol/cloud, and minor species in both day and night. Various types of lidars have been developed and deployed with each effectively measuring certain altitude range of the atmosphere, but so far none of them can profile the wind and temperature through the whole atmosphere. Neither does any existing instrument including radar and passive remote sensing. This situation significantly hampers the study of the whole atmosphere science, especially the wave coupling from lower to upper atmosphere. Absolute vertical wind is the key information characterizing the general circulations in the stratosphere and the mesosphere, but no instruments other than lidar can provide reasonable measurements. Even for lidar, to achieve cm/s accuracy and precision needed for the absolute vertical wind measurements is a big challenge. Fortunately, the technologies to achieve the absolute vertical wind measurements and to achieve the whole atmosphere temperature and wind profiling are emerging. Therefore, we propose a new concept – the whole atmosphere lidar that can profile wind and temperature through the whole atmosphere from ground to 120 km with superlative accuracy and resolution for advancing the whole atmosphere science.

Middle and upper atmosphere lidar began with cw searchlights in bistatic configuration for stratospheric aerosol, density, temperature, and pressure study from 1930s to 1950s [Elterman, 1954]. The invention of lasers in 1960s enabled the development of broadband Mie, DIAL, Rayleigh, and resonance fluorescence lidars with Q-switched ns pulsed lasers in monostatic configuration (see a paper collection by Grant et al., 1997). To date Rayleigh lidars have matured in detecting atmospheric density and temperature for study of thermal structure, gravity, tidal, and planetary waves in the aerosol-free region of 30-90 km [Keckhut, 1996]. Differential absorption lidars have been used to monitor stratospheric ozone for two decades [Keckhut et al., 2004]. Resonance fluorescence lidars have advanced significantly in the last 30 years capable of profiling temperature, wind, and density in middle atmosphere for study of thermal structure, composition, chemistry, and wave dynamics (see a summary by Chu and Papen [2005]). These lidars also function as Mie lidars to detect the stratospheric and mesospheric aerosols and clouds.

The whole atmosphere lidar concept was conceived during the development of a Major Research Instrumentation (MRI) mobile Fe-resonance/Rayleigh/Mie Doppler lidar [Chu et al., 2008, 2010]. Several revolutionary ideas and state-of-the-art laser and spectroscopy technologies were introduced to MRI Fe Doppler lidar, including utilizing the Fe-Doppler free spectroscopy and optical heterodyne detection (OHD) technologies to accurately analyze and control the frequency and spectrum of each lidar pulse [Chu and Huang, 2010]. It is worth noting that we were inspired by coherent Doppler lidar to develop the OHD technology for the MRI lidar. The development of the MRI Fe Doppler lidar represents an ambition to achieve the highest-accuracy (bias-free) resonance Doppler lidar for advancing the middle and upper atmosphere physics and

chemistry. For the future atmosphere science needs and challenges, a key piece of information that is still missing is the absolute vertical winds associated with the gravity-wave-driven residual meridional circulation in the mesosphere and associated with the planetary-wave-driven Brewer-Dobson circulation in the stratosphere. The successful development of the MRI Fe Doppler lidar and further technology advancement will enable the vertical wind measurements in the mesosphere and lower thermosphere (MLT) to ~ 1 cm/s accuracy and precision.

The whole atmosphere lidar is built upon such resonance Doppler lidar for making high-accuracy temperature and wind measurements in the MLT region but it is necessary to extend the wind and temperature measurement range to the stratosphere and troposphere, i.e., achieving the whole atmosphere lidar profiling of wind and temperature. It can be achieved by combining the resonance Doppler lidar technologies with several other lidar technologies, including Rayleigh/Mie Doppler lidar using edge-filter technologies [Gentry et al., 2000], vibrational-rotational Raman and pure rotational Raman lidar [Gerding et al., 2008], and the coherent Doppler lidar [Tucker et al., 2009]. Major advancements of coherent Doppler lidar and edge-filter direct detection Doppler lidar were made by NOAA and NASA teams. They can be integrated into upper atmosphere resonance Doppler lidar to form a whole atmosphere lidar. This, of course, requires relatively complicated transmitter and receiver development, but they all are within reach. A relatively large observatory is needed to implement such a whole atmosphere lidar system. But the modern laser, photonics and electronic technologies will make the system to be robust and reliable for routine operations.

The major justification for such a whole atmosphere lidar is to cover the entire atmosphere region from ground to 120 km so that one can track gravity waves from its source region through propagation to the region where they break and dissipate [Geller et al., 2007; Liu et al., 2010]. This will help remove the observational filtering effects on the study of gravity wave spectra and forcing. The whole atmosphere lidar will definitely be a key tool in understanding the whole atmosphere that is being studied extensively by sophisticated atmospheric models like WACCM and other general circulation models [Garcia et al., 2007].

Once the whole atmosphere lidar technology is achieved, it is envisioned to make observations at geophysically interesting locations like the gravity wave hot spots, equatorial region to observe the Brewer-Dobson circulation, and at high latitudes to observe the single-cell meridional circulations. It will be scientifically beneficial to place similar whole atmosphere lidars at several locations simultaneously to form a global network. Although it is still technically challenging, the whole atmosphere lidar is definitely within reach to address frontier issues in atmospheric and space sciences.

References

- Chu, X., and G. Papen, Resonance fluorescence lidar for measurements of the middle and upper atmosphere, Chapter 5 in *Laser Remote Sensing*, edited by T. Fujii and T. Fukuchi, published by CRC Press, ISBN: 0-8247-4256-7, pp. 179-432, 2005.
- Chu, X., W. Huang, J. S. Friedman, J. P. Thayer, MRI: Mobile Fe-resonance/Rayleigh/Mie Doppler lidar principle, design, and analysis, in Proc. of 24th Int. Laser Radar Conf. (ILRC), Boulder, CO, June 23–27, 2008, pp. 801-804, 2008.
- Chu, X., W. Huang, J. P. Thayer, Z. Wang, and J. A. Smith, Progress in MRI Fe-Resonance/Rayleigh/Mie Doppler Lidar, *Proceeding of the 25th International Laser Radar Conference*, pp. 947-950, St. Petersburg, Russia, 4-9 July 2010a.

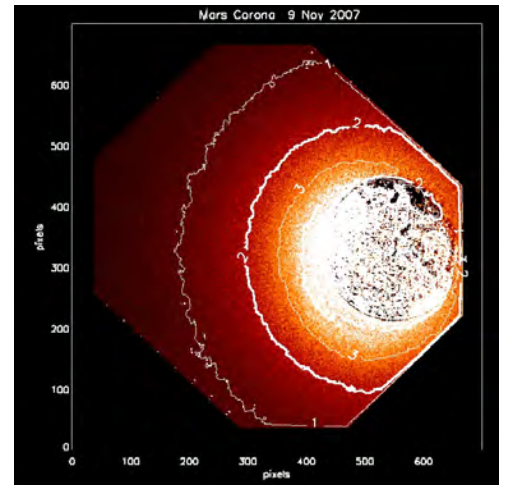
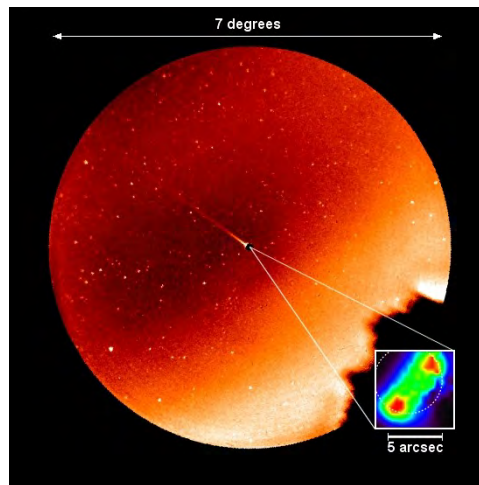
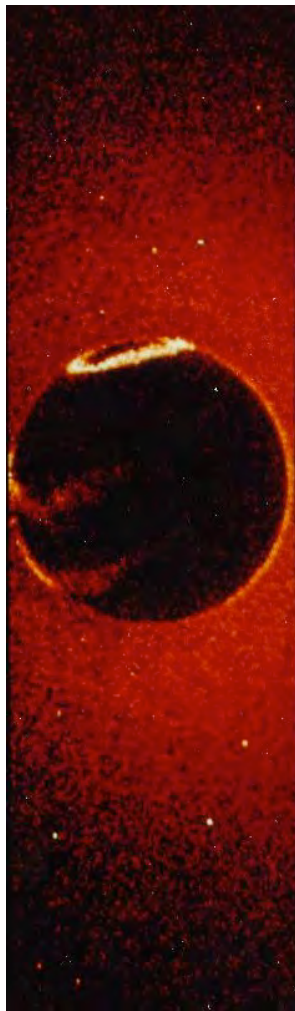
- Chu, X. and W. Huang, Fe Doppler-Free Spectroscopy and Optical Heterodyne Detection for Accurate Frequency Control of Fe-Resonance Doppler Lidar, *Proceeding of the 25th International Laser Radar Conference*, pp. 969-972, St. Petersburg, Russia, 4-9 July 2010.
- Elterman, L., Seasonal trends of temperature, density, and pressure to 67.6 km obtained with the searchlight probing technique, *JGR*, 59, 351-358, 1954.
- Garcia, R. R., D. R. Marsh, D. E. Kinnison, B. A. Boville, and F. Sassi (2007), Simulation of secular trends in the middle atmosphere, 1950–2003, *J. Geophys. Res.*, 112, D09301, doi:10.1029/2006JD007485.
- Geller, M.A., H.L., Liu, J.H., Richter, D. Wu., and F., Zhang (2006), Gravity Waves in Weather, Climate, & Atmospheric Chemistry: Issues & Challenges for the Community, Gravity Wave Retreat Report.
- Gentry, B. M., H. Chen, and S. X. Li, “Wind measurements with 355-nm molecular Doppler lidar,” *Opt. Lett.* **25**, 1231-1233 (2000).
- Gerding, M., J. Hoffner, J. Lautenbach, and M. Raute, Temperature structure of the atmosphere between 1 and 105 km at 54 N from resonance, Rayleigh, and Raman lidar, *Proceeding of 24th International Laser Radar Conference*, Boulder, CO, June 23–27, pp. 781-784, 2008.
- Grant, W. B., E. V. Browell, R. T. Menzies, K. Sassen, and C.-Y. She (editors), Selected papers on laser applications in remote sensing, *SPIE Milestone series, vol. MS 141*, ISBN 0-8194-2681-4, 1997.
- Keckhut, P., Rayleigh temperature lidar applications: Tools and methods, *J. Phys. IV France*, 139, 337-360, 2006.
- Keckhut, P., et al., Review of ozone and temperature lidar validations performed within the framework of the Network for the Detection of Stratospheric Change, *J. Environ. Monit.*, 6, 721, 2004.
- Tucker, S. C., W. A. Brewer, R. M. Banta, C. J. Senff, S. P. Sandberg, D. C. Law, A. N. Weickmann, and R. M. Hardesty, “Doppler lidar estimation of mixing height using turbulence, shear, and aerosol profiles,” *J. Atmos. Ocean. Tech.* 26, 673-688, 2009.

White Paper on Comparative Planetary Exospheres

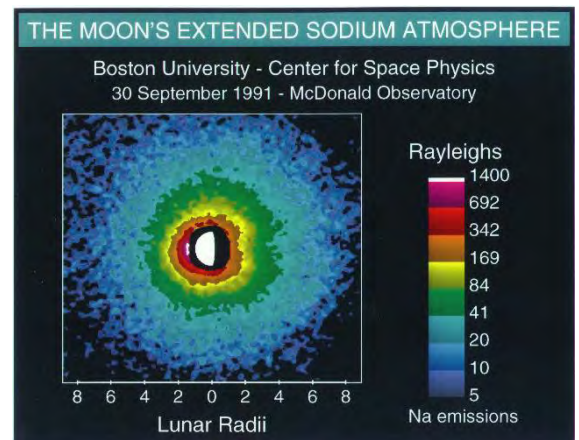
J. Clarke, C. Schmidt, J. Baumgardner, C. Carveth, M. Matta, M. Mendillo,
L. Moore, and P. Withers

Boston University
Center for Space Physics

submitted to the *Space Physics Decadal Survey* Fri. 11/12/2010



Figures: From left, images of the exospheres of the Earth (H Ly alpha emission), Mercury Na D line emission), Mars (H Ly alpha emission), and the Moon (Na D line emission).



Introduction and Justification:

Planetary exospheres form the boundary between the planetary atmosphere and the space environment. These regions encompass the interaction with the solar wind and are the source of atmospheric escape. Exospheres consist of neutral atoms and molecules that are collisionless on ballistic, satellite, and escaping trajectories. They are commonly greatly extended compared with the planet's radius (and decreasing "g"). They are subject to

complex interactions via charge exchange with solar wind protons, radiation pressure from solar emissions, and tidal effects from other gravitating bodies, while the source is determined by the properties of the underlying atmosphere or surface (e.g. Mercury).

Our understanding of the detailed physics of the Earth's exosphere can be greatly improved by studying other objects in the solar system where the conditions are very different in terms of the strength of gravity and the magnetic field, the intensity of solar radiation, and the size scale of the extended exospheric species.

This white paper proposes enhanced ground-based observations of the extended regions of other planets and satellites and a space-based mission that could greatly improve our knowledge at a modest cost. In addition to the Earth, the resulting improved understanding of exospheres can be applied to the cases of exo-planets, where super-Earths or Jovian planets exist much closer to stars than in our solar system. These systems are believed to undergo much more rapid escape and evolution of their atmospheres due to the large energy inputs, but little of the detailed work of modeling these systems has yet been carried out.

Measurement Objectives:

In the appendix we briefly tabulate the main properties of Mercury, Venus, Earth, Mars, Jupiter, Io, Europa, Titan, Uranus, , Triton, and Pluto, and the important emissions that can be observed to reveal the state of the exosphere of each object.

Recommendations / Mission Description:

A significant step toward improving the impact of comparative studies on the field of heliophysics would be achieved by the NASA Heliophysics R&A programs specifically soliciting/prioritizing such studies. Despite such studies falling within the scope of current R&A programs, experience has shown that proposals that go beyond the single planet Sun-Earth Connection are rarely rated highly.

Greatly expanded ground-based observations of planetary and satellite exospheres would improve our understanding of the detailed physics of their many interactions. A more comprehensive time series of measurements would allow us to study the dependence of exospheric densities and distributions with other parameters, i.e. solar wind strength, solar radiation, tidal forces, and the conditions in the lower atmosphere (the source). Ground-based observations can resolve exospheric properties with a resolution of ~ 1 arc sec, with some limits due to light scattering in the Earth's atmosphere (e.g. faint exospheric emissions next to the bright limb of the planet). We propose that 1 or more 1-meter class telescopes be outfitted with the needed instrumentation and dedicated to observations of planetary exospheres. Since the needed aperture is modest, existing telescopes no longer in active use for other observations could be modified to accomplish this. Since time coverage is an important parameter, it would be best if at least 2 telescopes with a wide range in longitude could be employed.

In addition, a small Earth-orbiting visible/UV telescope would permit higher resolution images of the key species, including simple atoms at UV wavelengths, with much lower scattered light than possible from the ground. This would provide measurements close to each planet's exobase, the source region of the atoms. In round numbers, visible imaging with a 6 inch aperture would provide ~ 0.5 arc sec imaging with a several arc min field of view of Na, K, and other species. Far-UV imaging with an ~ 30 cm aperture would provide similar resolution, field of view, and a higher sensitivity to the fainter emissions. Prior phase A studies of the proposed JMEX mission have shown that the resolution, pointing stability, data rates, etc. fit easily within SMEX mission parameters. A bore-sited two telescope system would in fact be significantly simpler, lower risk, and lower cost than the previously proposed JMEX mission. To keep this white paper short, we will not include all the specific data for such a mission, but we will be happy to provide more details if asked.

Summary:

In summary we propose comprehensive observations of planetary and satellite exospheres, including:

- Including comparative planetary exospheres in the stated priorities for proposals to NASA and NSF
- Supporting comprehensive ground-based observations of planetary exospheres, either with existing facilities or new telescope(s), including the development of the appropriate focal plane instrumentation
- Including in the next SMEX AO a priority for comparative planetary exospheres with a spacebased telescope payload.

Body	Species Observed	Scale Height	Extent Observed	Loss rates	Reference
Mercury	H (1216Å)	1330 km (day) 230 km (night)	0.4 R _M	<8.0e23 s ⁻¹ (Jeans)	Hunten et al., in <i>Mercury</i> 1988
	He (584Å)	330 km (day) 57 km (night)	1 R _M	3.0e22 s ⁻¹ (Photo-ionization)	Hunten et al., in <i>Mercury</i> 1988
	Na (5891Å & 5897Å)	150 km (supra-thermal)	1500 R _M	≤1.3e24 s ⁻¹ (Radiation Pressure (Varies))	Killen et al., <i>Planet Space Sci.</i> 1999; Schmidt et al., <i>Icarus</i> 2010
	Ca (4227Å)	0.3-0.5 R _M (supra-thermal)	3.5 R _M	?	Killen et al., <i>Icarus</i> 2005; Vervack et al., <i>Science</i> 2010
	Mg (285Å)	0.4-2.0 R _M (supra-thermal)	3.5 R _M	?	Killen et al., <i>Icarus</i> 2010; Vervack et al., <i>Science</i> 2010
Venus	H (1216Å)	1000 km (supra-thermal)	8 R _V	2.0e25 s ⁻¹ (Jeans)	Delva et al., <i>GRL</i> 2009; Galli et al., <i>JGR</i> 2008
	O (1304Å)	20 km (day)	0.8 R _V	2.0e25 s ⁻¹ (Charge Exchange)	Nagy et al., <i>Ann. Geophys.</i> 1990; Galli et al., <i>JGR</i> 2008
Earth	H (1216Å)	6.2 km	15 R _E	1.4e27 s ⁻¹ (Jeans)	Hunten, in <i>Atmospheres of the Solar System</i> , 2002
	O (1304Å)	10.9 km	?	?	Chamberlain, in <i>Phys. Aurora and Airglow</i> 1961
Mars	H (1216Å)	740 km (day) 230 km (night)	4.0 R _M	1.3e26 s ⁻¹ (Jeans)	Anderson and Hord, <i>JGR</i> 1971; Clarke et al., <i>DPS</i> 2009
	He (584Å)	185 km (day) 85 km (night)	?	7.0e23 s ⁻¹ (Jeans)	Krasnopolsky and Gladstone, <i>JGR</i> 1996.
	H ₂	370 km (day) 115 km (night)	?	3.3e24 s ⁻¹ (Jeans)	Krasnopolsky and Feldman, <i>Science</i> 2001.
Titan	H (1216Å)	900 km	0.7 R _T	6.95e26 s ⁻¹ (Jeans)	Hedelt et al., <i>Icarus</i> 2010
	H ₂	?	?	5.0e27 s ⁻¹ (Hydrodynamic)	Strobel, <i>Icarus</i> 2008
	CH ₄	?	?	2.0e27 s ⁻¹ (Hydrodynamic)	Strobel, <i>Icarus</i> 2008
Triton	H (1216Å)	0.7 R _T	?	2.0e28 s ⁻¹ (Jeans)	Hunten, in <i>Atmospheres of the Solar System</i> , 2002
	N	70 km	?	2.0e27 s ⁻¹ (Jeans)	Hunten, in <i>Atmospheres of the Solar System</i> , 2002
Pluto	N ₂	15-30 km	?	0.4-6.7e27 s ⁻¹ (Hydrodynamic)	Strobel, <i>Icarus</i> 2008

Constructively growing the sounding rocket program: A technology development line of sounding rocket launches

Executive Summary:

The rocket program has benefitted from recent funding increases that have helped to stabilize it. However, there are continuing, varied, and consistent programmatic and political reasons to continue to grow the program, as listed in the NRC Suborbital report. These include both long-term technological development needs, frequent access to space needs, and personnel development needs, both involving science students and NASA engineering staff. We propose that a competitive line of sounding rocket launches be added to the existing scientifically-competed program. This new line would be dedicated to technology development, enabling sounding rocket missions with long-term science goals but not necessarily immediate science closure.

The Problem of Limited Opportunities to Launch New Payloads

Since its inception, a major goal of the NASA sounding rocket program has been to facilitate low-cost rapid turnaround development of innovative new sensors, instruments, and mission concepts for space flight. Many of the nation's current generation of senior space scientists established their careers through instruments flown on sounding rockets while they were graduate students. Indeed, as noted in the 2010 report to the National Research Council by their "Committee on NASA's Suborbital Research Capabilities", the sounding rocket program has historically been "a vibrant and productive source of seed corn for NASA technology, science, and engineering talent."

Yet the section of this same report that summarizes the current status of this program is titled "Erosion of a National Asset". The general picture presented is that of a critical and once vibrant program that has in recent years been afflicted by steady decline. To quote from the report "According to the SRPO and NASA Headquarters, the average flight rate over the last 10 years was 17 core science and 5 reimbursable flights per year (22 total flights) with a total of 2 field campaigns in the entire period. In the years prior to that decade, the average flight rate was 28 flights per year and there were yearly remote campaigns." Considering that the reimbursable missions are not core science, the current rate of science missions is effectively 60% of what it was just 10 years ago.

The rocket program has benefitted from recent funding increases that have helped to stabilize it. However, there are continuing, varied, and consistent programmatic and political reasons to continue to grow the program, as listed in the NRC Suborbital report. These include both long-term technological development needs, frequent access to space needs, and personnel development needs, both involving science students and NASA engineering staff. The immediate scientific return of the sounding program is likely to scale roughly in proportion to the number of missions flown. However, *the decline noted above will likely induce a number of "multiplier" effects, such that the overall impact on the long term future of the nation's space science program is likely far more severe than*

a simple linear downscaling. These multiplier effects arise for a number of reasons, including:

- Fewer new sensors, instruments, and mission types can be developed;
- The relative scarcity of launches means that the success of each one is much more important, which stifles innovative or potentially higher risk missions;
- It is much harder to attract and train students to work on developing new flight hardware, because the chances are so low that they'd ever get to see their instruments flown;
- Research groups that currently are the main repositories of expertise in this area may close down as a result of their inability to maintain consistent viable funding.

Reinvigorating NASA's Suborbital Space Science Program

We propose that the most immediate way to reinvigorate suborbital space science is to target the “multiplier” effects mentioned above. We suggest that in addition to the science focus of the rocket program, which we consider vitally important, NASA should invest heavily in the short term on development of new and innovative mission technologies. Such a new technology development effort will provide the community with a new “toolbox” of tested and flight-hardened instruments and observation concepts to apply to their science questions, while at the same time injecting a cohort of new students, post-docs, and field-trained engineers into the field. A return to nearly 30 flights per year distributed roughly between core science and technology development constructively grow the program, and greatly facilitate scientific advancement and the development of new ideas, approaches, and people.

Three program thrusts are suggested to reinvigorate technology development. First, *a substantial expansion is suggested for the existing SR&T programs*, to enable development of new sensor concepts and technologies, and to test them under ground based laboratory conditions. These programs have not been substantially funded in recent years, but an increase now would facilitate the rapid development of instruments that could be used on sounding rockets. There is an existing program element specifically for this purpose, with established systems for proposal review, award management, and reporting of results.

In parallel with this, a new and separate initiative is suggested with the specific purpose to dramatically *increase the sub-orbital launch opportunities available for testing and demonstrating newly developed technologies under flight conditions*. The hallmarks of this proposed program would be fast turnaround, the fostering of innovation and its associated risk, and a culture that did not automatically penalize failures. Rather, failures that were understood and remediated would be seen as healthy and desirable outcomes of this flight testing program. Graduate and undergraduate student participation should be strongly encouraged, and internships by NASA engineering staff could be explored. This fostering of innovation would be applied to the efforts of both the science teams bringing new instruments forward for test flights, and to the contractors tasked with executing these missions.

This initiative could entail new mission development models at WFF/SRPO. One suggestion is that NASA emphasizes the development and dissemination of commonly used experiment components. Currently there is little sharing of designs among experimenters. If NASA assumed the role of maintaining a data base of commonly used experiment hardware that all teams could access, then some of the design time, and therefore development time and cost, could be reduced. A set of standards, such as are used in CubeSat development, could also be developed. These would be available to all experimenters and used to the extent that they facilitate the development process. These designs and standards are not forced on the experiments as it is crucial that the program be flexible and encourage innovation.

Another possible component of this initiative for increased launch opportunities would be for one of the existing launch ranges to host an annual campaign during which several complete sounding rockets would be tested, integrated and launched – with each vehicle carrying multiple payloads under test. These campaigns would be intensive “space camp” style activities during which multiple experiment teams would work in parallel both on their own payloads, as well as hopefully helping each other out to bring the completed vehicle packages to the launch rail. Students would be involved and would benefit greatly from such collaboration. Cost and development times would be reduced.

Various mechanisms can be imagined for how best to implement a line of competed launches dedicated to technology development. The most important element, though, would be a growth of the program sufficient to allow a significant increase in the number of competed launches, by inaugurating a new program dedicated to the long-term goals of technology development. Community discussion will provide input on how these initiatives should be applied. However it is implemented, such a program would provide benefits to the community through increased opportunities for student development, for engineering personnel development, and for science through high-level opportunities for technology development.

The third thrust is aimed at filling the logarithmic gap between sounding rockets (~\$2M) and explorers (~\$200M). A recent study conducted by the NASA Wallops Flight Facility (WFF), commissioned by the NASA astrophysics sounding rocket assessment team (ASRAT), showed how *advances in commercial launch technology could be used to establish an Orbital Sounding Rocket (OSR) program that would provide extended duration sounding rocket flights of 1 - 30 days*. The OSR program would be managed and operated by the NASA Wallops Flight Facility (WFF), whose low-cost, risk- tolerant culture is ideally suited to this effort. The OSR program is analogous to the highly successful Long Duration Balloon Program, in which the community uses balloons to perform breakthrough science. OSR payloads will be competitively chosen from successful sounding rocket experiments based on scientific and implementation merit. The goal is ~ 1 flight/year for 1000 lb payloads. This program would also provide routine opportunities for small (1-100lb) secondary payloads. The OSR cost, based on a preliminary WFF study is \$15M/flight, including payload modification, test and integration, launch vehicle procurement and processing, launch and mission operations.

A launch vehicle, for example the SpaceX Falcon 1, could deliver a previously proven sounding rocket payload to orbit for several days, which would provide much more than a factor of 1000 improvement in science return in comparison to 400 seconds typically available with today's sounding rockets. Such frequent access would allow the pursuit of high-risk high-value science, be and innovative approaches, and rekindle the competitive spirit of the early years.

References

Committee on NASA's Suborbital Research Capabilities; National Research Council (2010), "Revitalizing NASA's Suborbital Program: Advancing Science, Driving Innovation, and Developing a Workforce", The National Academies Press, ISBN-10: 0-309-15083-3 or online at http://www.nap.edu/catalog.php?record_id=12862.

Astro2010: The Astronomy and Astrophysics Decadal Survey, Position Papers, no. 36 (2009), "Reinvigorating the Astrophysics Sounding Rocket Program: Strategic Investment in the Future of Space Astronomy"
http://sites.nationalacademies.org/BPA/BPA_049492

Global scintillation prediction

Odile de la Beaujardiere, Cassandra Fesen, AFRL

We propose a systems-approach mission to predict scintillations from UHF to L-band at all latitudes.

The science questions to be addressed are:

1. Do irregularities need a “seed” to start?
2. When and why do irregularities grow after midnight, rather than at dusk?
3. Would irregularities form if there were no wind in the E- or the F-region?
4. Given a set of solar wind conditions, what is needed in terms of models and measurements to forecast the scintillation strength, time of onset and duration at a given place?
5. What is the role of the lower atmosphere and tides to control the onset and growth of scintillations?
6. On average, are scintillations evenly distributed across the polar cap and auroral oval?

A combination of space, ground observations, powerful neutral atmosphere/ionosphere/magnetosphere coupled models and careful analysis of the impact of scintillation on specific systems as well as the transition to operational prediction centers are required.

The aim is to improve our predictive capability to the point where scintillation can be predicted up to 24 hrs in advance with an accuracy of 67%, and up to 2 hrs in advance with 90 % accuracy.

The centerpiece of the mission is a complex satellite constellation with 8 satellites, 2 high (650 km) and 2 low (300 km) at high inclination, and a similar constellation at a low inclination. In addition, ground based sensors, coupled I-T-M models and system-impact specific algorithm are required.

LONG-TERM CHANGES IN THE IONOSPHERE/MAGNETOSPHERE SYSTEM AND RELIABLE PLATFORM FOR INNOVATIONS IN SPACE SENSORS

O. de la Beaujardiere, D. Ober --- AFRL

The recent deep solar minimum has taught us that the heliospheric system is not as predictable as previously thought, and that there is a need to investigate long term changes in the ionosphere/magnetosphere system. By the time the last DMSP satellite reenters, we will have had more than 40 yrs of continuous measurements of ionospheric and magnetospheric parameters.

Don't break the chain!

The DMSP s/c will be replaced in the same dawn-dusk orbit by the Defense Weather Satellite System (DWSS). Although the plan is to include a particle detector in the DWSS s/c, there are presently no other plans for space weather sensors.

This satellite should include a minimum set of sensors providing the same parameters as DMSP has: ion velocity, density, composition and temperature; electron and ion precipitating particles; magnetic field. These simple measurements are the backbone of numerous scientific investigations. They have become similar to geomagnetic indices. To stop them would be as devastating as to stop measuring the Dst index.

In parallel, these new s/c provide an opportunity for innovative instruments. There will offer ample opportunities for innovative instruments to be added to the set of operational instruments. By flying new instruments in the same satellite as more simple instruments, reliable sensor test and validation will be provided.

Examples of instruments to be tested could be both very small sensors to be included in nanosats or CubeSats as well as large and sophisticated instruments. For example, a Tiny Ionosphere Photometer (TIP) can be flown as well as a more complex GUVI.

In the longer-term, NWSS will facilitate the research to operation transition into the appropriate set of instruments that are needed for DoD and NOAA operational space weather models, while simultaneously providing the needed continuous long term data set, and providing a reliable test facility to increase instrument TRL. These s/c would continue to be the cradle for increasingly innovative instruments. These aims would be achieved through an interagency agreement between NASA and the AF. It is a win-win situation.

In conclusion, DWSS, the DMSP follow on, ought to accommodate instruments that provide the observations required for long-term trends in the magnetosphere ionosphere system as well as provide a reliable "home" for flight opportunity to test new instruments.

Submitted to: A Decadal Strategy for Solar and Space Physics (Heliophysics); Panels on Atmospheric-Ionospheric-Magnetosphere Interactions (AIM) and Solar Wind-Magnetosphere Interactions (SWM)

Title: On the Utility of Operational Satellite Data to Solar and Space Physics Research

Contributors: W.F. Denig (NOAA), W.K. Peterson (University of Colorado/LASP), G. Crowley (ASTRA), P.T. Newell (JHU/APL), M. Øieroset (University of California Berkeley/SSL), C.J. Farrugia (UNH), F.J. Rich (MIT/LL), E.F. Donovan (University of Calgary), R. Redmon (NOAA) J.L. Machol (Colorado State University/CIRA), J.V. Rodriguez (University of Colorado/CIRES), P.T.M. Loto'aniu (University of Colorado/CIRES) and B.A. Emery (HOA/NCAR)

Executive Summary – The heliophysics research community greatly benefits from the availability of operational satellite data from the U.S. Departments of Commerce, Defense and Energy. These agencies acquire and maintain numerous assets in space that collect space environmental data for a variety of operational purposes. While the needs of the heliophysics research community do not and should not drive operational requirements for space environmental data, the availability of such data continues to be leveraged for research purposes. To some degree, operational space environmental data provides a stand-alone capability for research. However we believe that the greatest value of operational space environmental data is in providing the context for multi-instrument studies of the interconnected sun-earth system. Additionally, the availability of extended historical records of the space environment is valuable to the scientific community since these records provide long-term statistical databases for space and terrestrial climatologies. While there is some degree of interaction between the operational and scientific communities, it is recommended that the Decadal Survey committee encourage further interactions to ensure that the potential research benefits of operational space environmental data are considered when operational decisions are made.

1. Introduction

This “white paper” is focused on the heliophysics research benefits of National operational satellite data. The United States (U.S.) Departments of Defense (DoD), Commerce (DOC), and Energy (DOE) have stated requirements to monitor and measure the near-earth space and solar environments for a variety of operational needs including space weather prediction, space situational awareness and nuclear compliance. For simplicity, within this paper we use the term “space environment” to include measurements and observations of the sun, interplanetary space and near-earth environment. Also, we note that these departments and other departments within the U.S. Government maintain operational ground facilities that contribute space environmental data for heliophysics research; namely, the DoD U.S. Air Force (USAF) Solar Electro-Optical Network (SEON) and Digital Ionospheric Sounding System (DISS), the DOC National Oceanic and Atmospheric Administration (NOAA) Continuously Operating Reference Stations (CORS) and the U.S. Department of the Interior (DOI) U.S. Geological Survey (USGS) magnetometer network. The contributions of these ground assets are, to some degree, addressed in other white papers (J.J. Love, USGS, private communication, 2010) and are not repeated here. Also not discussed are space environmental datasets available from commercial platforms such as Iridium Communications, Inc. for the Active Magnetosphere and Planetary Electrodynamics Response Experiment (AMPERE).

The next section discusses the available operational data sources and the benefits of operational data to heliophysics research. These benefits stem from the availability of long-term continuous data records using comprehensive sensor suites in multiple orbits. Section 3 discusses various factors affecting the performance of current operational assets and data availability issues. Section 4 discusses the perceived near-term (~10 year) future of operational space environmental monitoring. While in some cases the

current acquisition programs have been fully scoped, whether to the benefit or detriment of heliophysics research, in other cases the research community may be able to effectively influence the planning of future operational satellite missions. The intent is not to suggest that research needs should drive operational requirements or endanger homeland security but that a respectful consideration of leveraged research capabilities should be included in operational satellite mission planning and deployment.

2. The Value of Operational Satellite Data to Heliophysics Research

Space environmental data from operational satellites have been effectively used by the heliophysics research community for more than 30 years. Of particular note are the Defense Meteorological Satellite Program (DMSP) spacecraft, the NOAA Polar Operational Environmental Satellites (POES) and Geostationary Operational Environmental Satellites (GOES), and the various satellite systems which are the responsibility of the DOE. A simple search of abstracts contained in the Astrophysics Data System database finds almost 2,000 references to DMSP, 200 references to POES, and over 3,000 references to GOES. While not all of these references are specific to the space environment these numbers do indicate, nevertheless, that space environmental data from these operational satellite systems are widely used by the scientific community.

The value of operational satellite data for heliophysics research was first recognized in the early 1970's with the availability of white-light auroral imagery from the DMSP Operational Linescan System (OLS). The primary purpose of the OLS was to measure cloud patterns during the day and at night although the sensitivity of the OLS was sufficient to detect aurora in the absence of clouds and other stray light sources; i.e., daylight and moonlight. Pike and Whalen [1974] first used DMSP imagery to study the morphology of the auroral zone during geomagnetic substorms. Soon thereafter the DMSP program included an additional "special sensor" for measuring the characteristics of precipitating auroral charged particles which Meng [1976] then used to compare to the optical observations. Over time the complement of DMSP sensors relevant to heliophysics research increased to the current full suite which now includes the SSJ auroral particle detector, the SSIES ionospheric thermal plasma monitor, the SSM magnetometer, and the SSUSI and SSULI ultraviolet (UV) spectrographic imagers. Space environmental data from the DMSP satellites are available by request from the Air Force Research Laboratory and also freely available from various other data portals.

Also in the 1970's NOAA began to include space sensors on its geostationary satellites and its polar-orbiting platforms in low-earth orbit (LEO). The GOES-1 satellite, launched in 1975, was the operational follow-on to the National Atmospheric and Space Administration's (NASA) Synchronous Meteorological Satellite (SMS) which included a space environment monitor (SEM) package for measuring in-situ magnetic fields and energetic particles. The original GOES SEM also had an X-ray sensor used for measuring the solar background and to classify the magnitudes of solar flares. A solar x-ray imager (SXI) was added to the GOES space sensor complement in the early 2000's [Hill *et al*, 2005]. Starting in 1978 NOAA included a tailored SEM package (particles only) on the polar-orbiting Television InfraRed Observing Satellite (TIROS) spacecraft as the predecessor satellite system to the current POES. Historical data from the GOES and POES satellites are available through the NOAA National Geophysical Data Center (NGDC) and partly through the NASA virtual observatories and the Coupling, Energetics and Dynamics of Atmospheric Regions (CEDAR) database.

For more than 40 years, the Los Alamos National Laboratory (LANL) has provided operational sensors on various military spacecraft in geosynchronous orbit [Perez and Belian, 2010]. Since 1992, data from LANL sensors; i.e., the Charged Particle Analyzer (CPA), the Synchronous Orbit Particle Analyzer (SOPA) and Magnetospheric Plasma Analyzer (MPA), have provided detailed measurements of the geosynchronous

charged-particle environment that have been used in over 400 scientific publications [Bame *et al.*, 1992]. In the past these data were available through LANL although, more recently, the data have been greatly restricted and are currently not publicly releasable.

2.1. Space Environment Climatologies and Modeling

The evolving climatologies derived from space environmental data provide context and focus for future research missions. Inter-calibrations and comparisons of episodic data between operational and research grade instruments extends the reach of the research program and provides new insights to traditional space environmental data users.

The record of available space environmental data from operational satellites now extends well over thirty years and several satellite generations. During this time the orbital parameters for satellites in polar LEO at ~840 km and geostationary earth orbit (GEO) close to 36,000 km (~7 Earth Radii) have remained essentially the same. Also, while operational sensor designs have improved over this period the basic measurement principles have remained consistent and lend themselves quite naturally to long-term sensor inter-calibration [Hardy *et al.*, 2008; Emery *et al.*, 2008]. To this end, operational satellite data have been used for space environment climatologies [Kihn and Ridley, 2005; Gentile *et al.*, 2006] and related assimilative modeling of the space environment [Emery *et al.*, 1996; Lu *et al.*, 1994].

2.2. Providing a context for collaborative studies of the space environment

While the availability of operational datasets has intrinsic value, we believe that the greatest benefits are in collaborative studies of the space environment. Currently there are at least six operational satellites in the near-earth space environment (2+ DMSP; 1+ POES; 1 MetOp¹; 2+ GOES) providing continuous measurements of the space and solar environments. Historically this has basically been the case since the mid to late 70's. Also, and as an aside, we note that many sun-earth connection investigations are greatly facilitated today because of the ongoing availability of solar-wind and magnetic data near the L1 Lagrangian point from the NASA Advanced Composition Explorer (ACE) which many currently regard as a critical operational asset. Plans for future operational L1 measurements are discussed in Section 4.1.

Synoptic assessments of the near-earth space environment can be gleaned from operational datasets as was done for the 1989 great geomagnetic storm [Rich and Denig, 1992; Sojka *et al.*, 1994]. More focused collaborative efforts include using operational satellite data in combination with ground-based sensors [Singer *et al.*, 1985; Lockwood *et al.*, 1993; Denig *et al.*, 1993] and other scientific satellites belonging to NASA, the European Space Agency (ESA) and other nations [Maynard *et al.*, 1998]. Polar-orbiting operational satellites in LEO currently provide a global context for space physics in that each spacecraft provides a window into the larger magnetosphere during its 101-minute orbit. Indeed, many studies of the transfer of energy and momentum from the magnetosphere to the ionosphere depend critically on the availability of LEO data, particularly from DMSP, since these measurements provide information on the field-aligned current systems mediating the energy and momentum transfer [Farrugia *et al.*, 1998; Sandholt *et al.*, 2004; Sandholt and Farrugia, 2007]. Measurements available from operational assets in GEO also provide valuable contextual information as was the case on May 11th, 1999 when the solar wind dynamic pressure decreased to remarkably levels allowing the outer magnetosphere to become closely dipolar for an extended period of time [Farrugia *et al.* 2000].

¹ The MetOp series of satellites are owned and operated by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). Each MetOp satellite includes a SEM sensor provided by NOAA for operational space weather. The MetOp SEM is identical to the POES instrument and provides complimentary particle measurements in the midmorning orbit versus the POES early afternoon orbit.

Pre-planned collaborative studies can be conducted to measure background features in the space environment [Robinson *et al.*, 1987] although some of the more interesting investigations involve case studies where the operational satellite happens to be well positioned in the geospace; that is, located at the right place and at the right time. Two specific examples come to mind. During an early ESA/Cluster workshop it was found that the DMSP F13 satellite was nicely situated in the polar region on magnetic fields lines that closely mapped to the Cluster spacecraft locations while a noontime transient event was in progress [Lockwood *et al.*, 2001]. More recently, the availability of data from multiple GOES satellites in the vicinity of the Galaxy 15 spacecraft at the time of a debilitating satellite anomaly was critical in specifying the space environmental conditions that may have contributed to this satellite failure [Allen, 2010; see also Denig *et al.* 2010a]. The point of these two examples is that the consistency and coverage offered by operational satellites can provide an effective level of “space situational awareness” and can often be fortuitously positioned to contribute to particular case studies.

2.3. Intraspective studies of the space environment

Space environmental data available from operational satellite programs have been used intraspectively; that is, stand alone, to specify the space environment and to understand the underlying physics. DMSP measurements of precipitating charged particles from the SSJ sensor were initially used by Gussenhoven *et al.* [1983] to monitor the magnetic-activity related motion of the equatorial auroral boundary (see also Redmon *et al.* [2010]). In more detailed investigations, Hardy *et al.* [1985, 1989; see also Hardy *et al.*, (2008)] and Fuller-Rowell and Evans [1987] determined the statistical characteristics of auroral electron and ion precipitation using data collected over extended periods of time. Newell and Meng [1992] also used the SSJ data to identify magnetospheric source regions in high-latitude particle precipitation, a technique that has been extensively used in multi-sensor studies of the space environment. More recently, Newell *et al.* [2009] used 10 years of DMSP data to categorize auroral precipitation into diffuse ion or electron, monoenergetic and broadband, precipitation from dispersive Alfvén waves. Other examples of using DMSP data intraspectively for heliophysics research are the specification and understanding of high latitude convective flow patterns [Rich and Hairston, 1994] and the distribution of equatorial plasma bubbles [Huang *et al.*, 2001], to name just two.

As for other operational sources of space environmental data, the POES particle data have been used to derive an index for the total power deposited in the atmosphere from auroral precipitation [Foster *et al.*, 1986; Fuller-Rowell and Evans, 1987]. More recently, Rodger *et al.* [2010] used POES SEM data to examine radiation belt dynamics and the characteristics of energetic electron precipitation. Takahashi *et al.* [2010] used magnetic oscillations observed in GOES magnetometer data to determine solar cycle variations in the magnetospheric mass density of the outer magnetosphere. Similarly, LANL datasets have been used extensively used to understand the response of radiation belt particles to geomagnetic substorms [Reeves *et al.*, 1998].

3. Factors Affecting the Performance of Current Operational Space Assets

National operational satellite systems contributing to space physics research are the DMSP, the POES (MetOp), GOES and LANL sensors, as previously discussed. Operational requirements rightfully drive the architectural solutions for collecting operational space environmental data. However, careful planning with regards to the development, deployment and ground processing of operational space-sensor data and the public accessibility of operational environmental datasets have leveraged benefits for the space physics community which have, in turn, returned benefits to future operational missions through research-to-operations (R2O) efforts involving improved environmental modeling and ultimately improved next-generation instrument designs [Heelis and Hanson, 1998]. The availability of long-term, contiguous, high-quality datasets from operational assets has reaped substantial benefits for

heliophysics research. Unfortunately, there are examples where operational expediency has led to less than optimum benefits for the scientific community. Discussed herein are several examples from DMSP, POES, GOES and LANL, each of which illustrates a loss of leveraged capability.

The DMSP SSJ/5 sensor is an advanced tri-quadraspheric electrostatic analyzer (ESA) having multiple fields of view which are useful in determining the pitch-angle distribution of precipitating auroral particles. However as operationally deployed, the data from the 6 narrow field-of-view detectors (along a 105 degree arc segment) are internally summed together to produce a single measurement thereby losing all differential angle information. Fortunately, the future Space Environment Monitor – Next (SEM-N) operational sensor for the Defense Weather Satellite System (DWSS) will fully exploit the capabilities of the SSJ/5 and extend the particle measurements of ions (electrons) to 10 (1) MeV [Denig *et al.*, 2010b].

In a similar vein, the SEM-2 detectors on the operational POES and MetOp satellites perform some on-board processing which decimates the sensor outputs into a reduced set of “moments” used to calculate the differential energy flux of precipitating charged particles. Again, while the data from the SEM-2 is useful to the space physics research community [Emery *et al.*, 2008] the potential for this sensor to provide additional scientific information is limited by decisions reached for operational expediency. Also within NOAA, particle and field data from the GOES I-M series of spacecraft, specifically GOES 8 thru 12, were processed directly from raw telemetry into 1- and 5-minute averages of “corrected” flux values. Unprocessed, full resolution, sensor data are not readily available. Thus, the most accessible data products available to the scientific community are highly processed products that, notwithstanding their great value, have inherent processing assumptions and limitations.

With regard to public accessibility of NOAA environment data, the NOAA Administrative Order (NAO) 212-15 specifies policies for acquiring, integrating, managing, disseminating and archiving environmental and geospatial data and information. NOAA’s general policy is to provide unrestricted access to space environmental data collected by the POES and GOES spacecraft. Conversely, the USAF has no stated mandate for ensuring public access to historical space environmental data collected by the DMSP. The most accessible sources of DMSP data are through non-binding organizational commitments, by, for example, NGDC, the NASA virtual observatories and various universities. The situation with regard to LANL satellite data is on even more uncertain grounds in that recent changes in classification has greatly restricted public accessibility to these valuable historical datasets.

Again, the intent here is not to suggest that the research needs of the space physics community supersede or even compete with operational requirements. Rather, is it a recommendation that the research and operational communities remain respectively aware of each other’s respective needs and capabilities.

4. Future Capabilities of for Operational Space Environmental Data

The future of operational space weather is mixed. On the one hand, NOAA and the USAF are proceeding with plans to replace the aging ACE satellite which is vitally important to maintain the availability of upstream solar-wind data for heliophysics research. Similarly, NOAA is pursuing plans to field an operational replacement for the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC). For its part, the USAF may field an operational version of the Communications/Navigation Outage and Forecast System (C/NOFS) satellite [de la Beaujardiere and the C/NOFS Science Definition Team, 2004]. On the down side the future of operational space environmental sensing in polar LEO is diminished due to past changes in the National Polar-orbiting Operational Environmental Satellite (NPOESS) program and the resulting agency split into the separate NOAA Joint Polar Satellite System (JPSS) and the USAF Defense Weather Satellite System (DWSS).

4.1. Solar Wind Replacement (DSCOVR)

Seed funding for replacing the ACE satellite is included in the President's Fiscal Year 2011 budget request. NOAA's initial strategy for acquiring operational solar wind data is to acquire and refurbish NASA's Deep Space Climate Observatory (DSCOVR). This is a shared mission responsibility with the USAF expected to provide launch services to place DSCOVR near the earth-sun Lagrangian point (L1) located some 240 earth radii in the upstream solar wind. The anticipated launch date for the DSCOVR mission is 2014. Real-time data services to external users will be provided by the NOAA Space Weather Prediction Center (SWPC). Access to historical (>1 day) products will be through NGDC. Current resources also include funding to continue development of an operational solar coronagraph although the strategy for access to space is not yet determined. The follow-on strategy for acquiring operational solar wind data after the DSCOVR mission is still under consideration but may include a data-buy arrangement through a non-NOAA provider. Access to retrospective DSCOVR data will be through NGDC and most likely other data portals.

4.2. COSMIC-II

Also included in the President's FY11 budget request is limited funding for NOAA to proceed forward with the COSMIC-II program, also known as the Global Navigation Satellite System – Radio Occultation (GNSS-RO) mission. This program is a teaming arrangement between NOAA and Taiwan's National Space Organization (NSPO) with some added contributions by the USAF. While the greatest perceived benefit of COSMIC-II to NOAA is for atmospheric profiling, these data are also of value to operational space weather. As presently envisioned, the full up COSMIC-II constellation will consist of 12 satellites in LEO designed to collect some 8,000 daily soundings; i.e. occultations. The anticipated launch date for COSMIC II is 2014 in order to minimize the potential coverage gap between COSMIC and COSMIC-II.

4.3. GOES-R

For the most part, the GOES-R satellites will continue and extend NOAA's 30-year space environmental data record. However, the GOES-R sensors do offer some improvements over the current suite of space sensors. Specifically, the particle energy coverage for the GOES-R Space Environment In-Situ Suite (SEISS) will be extended down to ~30 eV, mostly to facilitate spacecraft charging assessments but also to provide additional details on the terrestrial radiation belts and ring-current particle populations. Heavy ions from He to the Fe group in the energy range of 10 to 200 MeV/nucleon will be added to monitor the flux of cosmic rays which can damage sensitive spacecraft electronics. Also included are improvements to the current SXI which will increase the sample cadence up to 3 or 4 images per minute. Acquisition planning for the GOES-R operational sensors benefited in the past from NOAA-hosted joint workshops with the research and operational communities.

4.4. Defense Weather Satellite System (DWSS)

By the time that the NPOESS program was canceled in February 2010, the complement of space environmental sensors originally envisioned for this program had already been severely compromised. Prior to 2003 the NPOESS Space Environment Sensor Suite (SESS) was planned to be an improvement over the capabilities offered by the combined DMSP and POES systems, including measurements of charged particles over a broad energy range, the thermal (ionospheric) background plasma, local magnetic fields and UV imaging of the ionosphere and thermosphere, all in multiple orbit planes. Changes in the NPOESS program in 2006 reduced the capabilities of SESS to particle measurements only and in only one plane. As an aside, we note that the cost overruns that led to the NPOESS program restructuring and reduction in SESS capability were not caused by SESS although, in some regards, SESS took the brunt of the fallout. In the aftermath of the NPOESS cancellation, the USAF intends to include a Space Environment Monitor – Next (SEM-N) particle detector on the DWSS although given the

uncertainty in the program such future measurements are not assured. The first launch of DWSS is planned for 2018. Also, during this next decade LEO particle measurements will be available from the predecessor SEM sensor on the MetOp satellites and from the DMSP as this latter program reaches its end of life prior to 2020.

4.5. Joint Polar Satellite System (JPSS)

The NOAA go-it-alone strategy for acquiring operational environmental data, following the cancelation of the NPOESS program, is to field the JPSS. Given an aggressive 2014 planned launch date for JPSS the NOAA has decided to adopt a NPOESS Preparatory Program (NPP) “build-to-print” approach. This decision unfortunately precludes flying a SEM-N on the JPSS-1. The SEM-N remains an unfunded “objective” for JPSS-2.

5. Summary

Over the next 10 years some elements of the Nation’s operational space environment monitoring infrastructure will be enhanced while others will remain basically the same or be diminished. On the plus side, NOAA and USAF plan to field; 1) a future operational solar-wind sentinel at L1, 2) an operational radio-occultation mission, and, possibly, 3) a replacement for CNOFS. The continuation of the GOES mission is also favorable both from the standpoints of available operational space data and continuity of the historical data record. While the continuity of precipitating particle measurements from polar LEO satellites is also of great value to heliophysics research, the loss of future capability in measuring other aspects of the LEO environment will be sorely missed.

The basic points of this white paper are to; 1) identify and articulate the leveraged research benefits of operational space data and 2) encourage the Decadal Survey panels make recommendations for the space physics community to remain engaged as appropriate in the planning and deployment of operational missions. Over the next 10 years the operational community will maintain GOES continuity and, in addition, likely field an operational ACE replacement and a COSMIC follow-on. Some decrease in capability is expected for space environmental monitoring in LEO due to the NPOESS past restructuring and eventual cancelation although continuity of particle precipitation measurements will be maintained by the DWSS, MetOp and, to some degree, by DMSP. Attention should also be paid to the expect quality of the datasets, including ground processing, and accessibility of operational space environmental data from restricted and classified programs.

References:

- Allen, J. (2010), The Galaxy 15 Anomaly: Another Satellite in the Wrong Place at a Critical Time, *Space Weather*, 8, S06008, doi:10.1029/2010SW000588.
- Bame, S.J., D.J. McComas, M.F. Thomsen, B.L. Barraclough, R.C. Elphic, J.P. Glare, J.T. Gosling, J.C. Chavez, E.P. Evans and F.J. Wymer (1993), Magnetospheric Plasma Analyzer: Initial Three-Spacecraft Observations From Geosynchronous Orbit, *Rev. Sci. Instrum.*, 64, 1026-1033.
- de La Beaujardière, O. and the CNOFS Science Definition Team (2004), C/NOFS: a mission to forecast scintillations (2004), *J. Atmos. Solar-Terres. Phys.*, 66, 1573-1591, doi:10.1016/j.jastp.2004.07.030.
- Denig, W., W. Burke, N. Maynard, F. Rich, B. Jacobsen, P. Sandholt, A. Egeland, S. Leontjev and V. Vorobjev (1993), Ionospheric Signatures of Dayside Magnetopause Transients: A Case Study Using Satellite and Ground Measurements, *J. Geophys. Res.*, 98, 5969-5980.
- Denig, W., D. Cooke, and D. Furgeson (2010a), Spacecraft Charging and Mitigation, *Space Weather*, 8, S10007, doi:10.1029/2010SW000632.
- Denig, W.F., T. Sotirelis, V. Grano, R. Hamilton, K. Wolfram, C. Brann and J. Manley (2010b), Status of the Space Environment Monitor for NPOESS (SEM-N), Abstract 357, 6th Annual Symposium on Future Operational Environmental Satellite Systems-NPOESS and GOES-R, 90th American Meteorological Society (AMS) Annual Meeting, 17-21 January 2010, Atlanta, GA.
- Emery, B., et al. (1996), Assimilative mapping of ionospheric electrodynamics in the thermosphere-ionosphere general circulation model comparisons with global ionospheric and thermospheric observations during the GEM/SUNDIAL period of March 28–29, 1992, *J. Geophys. Res.*, 101, 26681-26696.
- Emery, B. A., V. Coumans, D. S. Evans, G. A. Germany, M. S. Greer, E. Holeman, K. Kadinsky-Cade, F. J. Rich and W. Xu (2008), Seasonal, Kp, solar wind, and solar flux variations in long-term single-pass satellite estimates of electron and ion auroral hemispheric power, *J. Geophys. Res.*, 113, A06311, doi:10.1029/2007JA012866.
- Farrugia, C., P. Sandholt, W. Denig and R. Torbert (1998), Observation of a correspondence between poleward moving auroral forms and stepped cusp ion precipitation, *J. Geophys. Res.*, 103(A5), 9309-9315.
- Farrugia, C.J., H.J. Singer, D. Evans, D. Berdichevsky, J.D. Scudder, K.W. Ogilvie, R.J. Fitzenreiter and C.T. Russell (2000), Response of the equatorial and polar magnetosphere to the very tenuous solar wind on May 11, 1999, *Geophys. Res. Lett.*, 27(23), 3773-3776.
- Foster, J.C., J.M. Holt, R.G. Musgrove and D.S. Evans (1986), Ionospheric convection associated with discrete levels of particle precipitation, *Geophys. Res. Lett.*, 13, 656-659.
- Fuller-Rowell, T.J. and D.S. Evans (1987), Height-integrated Pedersen and Hall conductivity patterns inferred from the TIROS/NOAA satellite data, *J. Geophys. Res.*, 92, 7607-7618.
- Gentile, L. C., W.J. Burke and F.J. Rich (2006), A climatology of equatorial plasma bubbles from DMSP 1989–2004, *Radio Sci.*, 41, RS5S21, doi:10.1029/2005RS003340.
- Gussenhoven, M.S., D.A. Hardy and N. Heinemann (1983), Systematics of the equatorward diffuse auroral boundary, *J. Geophys. Res.*, 88, 5692-5708.
- Hardy, D.A., M.S. Gussenhoven and E. Holeman (1985), A statistical model of auroral electron precipitation, *J. Geophys. Res.*, 90, 4229-4248.
- Hardy, D.A., M.S. Gussenhoven and D. Brautigam (1989), A statistical model of auroral ion precipitation, *J. Geophys. Res.*, 94, 370-392.

- Hardy, D.A., E.G. Holeman, W. J. Burke, L.C. Gentile and K.H. Bounar (2008), Probability distributions of electron precipitation at high magnetic latitudes, *J. Geophys. Res.*, **113**, A06305, doi:10.1029/2007JA012746.
- Heelis, R.A. and W.B. Hanson (1998), Measurements of Thermal Ion Drift Velocity and Temperature Using Planar Sensors, in "Measurement Techniques in Space Plasmas: Particles", edited by R.F. Pfaff, J.E. Borovsky, and D.T. Young, *Geophysical Monograph 102*, pp. 61-71, American Geophysical Union, Washington D.C., 61-71.
- Hill, S. M., et al. (2005), The NOAA GOES-12 Solar X-Ray Imager (SXI) 1. Instrument, Operations, and Data, *Solar Phys.*, **226**, 255-281, DOI: 10.1007/s11207-005-7416-x
- Huang, C., W. Burke, J. Machuzak, L. Gentile, and P. Sultan (2001), DMSP observations of equatorial plasma bubbles in the topside ionosphere near solar maximum, *J. Geophys. Res.*, **106**, 8131-8142.
- Kihn, E.A. and A.J. Ridley (2005), A statistical analysis of the assimilative mapping of ionospheric electrodynamics auroral specification, *J. Geophys. Res.*, **110**, A07305, doi:10.1029/2003JA010371.
- Lockwood, M., et al. (2001), Coordinated Cluster, ground-based instrumentation and low-altitude satellite observations of transient poleward-moving events in the ionosphere and in the tail lobe, *Annales Geophysicae*, **19**, 1589-1612.
- Lockwood, M., W.F. Denig, A.D. Farmer, V.N. Davda, S.W.H. Cowley and H. Lühr (1993), Ionospheric signatures of pulsed reconnection at the Earth's magnetopause, *Nature*, **361**, 424-427, 1993.
- Lu, G. et al. (1994), Interhemispheric asymmetry of the high-latitude ionospheric convection pattern, *J. Geophys. Res.*, **99**, 6491-6510.
- Maynard, N., W. Burke, D. Weimer, F. Mozer, J. Scudder, C. Russell, W. Peterson and R. Lepping (1998), Polar observations of convection with northward interplanetary magnetic field at dayside high latitudes, *J. Geophys. Res.*, **103**, 29-45.
- Meng, C.-I. (1976), Simultaneous Observations of Low-Energy Electron Precipitation and Optical Auroral Arcs in the Evening Sector by the DMSP 32 Satellite, *J. Geophys. Res.*, **81**, 2771-2785.
- Newell, P. T. and C.-I. Meng (1992), Mapping the dayside ionosphere to the magnetosphere according to particle precipitation characteristics, *Geophys. Res. Lett.*, **19**, 609-612.
- Newell, P.T., T. Sotirelis and S. Wing (2009), Diffuse, monoenergetic, and broadband aurora: The global precipitation budget, *J. Geophys. Res.*, **114**, A09207, doi:10.1029/2009JA014326.
- Perez, M.R. and R.D. Belian (2010), The VELA success story and lessons learned, American Astronomical Society, AAS Meeting 215, Paper 316.04, 215 Meeting of the AAS, 03-07 Jan 2010, Washington D.C., *Bulletin of the American Astronomical Society*, **42**, p.317.
- Pike, C., and J. Whalen (1974), Satellite Observations of Auroral Substorms, *J. Geophys. Res.*, **79**, 985-1000.
- Redmon, R., W.K. Peterson, L. Andersson, E.A. Kihn, W.F. Denig, M.R. Hairston and R. Coley (2010), Vertical Thermal O⁺ Flows at 850 km in Dynamic Auroral Boundary Coordinates, *J. Geophys. Res.*, doi:10.1029/2010JA015589, in press.
- Reeves, G., R. Friedel, R. Belian, M. Meier, M. Henderson, T. Onsager, H. Singer, D. Baker, X. Li and J. Blake (1998), The relativistic electron response at geosynchronous orbit during the January 1997 magnetic storm, *J. Geophys. Res.*, **103**, 17559-17570.
- Rich, F. and M. Hairston (1994), Large-Scale Convection Patterns Observed by DMSP, *J. Geophys. Res.*, **99**, 3827-3844.
- Rich, F.J. and W.F. Denig (1992), The major magnetic storm of March 13-14, 1989 and associated ionospheric effects, *Canadian Journal Physics*, **70**, 510-525.

- Robinson, R., R. Vondrak, K. Miller, T. Dabbs and D. Hardy (1987), On Calculating Ionospheric Conductances from the Flux and Energy of Precipitating Electrons, *J. Geophys. Res.*, **92**, 2565-2569.
- Rodger, C.J., M.A. Clilverd, J.C. Green and M.M. Lam (2010), Use of POES SEM-2 observations to examine radiation belt dynamics and energetic electron precipitation into the atmosphere, *J. Geophys. Res.*, **115**, A04202, doi:10.1029/2008JA014023.
- Sandholt, P.E. and C.J. Farrugia (2007), Poleward moving auroral forms (PMAFs) revisited: responses of aurorae, plasma convection and Birkeland currents in the pre- and postnoon sectors under positive and negative IMF B_y conditions, *Ann. Geophys.*, **25**, 1629-1652.
- Sandholt, P.E., C.J. Farrugia and W.F. Denig (2004), Dayside aurora and the role of IMF $|B_y/B_z|$: Detailed morphology and response to magnetopause reconnection, *Annales Geophys.*, **22**, 613-628.
- Singer, H.J., W.J. Hughes, C. Gelpi and B.G. Ledley (1985), Magnetic disturbances in the vicinity of synchronous orbit and the substorm current wedge: A case study, *J. Geophys. Res.*, **90**, 9583.
- Sojka, J.J., R.W. Schunk and W.F. Denig (1994), The ionospheric response to the sustained high geomagnetic activity during the March 1989 great storm, *J. Geophys. Res.*, **99**, 21341-21352.
- Takahashi, K., R. E. Denton and H. J. Singer (2010), Solar cycle variation of geosynchronous plasma mass density derived from the frequency of standing Alfvén waves, *J. Geophys. Res.*, **115**, A07207, doi:10.1029/2009JA015243.

The Great Geospace Observatory and Simultaneous Mission of Opportunity

Eric Donovan, University of Calgary, Canada

In ISTP, and the missions that followed, the geospace fleet targeted the “system-level” view. This involved *in situ* and remote sensing observations that tracked the input and output (solar wind and the aurora, and ionospheric convection), with *in situ* observations in the cusp/cleft, the magnetosheath and magnetopause, auroral acceleration region, and high and low-latitude CPS at a range of altitudes. We saw our first true measurements of gradients (Cluster), as well as the first mission dedicated to the imaging of the geospace system as a whole (IMAGE). The armada gave us observations upon which much of our quantitative understanding of geospace is based, and significantly advanced our understanding of the geo-effectiveness of most important processes. We were left, however, with a thirst for more detailed observations of processes that sowed the seeds for the missions to follow.

As ISTP-era wound down, our understanding of the physics of geospace had taken great strides forward. The basic concepts of convection, including magnetic reconnection and bursty bulk flows had been established as fact. We had a heuristic understanding of the topology of geospace, and the key features of large-scale phenomena such as geomagnetic storms and substorms had been clearly established. At that juncture, however, the list of unresolved questions was long (even growing as a consequence of better but not quite good enough observations). Many of the fundamental unanswered questions relate to the larger system. These questions focus on, for example, the origin of structure in geospace, the role of different types of aurora in system-level dynamics, the mechanisms by which the CPS stages plasma for injection into the inner magnetosphere, the system level effects of dayside, distant, and near-Earth reconnection, feedback within geospace, and space weather affects the atmosphere and climate.

Geospace missions that will or are expected to fly over the next 10 years include RBSP, MMS, Swarm, and ePoP. These add to the fleet of extended missions including Polar, THEMIS, Geotail, Cluster, and Double Star, as well as the GOES and LANL geosynchronous constellation. Additional relevant missions that may fly in the next decade include Japan’s ERG and Scope. These missions all target specific processes with innovative new observations. MMS for example, will be four satellites probing scales down to electron-inertial, targeted at studies of fundamental processes such as reconnection. Almost the entire fleet is comprised of missions that target *specific* physical processes with *in situ* observations.

We are evolving towards a (mission level) focus on processes that occur on the smaller scales. We will understand geospace plasmas better, but our understanding of geospace as a whole will not move forward in step. If we are not careful, we will not have the global vantage to address questions that arise under the system-level themes that are starting to come to the fore in our field. Examples of such themes are Integrative Aeronomy around which the CEDAR program is rallying, and the Climate Impacts of Space Radiation and Dynamic Geospace Coupling initiatives discussed in the most recent NASA (2009) roadmap.

It is interesting that, although the missions that agencies and communities are choosing to fly are increasingly focussed on direct *in situ* observations of geospace plasma physical processes, interest in observations of the aurora is seeing a dramatic increase. This is evident on multiple fronts, including increasing numbers of aurora-related sessions at AGU meetings, and increased subscription of those sessions, and the number of aurora-related white papers submitted to this panel. There is obvious increased pressure for global auroral imaging from space, evidenced by numerous proposals for such missions, such as Ravens, Ravens-Europe, KuaFu-B, CASTRO, and others.

It would be valuable to fly an auroral imager *as good as* Polar UVI, or a global geospace remote sensing mission similar to IMAGE, but the value of re-flying such packages is either in the service of other (e.g., the *in situ*) missions, or to capitalize on synergies imaging along side of each of the upcoming missions. These are good but not great motivations, because inevitably this type of endeavour ends up being viewed as providing “context” for more interesting contemporaneous largely *in situ* projects, and context is not very marketable. In order for a mission targeting geospace at the system level to happen, it needs to offer something truly new. What would it take to do something truly new?

Truly new geospace imaging capabilities are well within our grasp. In terms of individual measurements, such as global auroral or Energetic Neutral Atom (ENA) imaging, the best that has been accomplished to date pales in comparison with what could be accomplished if we could bring readily available technology to accomplishing the tasks. Considering global auroral imaging, for example, after nearly forty years the best we have accomplished is roughly 100 km and 1 minute space and time resolution, and with poor sensitivity, meaning less than desirable SNR. This has been exacerbated by satellites that have afforded only inaccurate pointing information and ill-characterized wobble. For ENA imaging, limitations on imager mass have led to low spatial and temporal resolution. With advances in detector and other technologies, as well as experience from past missions, doing significantly better for these and other global observations is possible. Simply increasing the mass allocation for ENA imagers, and flying them on platforms affording better stereoscopic viewing of the inner magnetosphere would yield better resolved higher cadence sequences of images of the time-evolving inner CPS and Ring Current. For global auroral imaging, utilizing advanced dielectric filters, larger CCDs, three-axis stabilized satellites, and a more mass would enable an order of magnitude improvement in spatial resolution, temporal resolution, and spectral isolation. The result would be a window into new physics occurring on smaller and shorter scales, and the first-ever true dayside auroral imaging. One can easily envision similar advances for X-ray, and EUV imaging, as well as fundamentally new techniques enabling direct imaging of the magnetosheath and cusp plasmas [see e.g., *Robertson et al.*, JGR, 2006].

Any mission that makes significant advances in any geospace imaging technique would move our science forward in dramatic ways. The fact is, however, that we could accomplish a true revolution in geospace sciences by choosing to fly a number of different state-of-the-art imagers simultaneously. If one thinks very big, one can imagine Super-IMAGE, with multiple three—axis stabilized satellites in strategically selected orbits affording long duration 24/7 optimal views of the target geospace domains. One foresees such a mission giving one minute or better cadence simultaneous global views of the magnetosheath and cusp dynamics, the response of the (inner) CPS and ring current, and the simultaneous evolution of the plasmasphere and low and high energy auroral distributions. This would be complimented by a solar wind monitor (either existing or a new spacecraft such as Kuafu A), so the Great Geospace Observatory (GGO) would track energy input to, flow through, and impact on geospace as a system and geospace physical processes (as monitored by the *in situ* missions). By itself, the GGO view would rival that of the Sun provided by the fleet of solar imaging satellites, giving us a vision of planet Earth that would move geospace science to new heights and excite the public about our field.

The strategic benefits of GGO would be significant, but the mission carried out in the standard way (meaning all of the required instruments on one or several spacecraft built and flown specifically for the mission) would cost a significant fraction of a billion dollars. This may be possible, but it is worthwhile to consider alternatives that might deliver as much or even more than could be done by one mission. The fact is, if we were to take a philosophical/intellectual

step at the agency level, we could carry out GGO with remarkably little in the way of financial resources. Since geospace imaging is remote sensing, it does not matter where the satellite platform is, but rather what can be seen from that location. Furthermore, the orbit that gives the best view of the aurora is not necessarily the orbit that gives the best view of the magnetosheath or ring current region. This might be seen as a negative – meaning that the “mission” would need not one (or two if one wants 24/7 capabilities which one almost certainly does) but actually many satellites, each of which would be very expensive. However, this is where a change in our way of operating at the agency level could facilitate GGO without breaking the bank.

The approach that would facilitate carrying out GGO with modest resources is to utilize Simultaneous Missions of Opportunity (SMOs). The idea would be to identify platforms that would be ideal for each type of imaging. This could be, for example, commercial geostationary communications satellites for ENA imaging and the proposed Canadian Polar Communications and Weather (PCW) satellites for auroral imaging. Agencies such as NASA, ESA, CSA, CNSA, CAS, and others could decide which component of GGO they wish to support, with for example, Canada, the US, and China leading the auroral imaging, ENA imaging, and new solar wind observation components, respectively. The cost to each program would be only the cost of the specific mission of opportunity (say <\$40M per agency), with the exception of the cost of a solar wind observer. Without question, such an approach could deliver the best-case GGO for less than \$200M (not counting the solar wind observations) spread across four or five nations/agencies. This is eminently doable, but there is, of course, a catch. Under present practices, this is all but impossible. The fact is that if, for example, a US consortium bids on a mission of opportunity in an Explorer class competition, and if it seeks to have that mission fit into a larger set of parallel missions of opportunity, then there is a problem: if the answer to a question resembling “*Does the project’s success depend on a program or programs external to NASA?*” is “yes”, then this is a significant risk, bringing with it obvious difficulties in risk-averse programs. However, if there was a multi-agency decision to pursue GGO, this enhanced risk could be seen for what it is... enhanced risk, but with the possibility of a remarkable payoff, and with only moderate funds ventured by any one of the contributing parties. This last point is important... if the program elements are chosen carefully, then each one would be valuable in its own right. Stereoscopic ENA imaging with good spatial and temporal resolution, soft X-ray imaging of the cusp and magnetopause, or the kind of auroral imaging envisaged for PCW would each be valuable in their own right. Failure of every element but one would leave a surviving project still worthy of a mission of opportunity slot.

I have submitted this white paper as someone obviously outside of the US system, but I think that the message is relevant to your discussion. This note is meant, in part, to bolster the message being put forward in numerous other submissions (e.g., GEMINI) that there needs to be a reassessment of the value of system-level observations. Moreover, I hope that the simple message presented here can find some purchase in the US discussion regarding the future of geospace observations. Could several agencies get together and agree to sponsor three or four missions of opportunity, each of which targets a specific system-level observation in an innovative way? Could these fly together? Could these, together, make a Great Geospace Observatory that provides a system-level view of near-Earth space that parallels the view of the Sun that we now take for granted?

A Crucial Space Weather Effect: Meteors and Meteoroids

Lars Dyrud, Sigrid Close, Diego Janches, Meers Oppenheim, John Plane, John Mathews, Julio Urbina, Mihaly Horanyi, Jorge Chau, Jonathan Fentzke, Mike Sulzer, Sixto A. González, Shikha Raizada

Introduction

Every day billions of meteoroids impact and disintegrate in the Earth's atmosphere. Current estimates for this global meteor flux vary from 20,000-300,000 tons per year and estimates for the average velocity range between 14 km/s to 55 km/s, depending on sizes, masses and composition distributions. These particles arrive both during intense showers and as a nearly constant rain of sporadic meteors. Understanding the interplanetary meteoroid environment is important for several fields of study from solar system evolution, upper atmospheric physics, planetary atmospheres and ionospheres, planetary geology, and most critically to accurately assess the risk that these particles present to manned and unmanned space flight. Yet, the basic properties of this global meteor flux, such as the particle mass and velocity distributions and their chemical composition remain poorly constrained. It is known that at least four satellites have been severely damaged by meteoroids and one was destroyed during a meteor shower in the last 20 years; the shuttle's windows have been hit by over 20 meteoroids with enough energy to require replacement; and the Hubble telescope has received more than 5000 meteoroid impacts. A more accurate picture of the distribution of particle orbits and masses would help not only spacecraft designers but also scientists who model the outer solar system and its evolution based on the material observed near Earth, Earth and planetary atmospheric scientists and heliosphere scientists.

The largest uncertainties surround the most frequent meteors, mostly small (sand grain and dust size) sporadic meteors. We believe much of the mystery surrounding the basic parameters of the interplanetary meteor flux exists for the following reasons: the barely understood sampling characteristics of the different meteor observation techniques, and the scarcity of ground and spaced based measurements of meteors and interplanetary dust, which are used to derive or constrain most models.

Meteor physics is important to a broad range of disciplines. Meteors affect upper atmospheric and ionospheric physics and chemistry because meteoric material seeds the upper atmosphere with metallic atoms, ions, and dust [McNeil *et al.*, 2002, Plane, 2003]. These particles affect conductivities, create metal layers near 100 km altitude, and provide the condensation nuclei for noctilucent clouds [von Zahn, 2002]. Klekociuk *et al.* [2005], and subsequent press releases have recently suggested that meteoritic dust effects cloud formation and may influence weather, and large dust clouds from bolide explosion

may initiate abrupt climate change. Meteor radars have been used for decades to monitor winds in the mesopause and lower thermosphere [Chang *et al.*, 1999] and more recent attempts have been made to monitor temperatures [Tsutsumi, 1994]. This information is crucial for the understanding of the effects of global stratospheric warming in the MLT region and also acts as coupling between different regions of the atmosphere. Characterizing the high-speed meteoroid population is essential to understanding the current distribution of mass in the solar system and, by inference, its evolution and origin. The highest-speed meteors generally arrive from outside the solar system and characterizing this intergalactic population is of interest to astronomers [Meisel *et al.*, 2002]. Meteor burst communication relies on the reflectivity of meteor trails, bouncing high bandwidth radio signals off them.

Observing Micro-meteoroids

For decades, meteor observations were typically made with photographic and TV cameras and specular meteor radars. Specular radars only detect reflection from the trails of ionization formed *perpendicular* to the radar beam [Ceplecha *et al.*, 1998]. The resulting meteor parameters deduced from these radar observations are sensitive to this condition.

The last decade has seen a resurgence in US meteor research largely due to the increased attention from the Leonid Meteor Shower of 1998, which prompted numerous, large, military and civilian atmospheric radars to train in on the storm radiant. As a result, two new types of radar meteor reflections have become known and increasingly used. These reflections are known as meteor head echoes and non-specular trails and are largely observed and studied with high power and large aperture (HPLA) radars designed for incoherent scattering sensing of the ionosphere, and example of which is shown in Figure 1. Because these observations produce such detailed spectral signatures, they have stimulated intense study, they last few years of which has focused on using these observations for deriving more complex parameters about meteoroids and the atmosphere they interact with [Dyrud *et al.*, 2005; Zhou *et al.*, 2001, Malhotra *et al.*, 2007, Close *et al.*, 2002; Janches *et al.*, 2000, 2009; Mathews *et al.*, 2001, Janches *et al.* 2008, Chau and Galindo, 2008, Dyrud *et al.* 2008, Oppenheim *et al.* 2003a,b, 2009]. These observations have also shown that these small dust and sand grain sized micrometeoroids are considerably more numerous than previously measured, and are in largely retrograde orbits, unlike their larger sized counterparts. Additionally, it is becoming clear that observing radar meteors at two or more frequencies in common volume yields much more information on the details—e.g., meteoroid fragmentation [Mathews *et al.*, 2010].

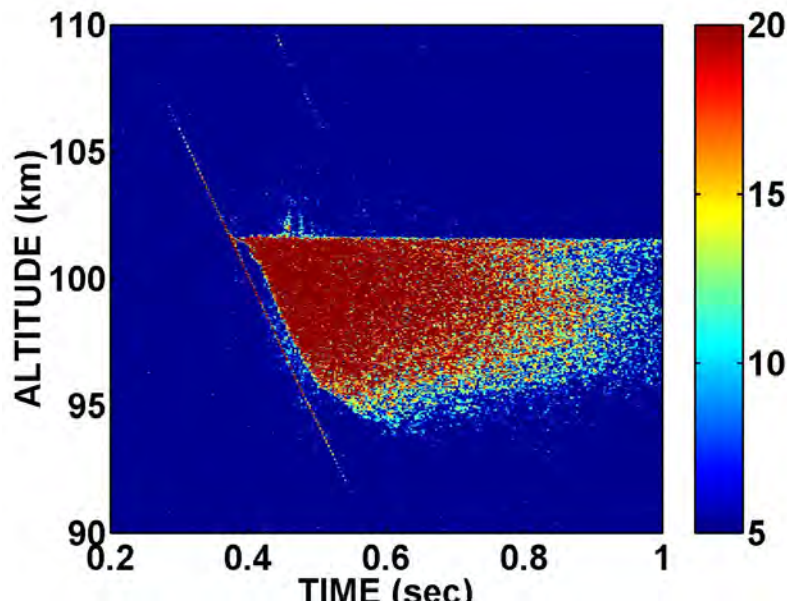


Figure 1 Altitude-time-intensity image of a head and subsequent non-specular echoes over extended range from ALTAIR VHF Radar. The diagonal line to the left is called a head echo, while the echoes spread in range and time to the right are the non-specular trail. Figure reproduced from [Close et al. (2002)]

Another driving force behind the resurgence in meteor research is the growing understanding of the impacts of meteoroids in planetary atmospheres. Starting with the Earth, meteoric ablation is the source of the global layers of metals atoms such as iron and sodium around 90 km altitude [Plane, 2003]. Metallic ions are the major constituents of sporadic E layers, which have important effects on radio communications. The metallic vapors also recondense to form nanometer size “smoke” particles, which provide condensation nuclei for noctilucent H₂O-ice clouds. Smoke also removes acidic gases and affects the formation of aerosols in the lower stratosphere, which impact on the ozone layer. The meteoroid flux probably explains the following recent observations in the atmospheres of other solar system bodies: sporadic layers on Mars and Venus; “noctilucent” CO₂-ice clouds around 80 km on Mars; and the detached haze layer on Titan around 500 km. We have just recently begun to understand the direct connection between the variability of the meteor flux with that of mesospheric layered phenomena [Gardner et al., 2010].

Recent work has demonstrated the importance of meteoric dust interaction with the solar wind flow [Mann et al., 2010]. Small nanometer scale particles are polarized and charged and interact with the solar wind and interplanetary medium, yet there is little modeling that exists on the overall influence both dust has on the solar wind, and the role that solar wind plays in the dust distribution.

Meteors as Space Weather-Space Craft Hazard

Although much emphasis has been placed on the space debris problem, which clearly dominates in low-Earth orbit (LEO), the meteoroid population travels at least twice the speed of debris and typically is a factor of 10 faster. In middle-Earth orbit (MEO) and geosynchronous-Earth orbit (GEO), the meteoroid and debris populations are at least comparable if not dominated by meteoroids. To determine an

approximate satellite impact rate, we can extrapolate flux measurements from ground-based radar. For example, the ALTAIR radar detects meteors formed by meteoroids with masses greater than $1\text{ }\mu\text{g}$. A typical measured sporadic peak meteoroid flux ($\sim 6\text{ AM}$ local time) is $192\text{ meteoroids/km}^2/\text{hr}$ for particles greater than $1\text{ }\mu\text{g}$. This detection rate, however, is highly dependent on the altitude of ablation and does not include small meteoroids that do not heat up enough to form meteors. If we instead look at the impact rate on LDEF, we find that microgram-sized particles impact at a rate of $360\text{ meteoroids/km}^2/\text{hr}$. Since the number of particles decreases (conservatively) as $1/m^2$ (where m is meteoroid mass), particles weighing 1 ng will impact a 1 m^2 spacecraft at least once per day. Furthermore, since meteoroids travel fast enough to ionize upon impact, they pose a risk of causing damage to satellites through electrical effects that depend more strongly on velocity than on mass.

The most common type of satellite damage is “mechanical,” which describes the penetration through the outer layer of a spacecraft. The electrical effects associated with an impacting meteoroid include electrostatic discharges (ESDs) and electromagnetic pulses (EMPs). An ESD occurs when there is a sudden discharge of accumulated electric charge on a satellite. This is usually caused when a satellite has a large buildup of charge that exceeds a voltage breakdown limit. ESDs are believed to occur frequently on spacecraft and are attributed to imbalances of surface currents in the space plasma charging process. In comparison, EMPs occur from the direct vaporization of an impacting meteoroid due to either radiation or charge transport mechanisms (Fig. 1). Although the link between meteoroid-induced EMP and satellite failure is not understood, the plasma production associated with a hypervelocity (i.e. meteoroid) impact has been studied for over 30 years using ground-based experiments. It is believed that both the meteoroid as well as a fraction of the target is vaporized and ionized, forming a plasma cloud instantaneously. This plasma cloud then expands into the surrounding vacuum with a separated charge caused by the different mobilities of the electrons and ions. The charge separation can cause an initial vehicle potential pulse, followed by plasma oscillations as the particles oscillate at their characteristic frequency. Since an asymmetry exists (such as a hemispherical plasma expanding away from a satellite), there will be an associated radiated power that may interfere with satellite operations or even cause catastrophic damage. Spacecraft lost due to electrical effects associated with meteoroid impact include Olympus in 1993, and SEDS and MSTI in 1994. In addition, the Landsat 5 satellite lost gyro stability during the peak of the Perseid shower in 2009 – similar to the Olympus satellite and suggestive of a meteoroid-induced electrical effect. Some of the spacecraft losses generically attributed to ESD may also be due to meteoroid impacts, since the catalyst for electrical anomalies typically remains undetermined.

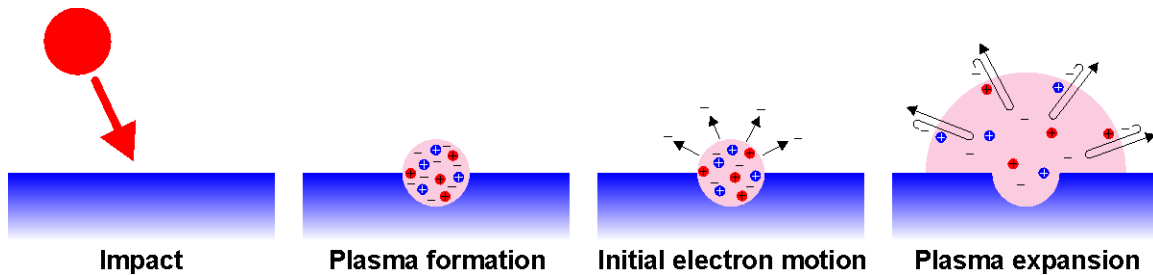


Figure 2. Depiction of plasma expansion process.

Recommendations

Because of the growing need from many within the Heliophysics and operational communities for an improved understanding of the interplanetary meteor and dust populations, we are strongly recommending funding for observations that support and constrain models of these populations. Further recognition of the interdisciplinary nature of this field of study via joint solicitations should also be prioritized in the coming decade.

1. In order to better understand meteor physics and to provide valuable observational constraints to meteoroid flux models, NSF should expand ground based utilities such as radars for observing meteors and meteoroids. This is necessary because each location and local time on Earth will preferentially view different sources of the meteoroid populations. Multi-frequency common-volume radars such as those available at Arecibo Observatory are particularly valuable. It is important to recognize that while these large radars are typically funded from NSF Geospace Facilities section (formerly Upper Atmospheric Facilities) the value of meteoroid observations extends well beyond and includes planetary science and heliophysics as well.
2. NASA should prioritize space-based methods for observing small meteors, meteoroids, and interplanetary dust in part to establish relationship between ground based observables and properties of interest such as meteoroid and dust composition and density.
3. Relevant agencies should establish cross-cutting solicitations and programs that facilitate the naturally multi-disciplinary nature of scientists with an interest and requirement to understand the dust and particle population and its effects. This will overcome programmatic hurdles that arise when the scientific motivation to understand the meteoroids/dust distribution possesses atmospheric, planetary, heliophysics, and space craft threat motivations.
4. NASA should facilitate and expand cooperation between basic researchers and space flight stake holders for improved characterization of the sporadic meteoroid and dust flux and threat to spacecraft well below the oft cited 1 cm particle size.
5. A relevant agency should establish, maintain, and distribute a meteoroid flux model to be used for scientific purposes by both Earth and planetary and heliophysics researchers that is

independent of engineering models, but with shared result to the benefit of improved space-craft threat models.

References

- D. W. R. McKinley. Meteor Science and Engineering, pages 1–309. McGraw Hill, New York, 1961.
- V. Bronshten. Physics of meteoric phenomena. d. Reidel Publishing Co., 1983.
- Z. Ceplecha, J. Borovic̃ka, W. Elford, D. Revelle, R. Hawkes, V. Porubc̃an, and M. S̃imek. Meteor phenomena and bodies. *Space Sci. Rev.*, 84:327–471, 1998.
- Oppenheim, M. M., Y. Dimant, and L. P. Dyrud. Large-scale simulations of 2-D fully kinetic Farley-Buneman turbulence, *Annales Geophysicae*, 26, 543–553, 2008a.
- P. Wiegert, J. Vaubaillon, and M. Campbell-Brown. A dynamical model of the sporadic meteoroid complex. *Icarus*, 201:295–310, May 2009.
- D. Janches, C. J. Heinselman, J. L. Chau, A. Chandran, and R. Woodman. Modeling the global micrometeor input function in the upper atmosphere observed by high power and large aperture radars. *J. Geophys. Res.*, 111(A10):7317, jul 2006.
- S. Close, M. Oppenheim, S. Hunt, and L. Dyrud. Scattering characteristics of high-resolution meteor head echoes detected at multiple frequencies. *J. Geophys. Res.*, 107(A10):1295, doi:10.1029/2002JA009253, 2002.
- L. Dyrud, E. Kudeki, and M. Oppenheim. Modeling long duration meteor trails. *Journal of Geophysical Research*, 112(A12), 2007
- M. M. Oppenheim, L. P. Dyrud, and A. F. vom Endt. Plasma instabilities in meteor trails: 2-D simulation studies. *Journal of Geophysical Research (Space Physics)*, 108:1064–+, 2003.
- J. Plane. A new time-resolved model for the mesospheric Na layer: constraints on the meteor input function. *Atmos. Chem. Phys.*, 4:39–69, 2004.
- T. Vondrak, J. Plane, S. Broadley, and D. Janches. A chemical model of meteoric ablation. *Atmos. Chem. Phys.*, 8:7015–7031, 2008
- J. Jones. Micrometeoroid engineering model-final report. Technical Report SEE/CR-2004-400, NASA, MSFC, June 2004.
- Mann, I., A. Czechowski, N. Meyer-Vernet, 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. Pantellini, 491-496, 2010

- M. Rapp, I. Strelnikova, and J. Gumbel. Meteoric smoke particles: Evidence from rocket and radar techniques. *Adv. Space Res.*, 40:809–817, 2007.
- J. Gumbel and L. Megner. Charged meteoric smoke as ice nuclei in the mesosphere: Part 1-A review of basic concepts. *J. Atmos. Sol. Terr. Phys.*, 71:1225–1235, Aug. 2009.
- D. M. Hunten, R. P. Turco, and O. B. Toon. Smoke and dust particles of meteoric origin in the mesosphere and stratosphere. *J. Atmos. Sci.*, 37:1342–1357, 1980.
- S. Love and D. E. Brownlee. A direct measurement of the terrestrial mass accretion rate of cosmic dust. *Science*, 262:550–553, October 1993.
- Z. Y. Fan, J. M. C. Plane, and J. Gumbel. On the global distribution of sporadic sodium layers. *Geophys. Res. Lett.*, 34:15808–+, 2007.
- J. Gumbel, Z. Y. Fan, T. Waldemarsson, J. Stegman, G. Witt, E. J. Llewellyn, C. She, and J. M. C. Plane. Retrieval of global mesospheric sodium densities from the Odin satellite. *Geophys. Res. Lett.*, 34:4813–+, 2007.
- M. Pätzold, S. Tellmann, B. H. Ausler, D. Hinson, R. Schaa, and G. L. Tyler. A Sporadic Third Layer in the Ionosphere of Mars. *Science*, 310:837–839, Nov. 2005.
- P. G. Brown, R. W. Whitaker, D. O. ReVelle, and E. Tagliaferri. Multi-station infrasonic observations of two large bolides: signal interpretation and implications for monitoring of atmospheric explosions. *Geophys. Res. Lett.*, 29(13):130000–1, 2002.
- B. W. Grime, T. J. Kane, S. C. Collins, M. C. Kelley, C. A. Kruschwitz, J. S. Friedman, and C. A. Tepley. Meteor trail advection and dispersion; preliminary lidar observations. *Geophys. Res. Lett.*, 26:675–678, mar 1999.
- Tsutsumi, M., Tsuda, T. and Nakamura, T., 1994. Temperature fluctuations near the mesopause inferred from meteor observations with the middle and upper atmosphere radar. *Radio Science* **29**, pp. 599–610.
- U. von Zahn, J. Hoffner, and W. J. McNeil. Meteor trails as observed by lidar. In E. Murad and I. P. Williams, editors, *Meteors in the Earth's Atmosphere*, pages 149–187. Cambridge University Press, 2002.

Commercial Access to Space for Scientific Discovery and Operations

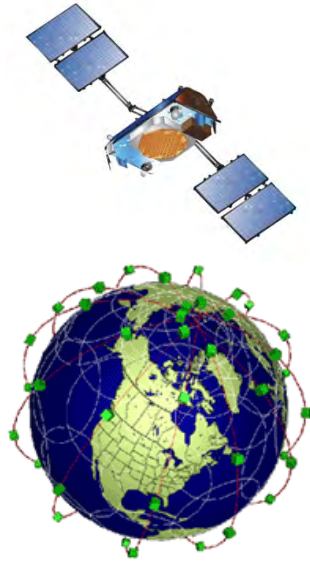
Leveraging broad new opportunities for cost-effective access to space for science

Lars Dyrud, H. Todd Smith, Larry Paxton, Chad Fish, Charles Sweson, Pontus Brandt, Gary Bust

Fewer than 500 scientific satellites and 500 people have been launched into space by government agencies. Commercial Space (non-DOD/NASA) have led the way in lowering costs and these developments promise to enable greater numbers of opportunities each year. The commercial use of space is anticipated to accelerate and will dramatically open new opportunities to address the frontiers of science. Addressing some of the most vexing basic science questions in magnetosphere, ionospheric and atmospheric science, concerns non-local non-linear coupled system dynamics. The observations required to achieve the greatest scientific advances will require arrays of scientific sensors distributed throughout the system gathering data simultaneously. This capability requires a change in the traditional approach in the implementation of scientific missions that are based on single or a few spacecraft and each typically cost \$200M - \$1B. While not all scientifically important orbits are available, sub-orbital, LEO, MEO and GEO provide excellent platforms for a range of in-situ and remote sensing sensors. Judicious and aggressive use of these opportunities for terrestrial magnetospheric, ionospheric, and upper atmospheric science could free up valuable resources for targeted missions addressing critical regions (as for instance MMS) and also for exploration missions to the Sun, L1, planetary magnetospheres and beyond.

Two broad types of commercial access to space are emerging manned/unmanned sub-orbital programs (Blue Origin, LLC; Dreamspace Group, Masten Space Systems, Inc.; Virgin Galactic LLC; XCOR Aerospace, Inc.), and hosted payloads on board commercial satellites (Iridium, Intelsat etc...). Each type of commercial space use has distinct scientific and cost saving potential. That these opportunities are real is demonstrated by the fact that the commercial sector is actively seeking commercial and federally supported hosted payloads and have accommodated this possibility in their space vehicle designs.

However there are several impediments that currently prevent scientists from taking advantage of these revolutionary opportunities. Pilot programs would be ideal avenues to advance and demonstrate the utility and benefit of removing these impediments and realizing the major cost, logistical, and time savings provided by commercial space flight opportunity. NASA's Commercial Reusable Suborbital Research Program (CRuSR) is a great example of such a program designed to help support and guide companies in meeting anticipated research needs however; this program is only currently modestly allocated and not yet funded.



Iridium NEXT constellation of 66+ LEO satellites



Examples of Commercial suborbital spacecraft (XCOR's Lynx Mark I and Virgin Galactic's SpaceShipTwo)

Benefits to scientific utilization of commercial access to space:

1. Dramatically lowered costs \$200,000 instead of \$2-3M (suborbital research), \$200M for constellation program instead of \$2-5B (Iridium)
2. Significantly more frequent access
3. Compressed timeline from scientific idea to space measurement
4. Reduced barriers to entry for scientists less familiar with space measurements
5. Address previously impossible scientific topics, i.e. constellations that make global measurements
6. Continue to address critical operational societal and scientific needs while lowering costs to the taxpayer.

Current impediments to realizing these benefits:

1. The mismatched schedules of commercial space and government funding agencies must be resolved. Commercial satellites are typically built and launched in 24-36 months. Funding agency programs are not set up to respond to these opportunities on such a short time scale.
2. Unresolved risk and liability issues between industry and Government.
Few or no sponsor mechanisms for providing support to projects utilizing the unique opportunities of commercial space.

To overcome these current obstacles and realize the significant benefits of utilization of commercial access to space we are making the following **recommendations**:

1. Firm government funding and support provided to the NASA CruSR program as well as adding additional funding opportunities for suborbital research, instrument development and education.
2. Initiate pilot programs to specifically fund hosted-payloads on commercial satellites to provide pathfinders for overcoming the above existing challenges.

We continue with a description of some existing efforts in both the hosted payload and sub-orbital areas that act to demonstrate some of the novel ideas within atmospheric and heliophysics that are being, and could be realized with future programs.

Program Name: AMPERE

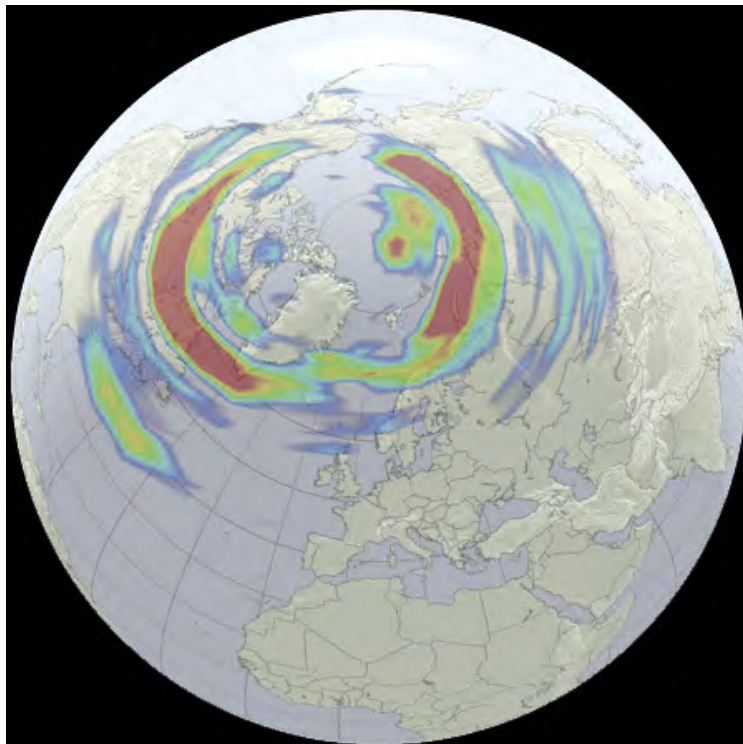
Program PI: Brian Anderson, JHU APL

Program Management: JHU APL

Funding Agency: NSF

The Active Magnetosphere Polar Electrodynamic Response Experiment (AMPERE) is a facility that provides global measurement of the field-aligned Birkeland electric currents that link the Earth's magnetosphere and ionosphere. The AMPERE facility utilizes the Iridium Communications Inc. constellation consisting of more than 70 satellites at 780 km altitude circular near-polar orbits evenly distributed among six equally spaced orbit planes. As part of the avionics, each satellite carries a vector magnetometer which is sensitive enough to detect these Birkeland currents. The AMPERE program acts to demonstrate that previously impossible science goals are attainable at modest cost with these proposed private public partner projects

The state and dynamics of the magnetosphere-ionosphere (M-I) system are intimately linked to the Birkeland electric currents that flow between the magnetosphere and ionosphere along magnetic lines of force and which mediate momentum and energy transport between high and low altitudes. The currents are the coupling nexus of the system and reflect the entire solar wind M-I interaction. By making continuous observations of the Birkeland currents with truly global and nearly uniform coverage



(see Figure 1) in both hemispheres independent of geomagnetic activity, AMPERE provides unprecedented capability to characterize the configuration and dynamics of the global field-aligned currents enabling major science advances across a broad range of M-I physics.

Figure 1 For 36 hours on Feb 14–15, 2010, AMPERE measured electric currents during a small magnetic storm. This view is from above the North Pole and slightly behind the Earth, with the Sun toward the top of the screen. Gray and blue colors represent weak currents, while greens, yellows and reds show

progressively stronger currents. This image shows the strong currents at the peak of the storm, intensifying as they move from the pole toward North America and Asia. The currents flow at heights of about 60–70 miles, but they can affect power grids, and major storms have caused widespread blackouts. Using AMPERE, scientists will soon be able to produce views like this nearly in real time – finally allowing us to see space weather as it happens.

Program Name: GEOScan: a Geosciences Facility from Space

Program PI: Lars Dyrud, JHU APL

Mission Management: JHU APL

Possible Funding Agency: NSF

Iridium Communications Inc. is launching a new generation of polar orbiting communication satellites in 2015-2017. Iridium has decided to place a hosted payload bay on each of the 66+ globally distributed satellites, which brings forth the possibility to host scientific sensors. Four primary factors make this an unprecedented opportunity for geoscience discovery, while holding the potential to affect a paradigm shift in the way we conduct science from space:

- 1) Truly global coverage provided by the constellation
- 2) Massively dense-space-based measurements enable revolutionary new techniques such as tomographic imaging.
- 3) Because Iridium Inc. is a telecommunications company, the logistical and cost barrier of transmitting massive amounts of data from 66+ satellites is REMOVED
- 4) Because we plan to build nearly 70 GEOScan pods we can take advantage of the cost savings of scale for science from space instead of the highly costly “one of a kind” methods of the past.

In order to take advantage of this new opportunity, we have devised a potential program named GEOScan, which consists of two complimentary main components, a community Hosted Sensor component where individual PI's provide and deliver unique sensors on one or more GEOSCAN pods, and 2) System Sensor science payloads, where two common sensors are placed on ALL GEOScan pods to conduct high impact and broad value system science and imaging investigations (examples include GPS radio occultation and atmosphere/ocean or limb imaging). Each 2-3U GEOScan pod that will be placed in all 66+ Iridium NEXT hosted payload bays and conform to the now defined Iridium SensorPod specification.

The GEOScan program will enable breakthrough science by uniformly sampling the Earth air and space environment in higher resolution with truly global coverage. The density of the measurements will enable new techniques such as tomographic imaging that allow viewing the Geo-environment in far greater detail than previously possible. These data will allow research into the untouched topics including global change that require both simultaneous and global sampling, and will provide an unparalleled source of data for both weather and space weather assimilative models. The success of the NSF funded AMPERE project that includes APL and Iridium acts as a model for GEOScan and provides inspiration for the possible science shifts in understanding that arise from global space based measurements.

Program Name: Storm Time Energy And Dynamics Explorer (STEADe)

Program PI: Dr. Charles Swenson, Utah State University

Mission Management: Space Dynamics Laboratory – Utah State University

Proposed Funding Agency: NASA

Program Description: Energy is deposited into high latitudes of Earth's upper atmosphere by a number of transporting sources, including gravity waves from below and solar originating electromagnetic, momentum (i.e., kinetic particle precipitation), and irradiation energy fluxes from above. The electromagnetic energy flux, due to continually changing solar-magnetospheric interactions, varies both temporally and spatially across the Earth's polar regions. It is a significant, and quite often the dominant, converging source in the ionosphere and thermosphere [e.g. [1]] as much of the polar region is dark throughout the year.

STEAD will globally measure electromagnetic energy flux into and the resultant heating of Earth's upper atmosphere at unprecedented temporal and spatial scales, thus providing a system response observation of sun-earth interaction. Specifically, STEAD will address the NASA Heliophysics Roadmap research focus area H2, "Understand changes in the Earth's magnetosphere, ionosphere, and upper atmosphere to enable specification, prediction, and mitigation of their effects", and the subsequent priority investigation of "How do the magnetosphere and the ionosphere/thermosphere systems interact with each other?"

Commercial Suborbital Opportunities

Numerous companies are in the process of developing suborbital spacecraft many of which are planned to start operational launches in 2011. Initially, companies developed a business plan that targets the space tourism industry with a cost of \$200,000 or less per seat. However, it has become apparent that the scientific community may provide a more stable and long term customer for these opportunities. Currently, most of these spacecraft are planned to attain approximately 100km above the Earth's surface for a period of time on the order of minutes in microgravity. The scientific community currently pays in excess of \$1.5M to conduct this type of research using traditional rockets. Therefore, commercial suborbital spacecraft offer the possibility of conducting the same research for a small fraction of the current cost, with the added benefit of potentially flying a scientist along with the experiment to ensure proper operation and safe return of the instruments.

Such low cost access to space offers new opportunities in our field in relation to scientific research, instrument development and STEM education. In-situ observations of the Earth's atmosphere are limited because this altitude is too high for balloons and too low for orbital spacecraft. Such observations (neutral particle composition and density, for example) conducted on a low cost and frequent basis will greatly increase our knowledge of the Earth's atmosphere. Additionally, such opportunities provide a low cost method of testing and validating new instrument concepts (beyond ground based laboratory testing) prior to flying them on heliospheric missions. This capability should more rapidly advance the technical state of the art. Finally, low cost commercial suborbital capabilities offers a much greater opportunity to directly involve university students in designing, developing and operating space-based scientific investigations and experiments. Such opportunities could dramatic improve future scientist education and retention.

These potential benefits from commercial suborbital spacecraft are rapidly advancing. Direct scientist interaction with emerging companies has already improved potential capabilities in regards to external payload bays for direct access to space and possible boosting of small payloads into orbit. Additionally, some of the companies are starting to design future spacecraft that will attain higher altitudes and thus expand the possibility for ionospheric (and above) research. This emerging technology is being developed and it is critical that the US continues to encourage this development through NASA supported and funded programs. Commercial spaceflight is advancing rapidly and if NASA does not support and lead this activity, there is a risk that NASA (and the US) could be left behind as the technology advances and becomes economically viable. This does not represent a loss of for NASA initiatives but more of an expansion to facilitate the safe and expedient exploitation of commercial access to space.

Geospace Climate Present and Future

A White Paper for the Decadal Survey of Solar and Space Physics

John T. Emmert

U.S. Naval Research Laboratory, Washington, DC

john.emmert@nrl.navy.mil

202-767-0467

Contributors and Signatories

Hugh G. Lewis	University of Southampton, UK
David E. Siskind	Naval Research Laboratory
Lucilla Alfonsi	Istituto Nazionale di Geofisica e Vulcanologia, Italy
Douglas P. Drob	Naval Research Laboratory
Ana G. Elias	Consejo Nacional de Investigaciones Cientificas y Tecnicas (CONICET), Argentina
John M. Holt	Massachusetts Institute of Technology
Jan Laštovička	Czech Academy of Sciences
Tao Li	University of Science and Technology of China
Richard J. Niciejewski	University of Michigan
Susan M. Nossal	University of Wisconsin-Madison
Craig A. Tepley	Arecibo Observatory

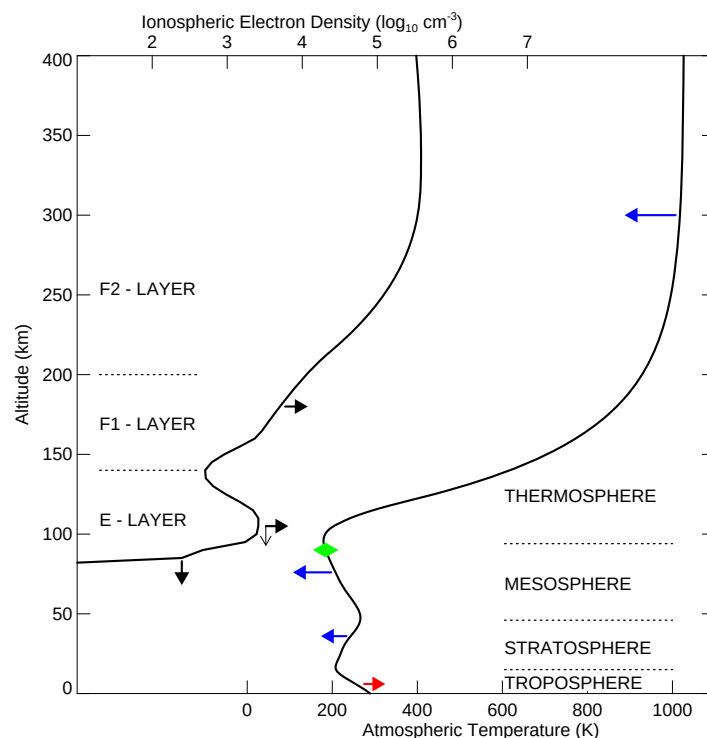


Figure 1. Qualitative summary of long-term changes that have been detected in ionospheric electron density (left) and atmospheric temperature (right). The directions of the observed changes, indicated by the arrows, are consistent with those predicted to occur as a result of increased CO_2 levels.

1. The Importance of Geospace Climate

It has long been known that Earth's thermosphere, ionosphere, and mesosphere respond strongly to variations in the Sun's photon, particle, and magnetic field output. More recently, it has become increasingly apparent that Earth's upper atmosphere is also intricately coupled to its underlying layers, all the way down to the Earth's surface. These interactions occur via a number of mechanisms, including radiative transport, chemical transport, propagation and selective filtering of waves of all types, and electrical fields and currents. **It is now clear that a meaningful understanding of upper atmospheric behavior – including both climate and weather – cannot be attained except by considering the entire atmosphere as a whole system that responds not only to direct solar influences, but also to internal variations and changes of both natural and anthropogenic origin.**

From a societal perspective, the climate of the upper atmosphere (here defined as the mesosphere, thermosphere, ionosphere, exosphere, and geocorona) is important for several applications, including satellite drag and radio signal propagation. Satellite drag is relevant to orbit and reentry prediction and to long-term mitigation of orbital debris. Societal use of the radio wave spectrum is fundamentally affected by the ionosphere, which alters the paths and properties of radio waves of all frequencies, including GPS signals.

The same gases that have been implicated in lower atmospheric climate change also play important roles in the energy balance of the upper atmosphere. CO₂ is the primary cooling agent of the upper mesosphere and thermosphere. Anthropogenic increases in CO₂ propagate into the upper atmosphere, and there are strong theoretical grounds and experimental evidence^{1,2} that these increases are causing the upper atmosphere to cool and contract. Changes in CH₄, H₂O, and O₃ may also be affecting upper atmospheric energy balance and structure³. Anthropogenic changes in CO₂, CH₄, and O₃ originate primarily from ground-based emissions (of CFCs in the case of O₃ changes), but upper atmospheric H₂O concentrations may be influenced by direct injection of rocket exhaust⁴ as well as by the background climate change⁵. NO is another major cooling agent; it is primarily produced by the interaction of the magnetosphere with the lower thermosphere and is transported down into the lower mesosphere and stratosphere, where it catalytically destroys O₃⁶. Although NO is produced naturally, theoretical simulations suggest that it will modulate the climate change induced by CO₂ increases⁷. Some of the changes predicted to occur as a result of greenhouse gas increases have been detected in historical upper atmospheric data, and a coherent picture is beginning to emerge^{2,8} (Figure 1).

There are also natural sources of geospace climate change. Long-term variations in the Sun's photon, particle, and magnetic field output will directly affect the structure and climatology of the upper atmosphere. Secular changes in the geomagnetic field will further alter the distribution of plasma and the location and intensity of ionospheric currents^{9,10}. Because the state of the upper atmosphere responds strongly to solar and heliospheric drivers, both anthropogenic and natural climate changes of terrestrial origin will, to varying degrees, depend on the phase of the solar cycle. Conversely, any terrestrial climate change is expected to alter how the upper atmosphere responds to solar variations.

The anthropogenic and natural evolution of geospace climate will have a profound effect on the low-Earth orbit (LEO) debris environment (Figure 2), because atmospheric drag is currently the only effective mechanism by which debris are removed from orbit. Of the 16,000 objects in the active satellite catalog (which only includes objects larger than 10 cm), 12,000 are debris or expended rocket bodies. The greatest density of debris is found between 500 and 1100 km altitude, where removal by satellite drag can take from 25 years to centuries¹¹. Recent studies

have found that even without any new launches, the amount of debris in LEO will increase via collisional fragmentation faster than it is removed from orbit by atmospheric drag¹². Active debris removal (ADR) strategies are therefore being explored, but little attention has yet been paid to how geospace climate change may affect these strategies. An initial study¹³ suggests that when long-term contraction of the thermosphere is taken into account, the effectiveness of ADR is dramatically reduced, thereby increasing the costs of achieving a stable debris population within the next hundred years (Figure 3). In addition to long-term contraction, anomalous reductions in thermospheric density, such as occurred unexpectedly during the recent solar minimum¹⁴, may also adversely affect efforts to control debris growth.

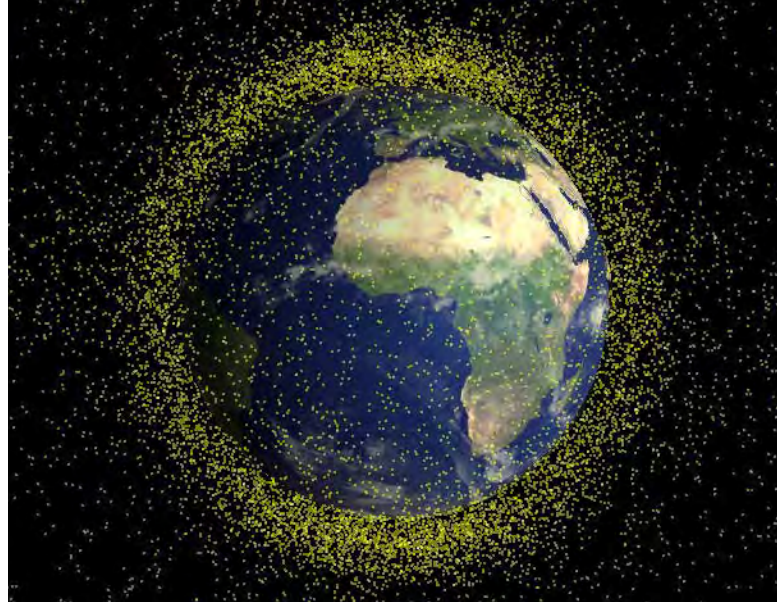


Figure 2. Snapshot of debris > 10 cm in low-Earth orbit on 1 May 2001. From Lewis et al., 2005⁴¹.

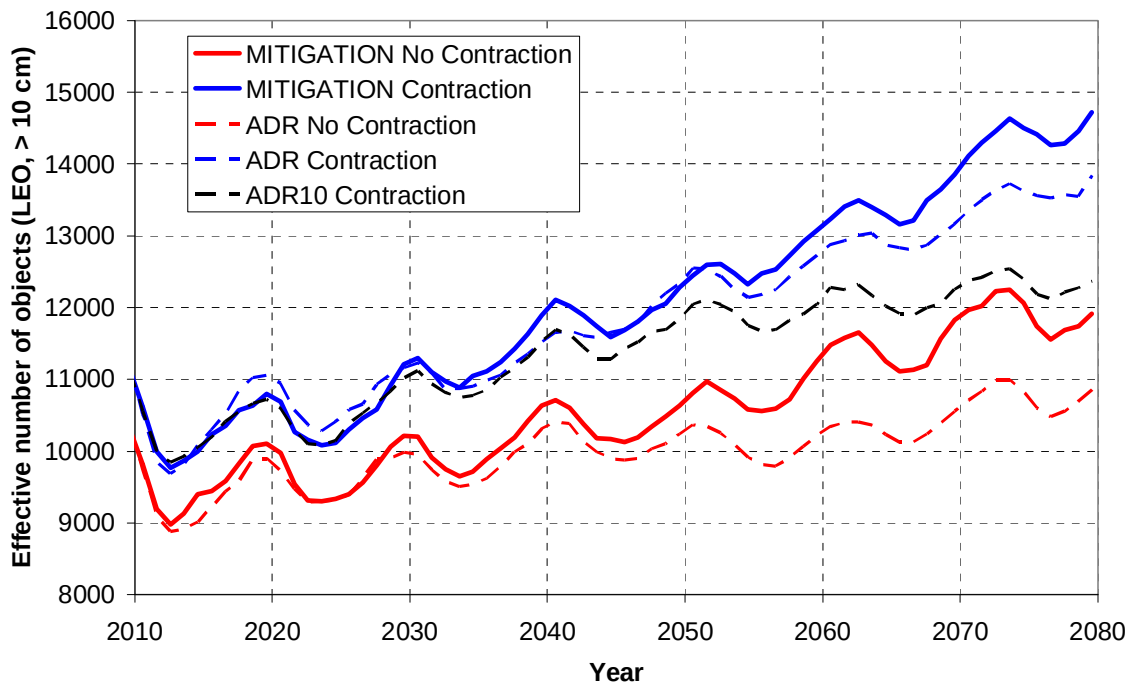


Figure 3. Simulation of orbital debris evolution under different scenarios of mitigation measures and upper atmospheric climate change. The red lines show the predicted number of objects assuming no climate change, and the blue lines apply a scenario of upper atmospheric contraction. The solid lines denote passive debris mitigation measures only, and the dashed lines denote active removal of 5 large objects per year. The black dashed line shows the predicted debris evolution when 10 objects per year are actively de-orbited, under the assumption of upper atmospheric contraction.

2. Advancing Geospace Climate Science

The climate questions facing the space science community are quite analogous to those that Earth scientists have been intensively investigating for the past few decades:

- *What is the typical systematic response of the upper atmosphere (which includes the ionosphere) to variations in solar and lower atmospheric drivers?*
- *How is the climate of the upper atmosphere evolving in response to anthropogenic changes in the lower atmosphere, and to long-term changes in natural drivers?*
- *How is upper atmospheric climate likely to change over the next 100 years?*
- *What are the implications of predicted changes for orbital debris management and the utilization of the radio wave spectrum?*

From a measurement perspective, investigation of geospace climate is likely to be more challenging than the detection and attribution of terrestrial climate, because of the generally greater difficulty of probing the vastly more extensive space environment and the lack of a long historical measurement record for most parameters. However, we believe that pivotal advances in the field can be affordably attained by adopting the following strategies:

A) Continuously monitor key upper atmospheric climate parameters.

The dominant rhythm of geospace climate is the 11-year solar cycle. Historically, space science has focused on the dramatic short-term perturbations forced by solar and geomagnetic activity, which are most prevalent during and immediately after the period of the sunspot maximum. However, it has long been recognized that solar maxima differ from one cycle to the next, and the 2008 solar minimum has illustrated that even the ‘ground state’ of the solar-terrestrial system is not reliably constant^{14,15}. Furthermore, theoretical and empirical evidence^{16,17,18,19} suggests that climatic changes of terrestrial origin become most apparent during solar minimum. **Consequently, because the system cannot be regarded as stationary even over century time scales, it is essential that it be continuously measured.**

Unfortunately, space-based measurements to date have generally been short-lived (less than a solar cycle), sporadic, and sparse. Unlike the Earth science community, which two decades ago implemented dedicated monitoring of 24 crucial observables, geospace remains poorly defined observationally (and theoretically, lacking sufficient model validation). Large temporal gaps between missions and the scarcity of data from previous solar cycles impose serious obstacles to the understanding and attribution of geospace climate. For example, the lack of lower thermospheric temperature and composition data during previous solar minima is a severe impediment to interpretation of the anomalous behavior of the entire thermosphere and ionosphere during the 2008 minimum. Ground-based instruments have historically provided much better temporal continuity than space-based measurements, but they are spatially very sparse for most types of instruments.

To advance space climate science, it is essential to improve the temporal coverage of space-based measurements and the spatial coverage of ground-based measurements. A mixture of space-based and ground-based assets is needed: Space-based measurements provide the coverage

necessary for detecting global shifts in climate, while ground-based measurements can provide a stable baseline and a link to past measurements. For all the required measurements, it is vital that continuity be maintained from one solar cycle to the next. **While a comprehensive, high-resolution specification of each and every upper atmospheric property is not required for geospace climate science, knowledge of ALL of the following key state parameters is essential:**

- *Neutral temperature* is perhaps the most important upper atmospheric climate parameter, since the vertical temperature profile and horizontal gradients fundamentally affect neutral and ion composition and dynamics.
- *Composition of major species O, O₂, N₂, He, H and their corresponding ions.* O is a key minor species in the mesosphere and lower thermosphere, but is the dominant species throughout most of thermosphere, and O⁺ is the dominant species of the ionosphere^{20,21}. The proportion of O near the base of the thermosphere, where O is strongly influenced by chemistry and dynamics, largely determines the composition of the upper thermosphere. He and H are minor species below about 400 km, but are the dominant species between 600 and 1000 km, where the greatest density of orbital debris resides. He and H exhibit large-scale spatial variations²², and anthropogenic increases in methane emissions may be affecting the H population in the geocorona^{1,23}.
- *Concentration of minor species CO₂, NO, O₃, H₂O, and CH₄.* CO₂, NO, and O₃ are (along with O) the dominant agents of radiative balance in the upper atmosphere²⁴. CO₂ is the primary cooling agent of the thermosphere. NO is also a major cooling agent, especially at high latitudes and during solar maximum; model simulations of CO₂ increases suggest that there may be significant feedbacks from NO⁷. CH₄ and H₂O are also radiatively active species, and CH₄ is a source of H₂O²⁵ and H¹. Observed changes in polar mesospheric ice cloud characteristics may be caused in part by H₂O trends, but hydrostatic effects associated with changes in the middle atmospheric temperature profile near the stratopause also play a major role⁵. At these altitudes, O₃ is the key constituent controlling the radiative budget.
- *Neutral wind* fundamentally affects, in a nonlocal manner, the electrodynamics and plasma distribution in the ionosphere²⁶, and also drives chemical transport^{6,27}. Currently, neutral winds are so poorly constrained that they are essentially free parameters in theoretical ionosphere models.
- *Wave activity* on all scales can affect the vertical and horizontal distribution of both neutral and charged particles. Small-scale waves such as gravity waves play an important role in the energy balance of the upper atmosphere²⁸.
- *Electric fields* drive ionospheric plasma dynamics and structure. They are generated by neutral winds, but are also imposed by the magnetosphere²⁹, and therefore must be measured independently of winds.
- *Electron density profile.* Knowledge of how the electron density profile is evolving is necessary for predicting long-term impacts of space climate change on communication and navigation systems. Ground-based instruments are an essential component for this parameter, because of their good geographic coverage, long measurement histories, and ability to measure all or most of the vertical profile.
- *Total mass density.* This parameter can be derived from the major species, but it is of major importance in its own right because of its direct relationship to satellite drag, which can be measured by accelerometers and derived from changes in satellite orbits. There is

an operational mandate to measure orbits of all objects larger than 10 cm, and it is expected that orbit data will be available indefinitely, so that funding for data processing is sufficient to produce orbit-derived mass density data. Accelerometer data provide much higher spatial and temporal resolution, and are therefore a valuable complement to orbit data.

- *Solar UV and EUV spectral irradiance.* Ultraviolet irradiance is the dominant energy source of the upper atmosphere, so accurate long-term knowledge of this driver is crucial. This measurement must be made from space and is prone to instrumental drift, so ample overlap between instruments and a vigorous calibration program are essential requirements. The Mg II core-to-wing ratio has proven to be a reliable indicator of UV irradiance and of the upper atmospheric response^{30,31}; this proxy overcomes some of the instrumental difficulties of UV measurements and is available beginning in 1978³². Continued monitoring and improvement of Mg II would provide a highly valuable asset for upper atmospheric climate studies.

Solar UV and EUV spectral irradiance is currently being monitored by TIMED, SORCE, and SDO. Otherwise, there are only two NASA satellites, C/NOFS and TIMED, dedicated to measuring critical state parameters of the upper atmosphere, and neither mission is expected to extend beyond 2014. C/NOFS was launched in 2008 and measures certain key ionospheric parameters via in-situ techniques at altitudes between 400 and 800 km. Since 2002, TIMED has remotely measured temperature, winds, and a few of the key minor species in the mesosphere and lower thermosphere (MLT); in the middle thermosphere, TIMED measured temperature and major species composition profiles until 2007. Thus, of the key parameters listed above, only temperature, winds, large-scale wave activity, and some minor species are currently being monitored in the MLT, and only sparsely. In the middle and upper thermosphere, the column O/N₂ ratio is the only quantity currently being provided by NASA missions. **There are no upper atmospheric missions planned for the next decade.** The Ionosphere Thermosphere Storm Probes mission, which received a high priority in the previous Decadal Survey (as part of the Geospace Network program)³³, following an extended and thorough science definition team study, is absent in NASA's plans^{34,35,36}, and presumably cancelled.

B) Support data analysis and modeling of upper atmospheric climate and upper atmospheric climate change.

Climate science is data-intensive by nature, since it requires large amounts of data of diverse types and varying quality to accurately determine the mean state or to validate theoretical predictions of the mean state, and rigorous statistical analysis is necessary to determine the evolution of the system. Funding opportunities should therefore place special emphasis on data analysis activities such as:

- Acquisition, preservation, digitization, and exploitation of existing data archives that can enhance the temporal and spatial extent of the upper atmospheric climate record. For example, unprocessed raw ionograms exist from several ionosonde stations, and standardized reprocessing of previously scaled ionograms can produce a uniform record of ionospheric parameters.
- Improvement of existing data inversion methods. Climate science requires continual assessment of biases among different data sets and within individual data sets. In many cases, such biases can be resolved by changes in data processing techniques.

- Core analysis of data for climate patterns and long-term change, including rigorous analysis of variances and uncertainties.
- Development of empirical and semi-empirical climate models. Data-based models are indispensable for tying together different data sets and different types of measurements in a physically consistent way, for filtering data to isolate specific climatic processes, and as a benchmark for assessing the quality of new data sets and for validating physics-based models.

As in lower atmospheric climate science, theoretical modeling should play a substantial complementary role. Important theoretical modeling activities include:

- Validation of model output against available data.
- Interpretation and attribution of climate patterns and trends.
- Long-term forecasting of upper atmospheric climate under different scenarios.
- Assessment of the impact of future changes on societal activities, and of the efficacy of possible mitigation strategies.

Funding from NASA for Heliophysics Research and Analysis projects has been low in relation to the size of the U.S. Heliophysics community and the number of worthy proposals submitted. Since 2003, the Heliophysics R&A proposal selection rate has been about 29%, which is slightly lower than the other Science Mission Directorate divisions (each of which funds over twice the number of R&A projects as Heliophysics)³⁷. In SMD as a whole in 2007 and 2008, only 45% of proposals rated Very Good were selected, and only 65% of proposals rated Very Good or higher were selected³⁷, suggesting that the community could be performing much more quality science than is currently funded. **Significant additional investment in Heliophysics R&A is badly needed in order to make meaningful progress in geospace climate science.**

C) Fill in critical gaps in the climate record.

In situ measurement of thermospheric neutral temperature and composition has not occurred since 1983. Mass spectrometers provided the biggest share of information about the upper atmosphere during the Explorer era, and today these instruments still have enormous capability to produce new discoveries and resolve long-standing problems. Unlike previous-generation instruments, modern mass spectrometers can differentiate between atomic and molecular oxygen, so it is now possible to fully probe the poorly understood composition between about 80 km, where O-O₂ chemistry and turbulent mixing are important, and 150 km, where diffusive separation is well established and O is free to become the dominate species of the thermosphere. Mass spectrometers can also provide needed validation of composition measured by the relatively new technique of UV remote sensing, as well as nightside composition and the density of species such as He and H, which cannot be inferred from UV emissions. Finally, mass spectrometers flown alongside accelerometers would provide badly needed closure on the problem of satellite drag coefficients and their dependence on object shape and ambient composition, which are the largest sources of uncertainty of mass density inferred from satellite drag.

CO₂ is the key species implicated in upper atmospheric climate change, but very few measurements of the vertical CO₂ profile have been made³⁸. CO₂ is assumed to be well mixed up to about 85 km, but this assumption has never been rigorously tested. New measurements of the complete CO₂ profile are needed to verify how quickly anthropogenic increases propagate into the upper atmosphere and to estimate how the influence of CO₂ on the upper atmosphere is evolving.

Other key upper atmospheric constituents, such as NO, have been measured for over 11 years, but not for more than one solar minimum: The UARS/HALOE NO time series begins in late 1991 and encompasses the 1996 solar minimum, but ends before the most recent solar minimum. Although the radiative effects of NO are considered to be relatively unimportant at solar minimum³⁹, it is a critical measure of the response of the lower thermosphere to both solar flux changes and temperature changes.

Thermospheric neutral winds play a fundamental role in the structure and dynamics of the ionosphere, but their behavior is still poorly understood, especially on the nightside between 110 and 200 km altitude, where only a handful of in situ measurements exist. In the upper thermosphere, there is currently very little space-based monitoring of winds (TIMED/TIDI began measuring Doppler shifts of 630.0 nm airglow in late 2008), and only a very small number of ground-based datasets have sufficient temporal extent to attempt a determination of long-term changes. Neutral winds must be taken into account in the interpretation of ionospheric climate, so it is vital that a reliable long-term measurement record be established.

3. Benefits and Relevance to Science and Society

The topic of upper atmospheric climate science and the strategies proposed above are highly relevant to the national and international space science enterprise, as follows.

Climate change in the geospace system was identified as a major theme of the NSF/CEDAR strategic plan for the next decade. Geospace climate is also directly relevant to Challenge 3 of the previous Decadal Survey, “Understanding the space environments of Earth and other solar system bodies and their response to external and internal influences”, and this challenge was partially addressed in NASA’s 2009 Heliophysics Roadmap with priority science target #7, “Climate Impacts of Space Radiation”. In addition to being a priority of U.S. research plans, space climate is also a cornerstone of the ongoing international SCOSTEP program “Climate and Weather of the Sun-Earth System”, in which major themes are solar influences on terrestrial and geospace climate and the geospace response to altered terrestrial climate.

Geospace climate science also contributes to the applications of orbit prediction, debris mitigation, and radio wave utilization. Understanding the climate of the upper atmosphere is prerequisite to fulfilling at least two directives of U.S. National Space Policy. The first is to preserve the space environment, in part by pursuing “research and development of technologies and techniques...to mitigate and remove on-orbit debris, reduce hazards, and increase understanding of the current and future debris environment” and leading “the continued development and adoption of international and industry standards to minimize debris.”⁴⁰ The second relevant space policy directive is to “take necessary measures to sustain the radiofrequency environment in which critical U.S. space systems operate.”⁴⁰

The strategies outlined in the previous section complement other observational systems and programs. Measurement requirements are very similar to those needed to advance space weather science; for climate science, the requirements are more basic and less stringent in terms of sampling density and short-term precision, but include a need for long-term monitoring. Most of the key climate parameters can be measured with well-developed, cost-efficient instrument technology. Geospace climate science contributes primarily to the Atmosphere-Ionosphere-Magnetosphere Interactions theme of the Decadal Survey, but there is also significant overlap with the Solar and Heliospheric Physics theme, because of the need for accurate specification of solar UV irradiance on time scales of days to decades.

A strong commitment to geospace climate research would contribute to two of the most important scientific questions facing the solar and space physics community. First, it would elucidate the intricate systematic interplay between solar and lower atmospheric influences on the near-Earth space environment. Second, it would address the natural upward extension of one of the most compelling science questions of our time: How are humans altering the climate of our planet?

The U.S. and international aeronomy communities are well-positioned to make significant advances in geospace climate science. There is already a vibrant, cooperative enterprise of study that only lacks a stable supply of data and support for analysis to bring our level of understanding up to par with that of terrestrial climate. There is also a strong culture of student development in the aeronomy community, across the three major skill areas needed for geospace climate science: instrument development and construction, data analysis, and theoretical modeling. The future is therefore highly promising for the specification and understanding of geospace climate, which is the foundation that underlies and informs all planetary space environment research.

References

- ¹Roble, R. G., and R. E. Dickinson (1989), How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere?, *Geophys. Res. Lett.*, **16**, 1441–1444.
- ²Laštovička, J., R. A. Akmaev, G. Beig, J. Bremer, J. T. Emmert, C. Jacobi, M. J. Jarvis, G. Nedoluha, Yu. I. Portnyagin, and T. Ulich (2008), Emerging pattern of global change in the upper atmosphere and ionosphere, *Ann. Geophys.*, **26**, 1255–1268.
- ³Akmaev, R. A., V. I. Fomichev, and X. Zhu (2006), Impact of middle-atmospheric composition changes on greenhouse cooling in the upper atmosphere, *J. Atmos. Solar-Terr. Phys.*, **68**, 1879–1889.
- ⁴Stevens, M. H., R. R. Meier, X. Chu, M. T. DeLand, and J. M. C. Plane (2005), Antarctic mesospheric clouds formed from space shuttle exhaust, *Geophys. Res. Lett.*, **32**, L13810, doi:10.1029/2005GL023054.
- ⁵Lübken, F.-J., U. Berger, and G. Baumgarten (2009), Stratospheric and solar cycle effects on long-term variability of mesospheric ice clouds, *J. Geophys. Res.*, **114**, D00I06, doi:10.1029/2009JD012377.
- ⁶Randall, C. E., V. L. Harvey, D. E. Siskind, J. France, P. F. Bernath, C. D. Boone, and K. A. Walker (2009), NO_x descent in the Arctic middle atmosphere in early 2009, *Geophys. Res. Lett.*, **36**, L18811, doi:10.1029/2009GL039706.
- ⁷Roble, R. G. (2010), "On Global Change in the Upper Atmosphere", 6th IAGA/ICMA/CAWSES workshop on Long-Term Changes and Trends in the Atmosphere, Boulder, Colorado, June 2010.
- ⁸Laštovička, J., R. A. Akmaev, G. Beig, J. Bremer, and J. T. Emmert, Global change in the upper atmosphere, *Science*, **314**, 1253–1254, 2006.
- ⁹Alfonsi, L., De Franceschi, G., and De Santis, A. (2008), Geomagnetic and ionospheric data analysis over Antarctica: a contribution to the long term trends investigation, *Ann. Geophys.*, **26**, 1173–1179.
- ¹⁰Elias, A. G., M. Z. de Artigas, and B. F. de Haro Barbas (2010), Trends in the solar quiet geomagnetic field variation linked to the Earth's magnetic field secular variation and increasing concentrations of greenhouse gases, *J. Geophys. Res.*, **115**, A08316, doi:10.1029/2009JA015136.
- ¹¹King-Hele, D. (1987), *Satellite Orbits in an Atmosphere: Theory and Applications*, Blackie, Glasgow.
- ¹²Liou, J.-C., and N. L. Johnson (2008), Instability of the present LEO satellite populations, *Adv. Space Res.*, **41**, 1046–1053.
- ¹³Lewis, H. G., A. Saunders, G. Swinerd, and R. Newland (2010), "Understanding the consequences of a long-term decline in thermospheric density on the near-Earth space debris environment", 6th IAGA/ICMA/CAWSES workshop on Long-Term Changes and Trends in the Atmosphere, Boulder, Colorado, June 2010.
- ¹⁴Emmert, J. T., J. L. Lean, and J. M. Picone (2010), Record-low thermospheric density during the 2008 solar minimum, *Geophys. Res. Lett.*, **37**, L12102, doi:10.1029/2010GL043671.
- ¹⁵Solomon, S. C., T. N. Woods, L. V. Didkovsky, J. T. Emmert, and L. Qian (2010), Anomalously low solar extreme-ultraviolet irradiance and thermospheric density during solar minimum, *Geophys. Res. Lett.*, **37**, L16103, doi:10.1029/2010GL044468.

- ¹⁶Rishbeth, H., and R. G. Roble (1992), Cooling of the upper atmosphere by enhanced greenhouse gases - modeling of thermospheric and ionospheric effects, *Planet. Space Sci.*, **40**, 1011–1026.
- ¹⁷Qian, L., R. G. Roble, S. C. Solomon, and T. J. Kane (2006), Calculated and observed climate change in the thermosphere, and a prediction for solar cycle 24, *Geophys. Res. Lett.*, **33**, L23705, doi:10.1029/2006GL027185.
- ¹⁸Qian, L., A. G. Burns, S. C. Solomon, and R. G. Roble (2009), The effect of carbon dioxide cooling on trends in the F2-layer ionosphere, *J. Atmos. Solar-Terr. Phys.*, **71**, 1592–1601.
- ¹⁹Emmert, J. T., J. M. Picone, and R. R. Meier (2008), Thermospheric global average density trends, 1967–2007, derived from orbits of 5000 near-Earth objects, *Geophys. Res. Lett.*, **35**, L05101, doi:10.1029/2007GL032809.
- ²⁰Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *J. Geophys Res.*, **107**, doi:10.1029/2002JA009430.
- ²¹Bilitza, D., and B.W. Reinisch (2008), International Reference Ionosphere 2007: Improvements and new parameters, *Adv. Space Res.*, **42**, 599–609.
- ²²Bailey, G. J., and R. Sellek (1992), Field-aligned flows of H⁺ and He⁺ in the mid-latitude topside ionosphere at solar maximum, *Planet. Space Sci.*, **40**, 751–762.
- ²³Kerr, R. B., R. Garcia, X. He, J. Noto, R. S. Lancaster, C. A. Tepley, S. A. Gonzalez, J. Friedman, R. A. Doe, M. Lappen, and B. McCormack, (2001), Secular variability of the geocornal Balmer-alpha brightness: Magnetic activity and possible human influences, *J. Geophys. Res.*, **106**, 28819–28830, doi:10.1029/1999JA900187.
- ²⁴Roble, R. G. (1995), Energetics of the mesosphere and thermosphere, The Upper Mesosphere and Lower Thermosphere, *Geophys. Monogr. Ser.*, **87**, 1–21.
- ²⁵Scherer, M., H. Vömel, S. Fueglistaler, S. J. Oltmans, and J. Staehelin (2008), Trends and variability of midlatitude stratospheric water vapour deduced from the re-evaluated Boulder balloon series and HALOE, *Atmos. Chem. Phys.*, **8**, 1391–1402.
- ²⁶Richmond (1995), The ionospheric wind dynamo: Effects of its coupling with different atmospheric regions, *Geophys. Monogr. Ser.*, **87**, 49–65.
- ²⁷Fuller-Rowell (1997), How does the thermosphere and ionosphere react to a geomagnetic storm?, *Geophys. Monogr. Ser.*, **98**, 203–225.
- ²⁸Yiğit, E., and A. S. Medvedev (2010), Internal gravity waves in the thermosphere during low and high solar activity: Simulation study, *J. Geophys Res.*, **115**, A00G02, doi:10.1029/2009JA015106.
- ²⁹Fejer, B. G. (1997), The electrodynamics of the low-latitude ionosphere: recent results and future challenges, *J. Atmos. Solar-Terr. Phys.*, **59**, 1465–1482.
- ³⁰Viereck, R., L. Puga, D. McMullin, D. Judge, M. Weber, and W. K. Tobiska (2001), The Mg II Index: A Proxy for Solar EUV, *Geophys. Res. Lett.*, **28**, 1343–1346.
- ³¹Lean, J. L., J. M. Picone, J. T. Emmert, and G. Moore (2006), Thermospheric densities derived from spacecraft orbits: Application to the Starshine satellites, *J. Geophys Res.*, **111**, A04301, doi:10.1029/2005JA011399.
- ³²Viereck, R. A., L. E. Floyd, P. C. Crane, T. N. Woods, B. G. Knapp, G. Rottman, M. Weber, L. C. Puga, and M. T. DeLand (2004), A composite Mg II index spanning from 1978 to 2003, *Space Weather*, **2**, S10005, doi:10.1029/2004SW000084.

- ³³National Research Council, Solar and Space Physics Survey Committee (2003), *The Sun to the Earth—and Beyond: A Decadal Research Strategy in Solar and Space Physics*, 196 pp., National Academies Press, Washington, D. C.
- ³⁴National Research Council, Committee on Heliophysics Performance Assessment (2009), *A Performance Assessment of NASA's Heliophysics Program*, 78 pp., National Academies Press, Washington, D. C.
- ³⁵NASA, 2009 Heliophysics Roadmap Team (2009), *Heliophysics: The Solar and Space Physics of a New Era*, 124 pp., available at http://sec.gsfc.nasa.gov/sec_roadmap.htm.
- ³⁶NASA (2010), List of Heliophysics missions under development, <http://science.nasa.gov/heliophysics/missions/>.
- ³⁷NASA (2010), Grant stats from the Service and Advice for Research and Analysis website, <http://science.nasa.gov/researchers/sara/>.
- ³⁸Fomichev, V. I., J.-P. Blanchet, and D. S. Turner (1998), Matrix parameterization of the 15 micrometer CO₂ band cooling in the middle and upper atmosphere for variable CO₂ concentration, *J. Geophys Res.*, *103*, 11,505–11528.
- ³⁹Mlynczak, M. G., et al. (2010), Observations of infrared radiative cooling in the thermosphere on daily to multiyear timescales from the TIMED/SABER instrument, *J. Geophys Res.*, *115*, A03309, doi:10.1029/2009JA014713.
- ⁴⁰United States (2010), *National Space Policy of the United States of America*, available at http://www.whitehouse.gov/sites/default/files/national_space_policy_6-28-10.pdf.
- ⁴¹Lewis, H. G., G. Swinerd, C. Ellis, and C. Martin (2005), “Response of the space debris environment to greenhouse cooling”, 4th European Conference on Space Debris, Darmstadt, Germany, April 2005.

Concept Paper: An Investigation of the Coupling of the Earth's Atmosphere to its Plasma Environment

Scott England¹, Harald Frey¹, Andrew Stephan², Joseph Huba², Christoph Englert², Geoff Crowley³, Gary Swenson⁴, Jonathan Makela⁴ and Astrid Maute⁵

Executive Summary

Ion-neutral coupling is a process that is common to all planetary atmospheres and stellar photospheres. The two-way coupling of the neutral atmosphere to the ionosphere at Earth is a natural, within-reach laboratory in which this process can be studied. The low-latitude ionosphere at Earth is a relatively limited system where it should be possible to unambiguously discern the effects of ion-neutral coupling from other atmospheric processes. Although many of the individual processes in this region have been identified before, our understanding of the complete system is seriously lacking. What is needed is a simultaneous study of both the ionized and neutral gases in a coordinated manner, where the measurements of each are ideally in a common volume or at least linked by the same magnetic field-line. Here we propose a mission concept that addresses this topic in the simplest way possible, using flight-proven instrumentation. This would provide the coordinated observations needed to make concrete advances in our understanding of this coupled system.

1 Scientific Motivation

The Earth's ionosphere is formed primarily via photo-ionization of the upper atmosphere. This forms the densest plasma in the geospace environment. Yet, even at its most dense, the ionosphere remains a partially ionized gas and as such the interactions between the ions and neutrals are critical to understanding this coupled system. Indeed, ion-neutral coupling is important in all planetary systems

(upper atmosphere / exosphere) and stars (photosphere). Such systems span a wide range of parameter space (density, thermal conditions, plasma β), but certain processes remain common to all. These interactions have been called out as central to our understanding of Solar and Space Physics. For example, the Science Plan for NASA's Science Mission Directorate (2007-2016), Heliophysics division calls out as a primary task to '*understand the coupling between planetary ionospheres and their upper atmospheres as mediated by strong ion-neutral interactions.*' The Earth's ionosphere-thermosphere is the easiest of these environments to probe and is thus a prime candidate for studying plasma-neutral coupling.

The Earth's low-latitude ionosphere resides on closed magnetic field lines and therefore compared to high-latitudes, magnetospheric effects are much weaker. Thus, during magnetically quiescent conditions, one may consider this to be a system consisting of just three elements: the atmosphere, the ionosphere and solar radiative input. Comprehensive models have been developed that represent the elements of this system and couple these to one another, either self-consistently or with limited feedback between the elements. One may therefore presume that within such a relatively closed system it should be possible to predict the state of the plasma. While such models do an acceptable job of simulating monthly-mean type climatologies, they consistently fail to simulate the day-to-day changes in the ionosphere. Indeed, no model exists that can forecast the global-scale state of this region with any reasonable confidence.

Figure 1 highlights the challenge such models face. This shows the vertical ion motion measured by the JULIA radar in Peru (magnetic equator) for August-October over a variety of solar flux conditions. While a day-night pattern is seen in the average of these

¹ University of California, Berkeley, California

² Naval Research Laboratory, Washington D.C.

³ ASTRA, Boulder, Colorado

⁴ University of Illinois, Illinois

⁵ National Center for Atmospheric Research, Boulder, Colorado

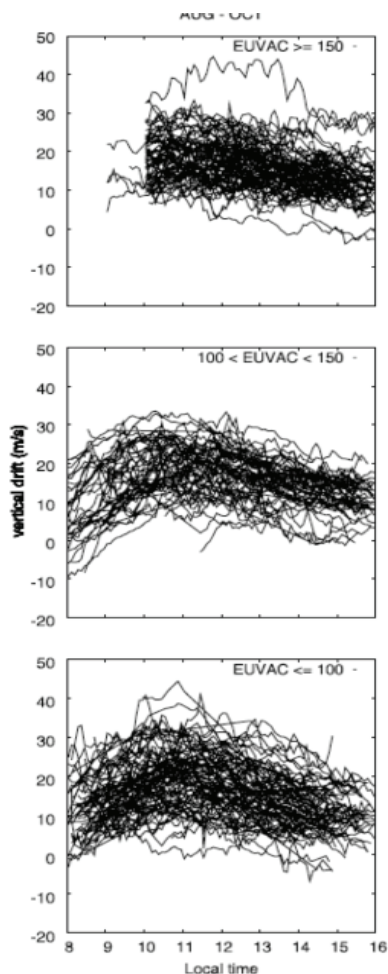


Figure 1: Vertical ion motion measured by the JULIA radar for magnetically quiet conditions. Data are divided into low, medium and high solar EUV flux. After Alken 2009.

drifts, it is clear that such patterns only describe the monthly-mean properties and do not represent the day-to-day state of the system. Indeed, the challenge is all the more evident when we note that this plot includes only data magnetically quiet conditions ($K_p < 3$) when magnetospheric and solar-wind inputs are not strongly affecting the system. With such a gap in our understanding of the ions in this region, our progress in defining their feedback to the neutral atmosphere is very much in its infancy, although observations of the neutral density close to the magnetic equator suggest this feedback could be surprisingly strong.

Numerous observations of both the ionosphere and thermosphere have been

made, both from the ground and space. These have allowed us to categorize the independent processes that exist, but our understanding of the behavior of the system is severely lacking. What is needed is an investigation that measures both the drivers and the response of this system of ion-neutral coupling in a systematic and coordinated way. In this paper we describe the parameters of the system that are needed (Section 2), the observations through which these can be obtained (Section 3), the state of the art instrumentation that would be capable of making these observations with sufficient accuracy and precision (Section 4) and one possible measurement scenario that would allow for such a coordinated study (Section 6). Indeed, given the extraordinary temporal and spatial variability seen in this system, even on a global scale, it is only through such a coordinated study that observes both the neutrals and ions simultaneously that we will come to a better understanding of this system.

2 Parameters Required for this Study

We seek to understand the pathways and nature of the coupling between the ionosphere and thermosphere at scales relevant for determining the global-scale behavior of the system. Global-scale models typically have resolutions of a few degrees horizontally and a few km vertically, and while interesting processes that create localized gradients and motions on scales significantly smaller than this do exist (e.g. meter-scale turbulence in plasma irregularities), we will focus on processes relevant to the bulk motion and evolution of the system on temporal scales of hours, rather than seconds.

On these scales, the key parameters of the neutral atmosphere are the horizontal winds, density and composition. For neutral winds, the hydrostatic approximation holds under all conditions found at low- and mid-latitudes, therefore only horizontal winds are needed for a complete understanding of the motion of the atmosphere. For composition, O and N_2 are the primary source and sink of O^+ , the main ionospheric species at Earth. The O^+ density must furthermore be sampled in conjunction with ionospheric conductivity,

Parameter Required (Section 2)	Observable Quantity (Section 3)	Instrument / Model (Section 4)
Horizontal winds, 2 vector components, 100-300 km alt	Doppler shift of 557.7 nm and 630.0 nm O emission	Interferometer
Neutral density, 150-250 km	Brightness of 135.6 nm O and LBH N ₂ band emissions	Far-UV spectrometer
Ratio of thermospheric O to N ₂	Brightness of 135.6 nm O and LBH N ₂ band emissions	Far-UV spectrometer
O ⁺ density, 150-500 km	Brightness of 83.4, 61.7 nm OII emission	Extreme-UV spectrometer
Electric field & plasma drift	Ion ExB motion in-situ	Ion Velocity Meter
Conductivity	Brightness of 83.4 nm OII emission	EUV spectrometer, assimilative model
Ionospheric currents	Scalar and vector magnetic field perturbations	Scalar and Vector Magnetometer
Solar EUV input	Brightness of 135.6 nm O and LBH N ₂ band emissions	Far-UV spectrometer
Magnetospheric input	High-latitude radar, magnetometer	Assimilative model
Horizontal winds, 2 vector components, 100-300 km	Doppler shift of 557.7 nm and 630.0 nm O emission	Interferometer

Table 1: The parameters needed for the proposed investigation, the observable quantities used to find these and the proposed category of instrument that could be used to obtain these observations. Each of the instruments is color-coded for easy reference later in the text.

the electric potential generated by the wind, gravity and plasma pressure gradient driven dynamos, the bulk motion of the O⁺ layer, and main currents that flow in the ionosphere. All of these must finally be put in context of solar EUV radiative input that drives much of the photo-effective processes in the upper atmosphere. However, it is not sufficient to sample these at one altitude as each parameter varies greatly over small vertical scales. For example, large-scale atmospheric waves can cause strong shears in the wind over an altitude range of 10s km.

All of these requirements are summarized in the first column of Table 1. Magnetospheric inputs are added to the list of key variables to ensure their impact is small during the period under study. This table also lists the observable quantities and proposed instrumentation / model that can be used to obtain these parameters. The following two sections will detail each of these respectively.

3 Observations and Models Required for this Investigations

None of the observations or instruments listed in this and the following section are new or require the development of new technologies. Rather it is the combination of these measurements and the observing scenario described in Section 5 that is critical to using these tested technologies to perform this fundamentally new investigation.

Of the parameters listed in Table 1, all but conductivity and magnetospheric inputs can be observed directly or indirectly from a low to mid-latitude, low-earth orbit spacecraft. Those that cannot be observed from this location can be obtained in near real-time using assimilative models such as AMIE and IDA3D (Bust *et al.*, 2007a,b; Richmond *et al.*, 1992).

Both the Earth's atmosphere and ionosphere have widely different spatial scales in the vertical and horizontal directions at low latitudes. As a simple results of hydrostatic balance and optical depth, the vertical scales over which parameters changes significantly are of the order of a few km, whereas horizontal

scales are 100s-1000s km. There is a need to sample the relevant fields over a range of locations and altitudes, which cannot be done from the ground or with nadir viewing instrumentation. Both of these requirements can however be met by a combination of limb-viewing optical instruments and in-situ measurements of the electric and magnetic fields when the spacecraft crosses magnetic field lines that intersect the limb observations. This combination of in-situ and remote sensing is the most obvious way to fulfill the requirements given in Table 1. The relevant optical wavelengths and other parameters for these observations are given in the second column of Table 1.

With regard to the horizontal winds, a remote vector measurement must be made. This has previously been accomplished by using 2 view directions at 90° to one another that sample a common volume as the spacecraft moves. There is a time separation between these measurements of typically 5-10 minutes for a LEO orbit, which is small compared to the scales we wish to study.

It is also worth noting that measurements of the electric field can be related to the potential generated at lower altitudes (dynamo region ~110-180 km) as the magnetic field lines that spacecraft intersects in this region are effectively equipotentials.

4 Instrumentation

Below we describe the most feasible instrument options to fulfill the requirements in Table 1. We will not propose specific instruments from different instrument groups, but rather the categories of instruments that have flight heritage and have demonstrated sufficient performance to make the observations described.

Interferometer

Limb-viewing Fabry-Perot and Michelson interferometers have been shown to be capable of measuring horizontal winds in the thermosphere. Fabry-Perots have flown on DE, UARS and TIMED (Hays *et al.*, 1981; Killeen *et al.*, 1999) and Michelsons on UARS (Shepherd *et al.*, 1993). Both instruments can be made to sample small fields of view that

are scanned to build up an altitude profile, or image a wider scene. Some Michelson interferometers also require stepping to sample four or more path differences introducing another moving part in these instruments. It has been demonstrated by the SHIMMER SHS interferometer (Englert *et al.*, 2007; 2010) that replacing the mirrors with grating can eliminate the need for this stepping. The relevant emission lines to cover the altitude range of interest are the 557.7 and 630.0 nm emissions from O, both of which have been previously used for such measurements.

Far Ultra-Violet Spectrometer

The measurement of O and N₂ in FUV emissions in the 135.6 nm line and Lyman-Birge-Hopfield band respectively has been demonstrated by the IMAGE, TIMED and DMSP satellites. The key processes involved and the inversion of the brightness to species abundances is well understood (Meier, 1991). In addition to the ratio of these two species, it is possible to use these same observations to obtain the absolute density and a measure of the solar EUV brightness required without

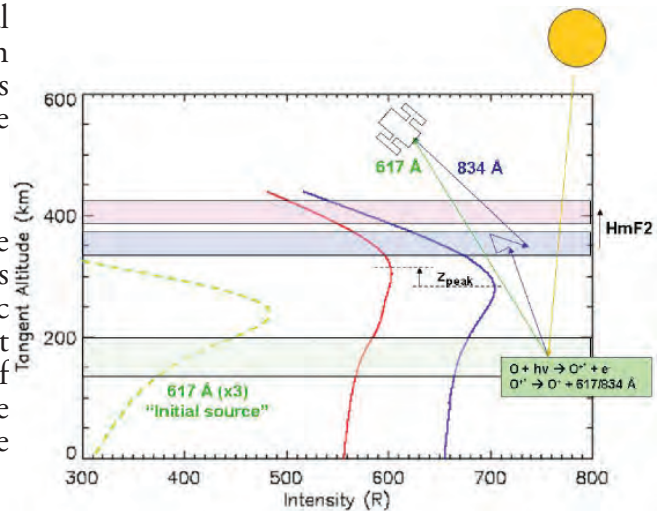


Figure 2: Depiction of the dayside 83.4 nm ionospheric O⁺ observation. Solar EUV excites bright terrestrial sources of 61.7 and 83.4 nm emissions which have the same initial state. The F-region ionosphere we wish to study is optically thin to 61.7 but scatters 83.4, allowing both the initial source brightness and the altitude profile of O⁺ to be established.

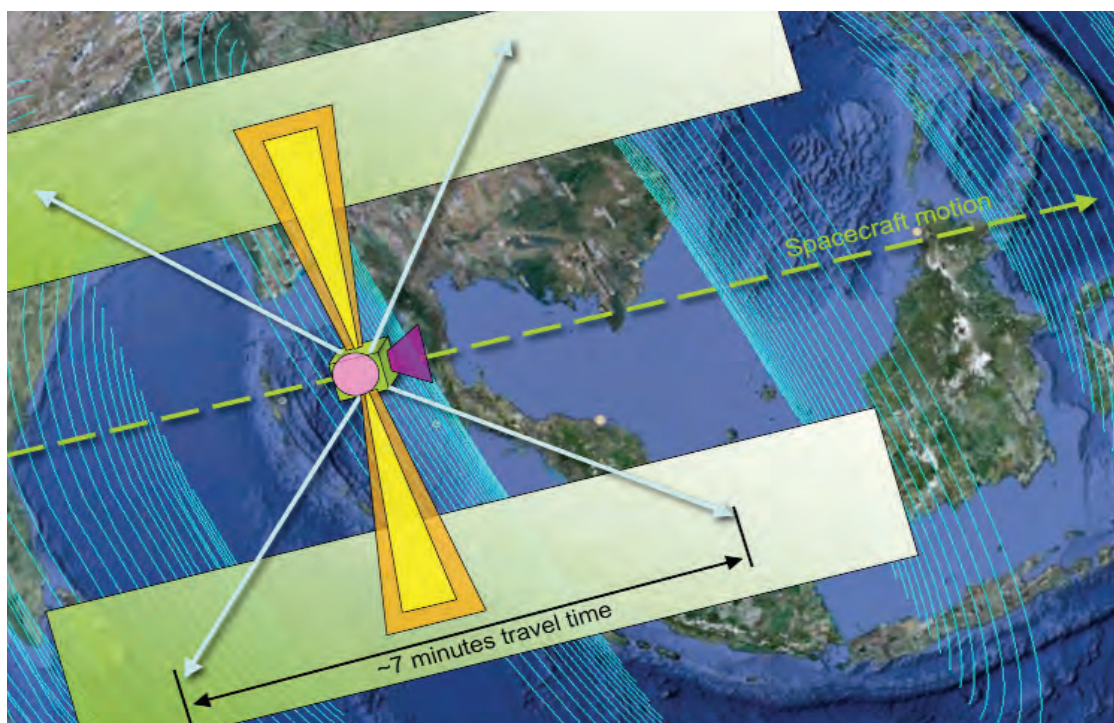


Figure 3: Proposed spacecraft with approximate instrument fields of view shown, color-coded to match Table 1. Approximate sample area of remote sensing instruments is shown (green) along with magnetic field lines (blue). As the spacecraft flies from left to right, it samples all fields, including vector winds in approximately 7 minutes.

an independent observation with another instrument. Either a 2-channel spectrographic imager (e.g. IMAGE-FUV) or an imaging spectrometer (e.g. TIMED-GUVI) would be capable of making the necessary observations with sufficient resolution to separate the 135.6 nm emissions from the much brighter 130.4 nm line.

Extreme Ultra-Violet Spectrometer

Remote sensing of the dayside ionospheric O^+ density using the 83.4 nm emission has been demonstrated by the RAIDS instrument onboard the International Space Station and the SSULI sensor on the latest DMSP satellite. The overall concept is depicted in Figure 2. Solar EUV ionizes an inner shell electron from O in the lower atmosphere. The resulting decay produces the 83.4 nm emission that radiates upwards and scatters through O^+ in the F-region ionosphere, yielding an altitude profile that characterizes the O^+ density during the day. The accuracy of this method is greatly improved if it is known what the magnitude is

of the initial photoionization source that then illuminates the ionosphere from below. This can be modeled, but a measurement of the optically thin companion emission at 61.7 nm, first measured routinely by the RAIDS experiment, can be used as a direct measure of that vital piece of information. Thus, this instrument must be capable of measuring both of these wavelengths simultaneously to characterize the absolute density of O^+ .

Plasma Drifts and Ionospheric E-Fields

The magnetic field lines at low latitudes are effectively equipotentials. This means that a spacecraft observing the local plasma drift (vector, in-situ measurement) is effectively measuring the electric field in the dynamo region along the intersected magnetic field line (E-region during the day). The combination of this observation made simultaneously with the remote sensing limb observations described above allows for a complete characterization of the driver of the E-region dynamo (horizontal winds), the resultant

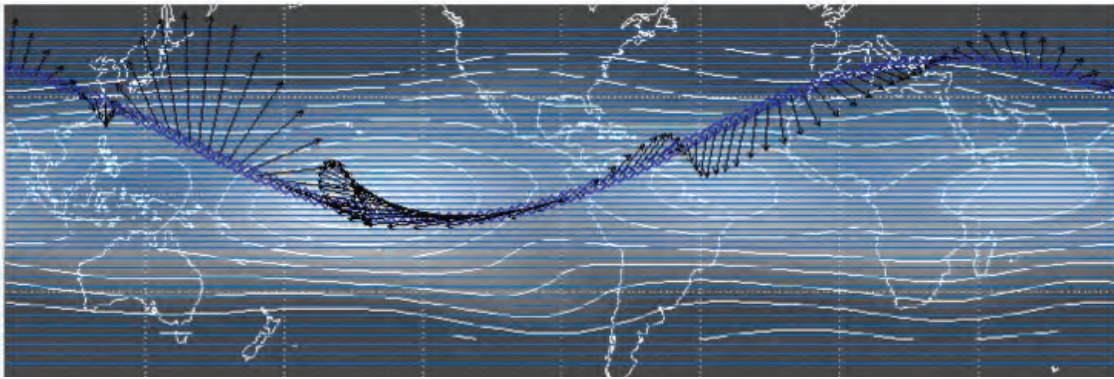


Figure 4: Latitudinal and longitudinal coverage of the limb-viewing instrumentation proposed. See text for details.

electric field and the O^+ plasma motion. Such measurements have been made independently before, but they have rarely overlapped and never been made in a coordinated fashion. Examples of the instrument required to make the in-situ drift measurements are the ion velocity meters on ROCSAT-1 and C/NOFS. Wire-boom electric field instruments are incompatible with the 3-axis stabilized platform required for the optical instruments.

Vector and Scalar Magnetometers

Currents flowing in the ionosphere can be measured remotely with magnetometers onboard a low Earth orbit (LEO) spacecraft. The best-known and strongest current system at low-latitudes is the equatorial electrojet, which flows along the magnetic equator. This can be observed with a scalar magnetometer (e.g. CHAMP; e.g. Lühr *et al.*, 2004). The measurement of the F-region dynamo can be made with a vector magnetometer (e.g. CHAMP; Lühr and Maus, 2006).

5 Requirements Placed on the Spacecraft

The instruments required (Section 4), measurements needed (Section 3) and coordinated observing scheme (Section 6) place constraints on the spacecraft that go beyond generic mass, power and field-of-view requirements. These are:

Orbit

A circular orbit is required for the limb-viewing instruments. LEO is required for sampling of the plasma motion and closed

magnetic field lines. An inclination of $\sim 25^\circ$ is required for sampling the magnetic equator at all longitudes, allows rapid sampling of all local times and regular revisiting of each longitude.

Attitude Control System

A 3-axis stabilized spacecraft is required for the limb-viewing instruments. Sufficient stability is required for the integration times of all instruments (similar requirements to the stability of e.g. TIMED is appropriate). Pointing knowledge is required for vector measurements and latitude profiles. Location knowledge is required for the translation between the inertial and Earth-centered frames (this can be fulfilled with either a GPS receiver or telemetry ranging from a ground-station).

Cleanliness

Magnetic cleanliness is required for the plasma drift and current measurements. The electric potential of the spacecraft must be kept small for the plasma drift measurements. Strict parts selection is required for all spacecraft components to prevent contamination of the EUV and FUV instruments.

6 Mission Scenario

The instruments and spacecraft described in the two previous sections must be used in the configuration depicted in Figure 3 to obtain all of the measurements described in Table 1 in a coordinated manner. The optical instruments point towards the limb where they sample altitude profiles of each remotely

sensed parameter. These are coordinated with the plasma motion and E-field measurements at the spacecraft whenever it intersects a magnetic field line in the vicinity of the magnetic equator (as shown). As the spacecraft flies from west to east, it samples all parameters, including the vector winds within ~7 minutes of each other, meeting the requirements described earlier. Measurements of the currents are made continuously and can be related to the other measurements as a complete picture is built up.

From a 25° orbit, the latitudes sampled by the optical instruments is shown in Figure 4 (blue). Overlaid is one orbit of vector winds from one side of the spacecraft (arrows). The white contours show simulated global-scale wave fields from a numerical model that such observations would help to constrain.

7 Cost Estimate and Basis of Estimate

We have proposed a relatively straightforward mission scenario that requires no new technologies to implement. The duplication of instruments as shown in Figure 3 should also help to limit payload costs. The LEO orbit at 25° can be easily reached from a Florida launch site. All of the instruments suggested have flown in some form in the last decade, so their masses and powers can be estimated from these heritage instruments. We estimate the total payload to weigh less than 200 kg and require less than 200 W of power. The total payload cost should be around \$100m in FY10. No very stringent requirements are placed on the spacecraft, so a launch vehicle with the ability to lift ~600 kg to LEO should be sufficient. Extrapolating from these numbers, a total mission cost of order \$400M should be realizable, making this a small mission as defined by the RFI.

References

Alken, P., 2009, A quiet time empirical model of equatorial vertical plasma drift in the Peruvian sector based on 150 km echoes, *J. Geophys. Res.*, 114, doi:10.1029/2008JA013751.

Bust, G. S., T. W. Garner, T. L. Gaussiran II, R.W. Meggs, C. N. Mitchell, P. S. J. Spencer, P. Yin and B. Zapfe, and G. Crowley, Four Dimensional GPS Imaging of Space-Weather

Storms, *Space Weather Journal*, in press, 2007a.

Bust, G. S., and G. Crowley, Tracking of Polar Cap Patches using Data Assimilation, *J. Geophys. Res.*, 112, A05307, doi:10.1029/2005JA011597, 2007b.

Englert C.R., D.D. Babcock, J.M. Harlander, Doppler Asymmetric Spatial Heterodyne Spectroscopy (DASH): Concept and Experimental Demonstration, *Applied Optics*, 46, 7297-7307, 2007.

Englert, C. R., M. H. Stevens, D. E. Siskind, J. M. Harlander, and F. L. Roesler (2010), Spatial Heterodyne Imager for Mesospheric Radicals on STPSat-1, *J. Geophys. Res.*, 115, D20306, doi:10.1029/2010JD014398.

Hays, P. B., T. L. Kileen, and B. C. Kennedy, The Fabry-Perot on Dynamics Explorer, *Space Sci. Instrumentation*, 5, pp 394-416, 1981.

Kileen, T. L., et al. (1999), TIMED Doppler Interferometer (TIDI), *Proc. SPIE Int. Soc. Opt. Eng.*, 3756, 289 – 301, doi:10.1117/12.366383.

Lühr, H., S. Maus, and M. Rother (2004), Noon-time equatorial electrojet: Its spatial features as determined by the CHAMP satellite, *J. Geophys. Res.*, 109, A01306, doi:10.1029/2002JA009656.

Lühr, H., and S. Maus (2006), Direct observation of the F region dynamo currents and the spatial structure of the EEJ by CHAMP, *Geophys. Res. Lett.*, 33, L24102, doi:10.1029/2006GL028374.

Meier, R. R., Ultraviolet Spectroscopy and Remote Sensing Of The Upper Atmosphere, *Space Sci. Rev.*, 58, 1, 1991.

Richmond, A.D., E.C. Ridley, R.G. Roble, A thermosphere/ionosphere general circulation model with coupled electrodynamics, *Geophys. Res. Lett.*, 19, 601, 1992.

Shepherd, G. G., and the WINDI team, et al., WINDI: The wind imaging interferometer on the upper atmosphere research satellite, *J. Geophys. Res.*, 98, 10725-10750, 1993.

**Spatial Heterodyne Spectroscopy:
An Emerging Optical Technique for Heliophysics and Beyond**

Christoph R. Englert

US Naval Research Laboratory, Space Science Division
4555 Overlook Ave SW, Washington DC 20375, USA

John M. Harlander

St. Cloud State University, Department of Physics, Astronomy and Engineering Science
720 Fourth Avenue South, St. Cloud, MN 56301, USA

Frederick L. Roesler

University of Wisconsin–Madison, Department of Physics
1150 University Avenue, Madison, WI 57306, USA

Abstract

At the end of the 1980s, an innovative optical technique called Spatial Heterodyne Spectroscopy (SHS) has become practical for scientific research applications with the availability of imaging detectors and more capable and affordable computers. SHS is similar to a conventional Fourier Transform Spectrometer since it also uses a Michelson type interferometer, but it samples the entire interferogram simultaneously in the spatial domain without the need for moving parts. In the past two decades, SHS has been used for laboratory spectroscopy, sounding rocket instruments, and space-borne experiments. The SHS concept has now reached a level of maturity that allows the serious consideration for NASA missions. Applications in the field of heliophysics include the measurement of upper atmospheric composition, winds and temperatures. Currently, observations of thermospheric winds present a major gap and their future availability will facilitate significant improvements in space weather prediction and our understanding of the atmosphere-ionosphere-magnetosphere interactions, including the effects of variable solar forcing from co-rotating interaction regions or coronal mass ejections. In addition to heliophysics applications, SHS is also a promising candidate for planetary applications, like the measurement of methane on Mars, or earth-science applications like the measurement of tropospheric winds. Continued development and application of this still novel technique will strengthen the position of SHS as a mature optical technique that is available for NASA missions. The unique properties of SHS have a strong potential to improve and enable future measurements, while minimizing resource requirements on the respective platforms.

1. This is how SHS works

1.1. The Basic SHS Concept is a Fourier Transform Spectrometer without Moving Parts

SHS is a method of Fourier transform spectroscopy based on a Michelson interferometer modified by replacing the mirrors in each arm with fixed diffraction gratings as shown in Figure 1 [Harlander *et al.*, 1992, 2003, 2004]. For each wavelength in the wave front entering the interferometer, two wave fronts with a wavelength dependent crossing angle between them exit the interferometer (see Fig. 1). The resulting superposition of Fizeau fringes with wavelength dependent spatial frequencies is localized near the gratings and imaged by exit optics on a position-sensitive detector. Since each spectral element at the input is modulated by a unique spatial frequency at the output, the fringe image is a constant plus the Fourier transform of the input spectrum about the heterodyne wavelength (the wavelength producing parallel output wave fronts). The Fourier transform of the fringe image produces the spectrum within a limited spectral range, with the zero of the transform corresponding to the heterodyne wavelength.

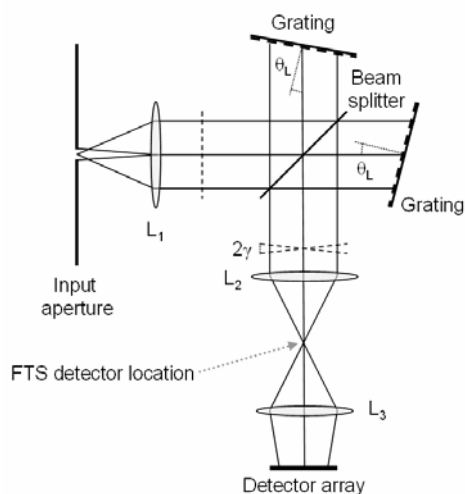


Figure 1: Schematic diagram of the basic SHS configuration. For each wavelength in the incident wave front, two wave fronts with a wavelength-dependent crossing angle between them exit the interferometer. The resulting superposition of Fizeau fringes with wavelength-dependent spatial frequencies is localized near the gratings and imaged by exit optics on a position-sensitive detector. The image is the Fourier transform of the input spectrum about the heterodyne wavelength (the wavelength producing parallel output wave fronts). The arrow indicates the typical location of the detector for FTS and Fabry-Pérot spectrometry. From this position the interferometer elements are completely out of focus.

With SHS, high spectral resolution can be achieved with only modest requirements on the spatial resolution of the detector. If the fringe pattern is imaged by N pixels in the dimension parallel to Fig. 1, $N/2$ spectral elements can be obtained without aliasing. An interference filter can be used to eliminate aliasing and multiplex noise from out-of-band light. SHS achieves a maximum resolving power equal to the theoretical resolving power of a dispersive (grating and prism) system; at the same time, its field of view is characteristic of interferometric spectrometers. When expressed as a solid angle, the field of view of a non-fieldwidened SHS is approximately 100 times larger than that of a conventional grating spectrometer of the same resolution.

Without degrading the resolving power, fixed field-widening prisms can be placed in the arms of the interferometer to increase the solid angle field of view by another factor of ~ 100 . The prism angle is chosen so that the gratings – when viewed from the exit of the interferometer – appear, from a geometrical optics standpoint, to be coincident to and perpendicular to the optical axis. Depending on the detector and optics before the interferometer, zero, one, or two dimensions of spatial information can be recorded. Compared to scanning Fourier Transform Spectrometers (FTS) and Fabry-Pérot spectrometers (FPS), SHS instruments have greatly relaxed alignment and surface figure tolerances. This is because in SHS, the elements of the

interferometer are nearly imaged on the detector. As a result, each detector pixel integrates only over a small area in the interferometer and misalignments of a few wavelengths shift the heterodyne wavelength but do not reduce the fringe contrast. As long as the optical quality of the elements is good over the small area sampled by a detector pixel, flatness errors on the gratings, prisms and beamsplitter, or index of refraction inhomogeneities, distort the Fizeau fringe pattern but do not reduce its contrast. The measured response to monochromatic light sources provides a measure of the phase errors that can be used in software to correct broadband interferograms with minimal impact on the signal-to-noise ratio in the recovered spectrum [Englert *et al.*, 2004]. The typical location for the detector in a FPS and FTS is indicated in Fig. 1. At this location, the interferometer elements are completely out of focus and optical misalignments and/or flatness errors of more than a fraction of a wavelength reduce the detected fringe contrast and the signal-to-noise ratio of the spectrum.

Figure 1 shows an SHS implementation that includes transmissive elements, like the beamsplitter. If only reflective, rather than transmissive optical elements are preferred for a specific application, for example because transparent, homogeneous materials are not available for a specific wavelength region, SHS can also be implemented in an all reflective configuration. Measurements in the far ultraviolet are an example for which an all reflective SHS might be a good choice [Harlander *et al.*, 1992].

1.2. The Basic SHS Concept Can be Modified to Meet Specific Performance Requirements

The basic SHS configuration lends itself to modifications in order to meet challenging measurement requirements. The bandpass of SHS can, for example, be increased by using echelle gratings, rather than first order gratings [Lawler *et al.*, 2008]. Another example is a tunable, all reflection SHS, which allows the rapid change of the spectrometer’s passband [Dawson and Harris, 2009]. A third type of modified SHS is a concept that is optimized for measuring Doppler shifts for remote sensing measurements of atmospheric winds. Since most of the high spectral resolution information about the line center of an atmospheric emission line is contained in a part of the interferogram that is offset from the zero path position (center of the detector array in Fig. 1), one can save mass, size, and power resources of the experiment by measuring this part of the interferogram only. The modification to the basic SHS concept necessary to accomplish that is to simply lengthen one interferometer arm with respect to the other. This technique is called Doppler Asymmetric Spatial Heterodyne (DASH) spectroscopy [Englert *et al.*, 2007]. It is very similar to the stepped Michelson technique which was applied very successfully by the WINDII experiment on the NASA UARS satellite [Shepherd *et al.*, 1993]. Some advantages over the stepped Michelson are that DASH has no need for moving parts, it provides many more interferogram samples which are measured simultaneously, and it can observe several emission lines simultaneously.

2. Past SHS applications provide technical readiness for future missions

SHS was first utilized for high spectral resolution measurements of diffuse sources in the ultraviolet [Harlander *et al.*, 1992, 1994]. For these applications, when compared to other conventional spectroscopic techniques, the main advantages of SHS are the high interferometric throughput, high spectral resolution, the compact, rugged package, and the relaxed alignment and fabrication tolerances. SHS instruments for the UV were also flown on sounding rockets [Watchhorn *et al.*, 2003, Stephan *et al.*, 2001] and the Space Shuttle [Cardon *et al.*, 2003] taking advantage of their sensitivity, ruggedness and lack of moving parts. The fact that the SHS measurements are performed in the spatial domain rather than the temporal domain, which makes

October 2010

them insensitive to unwanted effects from changing scenes, was a main driver to develop an SHS instrument for atmospheric remote sensing in the thermal infrared [Englert *et al.*, 2009].

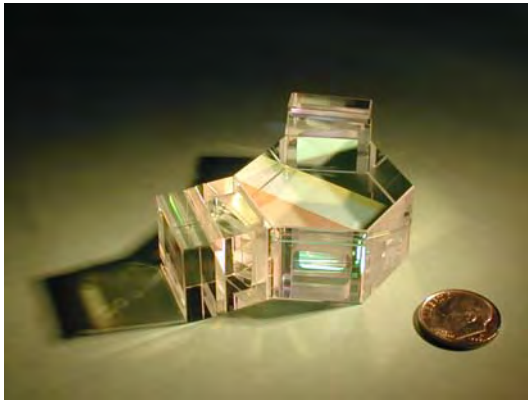


Figure 2: Photograph of the monolithic SHS interferometer for the near ultraviolet flown on as part of the SHIMMER payload on STPSat-1 [Harlander *et al.*, 2004, Englert *et al.*, 2010a].

The single most significant advance of the SHS readiness for future space flights was achieved with the first monolithic SHS interferometer [Harlander *et al.*, 2004], which as successfully flown on a free flying satellite in low-Earth-orbit for 2.5 years [Englert *et al.*, 2010a]. The monolithic interferometer is shown in Figure

2. Since all optical elements of the interferometer are optically contacted, it is extremely robust, light, and compact. It was integrated in an instrument called SHIMMER (Spatial Heterodyne Imager for Mesospheric Radicals) which was the primary payload of the STPSat-1 spacecraft. SHIMMER was designed to measure mesospheric hydroxyl using solar resonance fluorescence in the near infrared. A similar measurement was made from the Space Shuttle in the 1990s using a conventional grating spectrograph [Conway *et al.*, 1999]. However the SHS technique facilitated the reduction of size and power by about an order of magnitude. This allowed the SHS instrument to be integrated on a small satellite which led to about 2.5 years of data compared to two times two weeks from the Space Shuttle. The successful flight of SHIMMER demonstrated that SHS is suitable for long duration space flight applications and thus can be considered for future NASA missions.

For the DASH modification of SHS, the heritage is not only derived from SHIMMER on STPSat-1 [Englert *et al.*, 2010a], but also from the WINDII instrument [Shepherd *et al.*, 1993], which also used the interferogram phase at a finite optical path difference to determine wind speeds. Moreover, a monolithic DASH device was recently built, characterized and successfully used to measure thermospheric oxygen red line winds from the ground [Harlander *et al.*, 2010; Englert *et al.*, 2010b]. A photograph of the monolithic DASH interferometer is shown in Fig. 3.

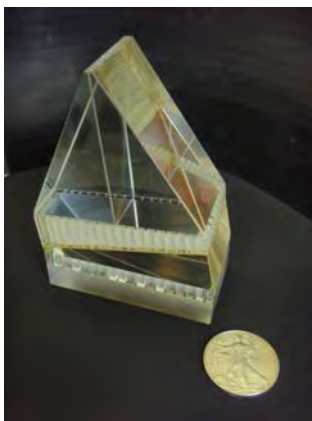


Figure 3: Monolithic, temperature compensated DASH interferometer designed for measuring thermospheric winds using the oxygen red line ($\lambda=630\text{nm}$). This particular design uses a Kösters prism beam splitter, which makes the two interferometer arms parallel and allows the use of a single field widening prism and grating for both arms. The spacers between the beamsplitter, field widening prism, and grating are made up of individual posts (“Stonehenge Design”) to minimize mechanical strain originating from the thermal expansion of the different glasses.

3. SHS can contribute significantly to future heliophysics missions

The advantages of SHS can make this approach the method of choice for of key heliophysics mission objectives. One example is the measurement of **thermospheric wind and temperature profiles**, which is absolutely necessary to address one of the priority objectives in the latest heliophysics roadmap [NASA, 2009]: “*Discover how winds and the composition of the upper atmosphere drive the electrical fields and chemical reactions that control the Earth’s ionosphere.*” Thermospheric wind measurements also have a long term societal impact, since they will facilitate significant improvements in the prediction of space weather effects that influence satellite operations, trajectory calculations (drag), and the propagation of electromagnetic waves through the ionosphere. Space weather has effects our daily life through GPS outages, power grid failures, or communication problems. Moreover, it is highly likely that commercial sub-orbital flights will become common in the near future [Foust, 2010]. These flights, with pilots, scientists, and tourists on board, will go to altitudes that are well within the thermosphere/ionosphere region and the providers will have a need for environmental data at all flight levels.

The above mentioned, modified SHS concept called **DASH is a prime candidate to measure thermospheric wind and temperature profiles** from low Earth orbit. The instrument would be **small and light enough (<35kg) to be included even in small missions**. For the thermosphere, typical measurements could be conducted **between 90-300km altitude with up to 5km altitude resolution** observing the oxygen red and green lines. With horizontal resolutions of a few hundred kilometers, one can achieve a wind speed **precision of less than 5m/s** and **temperature precision of less than 5K**, depending on the source brightness and altitude resolution. The **cost of such a thermospheric instrument is estimated to be in the \$20M range**.

DASH is also a candidate for measuring mesospheric winds and temperatures using molecular emissions, since it is able to simultaneously observe several spectral lines. A primary candidate for a suitable emission is the molecular oxygen band around 1.27 μ m. This effort is also highly relevant to the NASA 2009 Heliophysics Roadmap [NASA, 2009] which lists as one of its science investigations: “What is responsible for the dramatic variability of the ionosphere-thermosphere-mesosphere (ITM) region?” and “How do coupled middle and upper atmospheres respond to external drivers and to each other?” One major aspect of these questions is to understand how gravity waves propagate through, and break within the mesosphere. Since gravity wave propagation is intimately linked with the background wind flow [Holton, 1983], knowledge of the mesospheric wind field will help to answer these questions.

SHS also has the potential to contribute to solar physics investigations. For example, SHS is likely a suitable method that can meet the spatial imaging and spectral resolution requirements necessary for the first observation of the magnetic field in solar flares [Uri Feldman, private communication]. The instrument would measure the Zeeman splitting of an infrared solar emission line [Greenhouse et al., 1993; Feldman et al., 1984] within a flare, which gives direct information on the local magnetic field.

4. SHS Applications are Cross Cutting Between Heliospheric Themes and Beyond

We have demonstrated above with some examples that instruments based on SHS can contribute at a minimum to the two themes: “Atmosphere Ionosphere-Magnetosphere Interactions” and “Solar and Heliospheric Physics”. Other NASA objectives in planetary research and Earth science also have the potential to benefit significantly from the SHS technique. Two examples are the measurement of methane, water vapor, and carbon dioxide in the Martian atmosphere

[Englert *et al.*, 2008], and the use of a DASH interferometer as the back-end of a tropospheric wind LIDAR [Englert and Siskind, 2010].

The Mars application is a solar occultation observation with a very high resolution SHS spectrometer that measures CH₄, H₂O and CO₂ absorption lines in a narrow passband. Measuring CO₂ together with water vapor and methane is important so that the correct volume mixing ratios can be inferred in the presence of multi-scattering. The proposed, non field widened interferometer is no bigger than the case of a compact disk (CD), has a resolving power of 150,000 at 3.2μm, and uses immersion gratings.

DASH as a LIDAR back-end is a concept that was driven by the increasing need to acquire global tropospheric wind fields, which is a priority for the NASA Earth Science Division. The technology for this type of measurement is still under development in the US. The measurement of three-dimensional tropospheric winds is listed as a “Tier 3” class mission (i.e. “technology development underway”) in the last Decadal Survey and the 2010 NASA Science Plan [NASA, 2010].

References:

- Conway, R.R., M. H. Stevens, C.M. Brown, J.G. Cardon, S. E. Zasadil, and G.H. Mount (1999), Middle atmosphere high resolution spectrograph investigation, *J. Geophys. Res.*, 104(D13), 16,327–16,348, doi:10.1029/1998JD100036.
- Dawson O.R. and W.M. Harris (2009), “Tunable, all-reflective spatial heterodyne spectrometer for broadband spectral line studies in the visible and near-ultraviolet,” *Appl. Opt.* 48, 4227–4238.
- Englert, C.R., J.M. Harlander, J.G. Cardon, F.L. Roesler (2004), Correction of Phase Distortion in Spatial Heterodyne Spectroscopy, *Applied Optics*, 43, 6680-6687, 2004.
- Englert C.R., D.D. Babcock, J.M. Harlander (2007), Doppler Asymmetric Spatial Heterodyne Spectroscopy (DASH): Concept and Experimental Demonstration, *Applied Optics*, 46, 7297-7307.
- Englert C.R., J.M. Harlander, R. DeMajistre, M.H. Stevens (2008), Miniaturized Mars Methane Monitor, NASA Planetary Instrument Definition and Development Grant (proposal 08-PIDD08-0007).
- Englert C.R., D.D. Babcock, J.M. Harlander (2009), Spatial Heterodyne Spectroscopy for the long wave infrared: First measurements of broadband spectra, *Optical Engineering*, 48(10), 105602, doi: 10.1117/1.3250194.
- Englert, C. R., M. H. Stevens, D. E. Siskind, J. M. Harlander, and F. L. Roesler (2010a), Spatial Heterodyne Imager for Mesospheric Radicals on STPSat-1, *J. Geophys. Res.*, 115, D20306, doi:10.1029/2010JD014398.
- Englert C.R., J.M. Harlander, J.T. Emmert, D.D. Babcock, F.L. Roesler, Initial thermospheric wind measurements using a ground-based DASH interferometer (2010b), submitted to *Optics Express*
- Englert C.R. and D.E. Siskind (2010), Doppler Asymmetric Spatial Heterodyne Spectroscopy Light Detection and Ranging Receiver, US Patent Application.
- Feldman U., J.F. Seely, N.R. Sheeley, Jr., S. Suckewer, A.M. Title (1984), Magnetic field measurements in tokamak plasmas, *J. Appl. Phys.* 56, 2512-2518.
- Foust J. (2010), Suborbital research gets ready for liftoff, *The Space Review*, March 1, 2010, <http://www.thespacereview.com/article/1577/1>
- Greenhouse M.A., U. Feldman, H.A. Smith, M. Klapisch, A.K. Bhatia, A. Bar-Shalom (1993), Infrared coronal emission lines and the possibility of their laser emission in Sayfert nuclei, *Astrophys. Journal Supplement Series*, 88, 23-48.

“Spatial Heterodyne Spectroscopy: An Emerging Optical Technique for Heliophysics and Beyond”

Concept Paper For NRC Space Studies Board

October 2010

- Harlander, J. M., R. J. Reynolds, and F. L. Roesler (1992), Spatial heterodyne spectroscopy for the exploration of diffuse interstellar emission lines at far ultraviolet wavelengths, *Astrophys. J.*, 396, 730–740.
- Harlander J.M., F.L. Roesler, J.G. Cardon, C.R. Englert, and R.R. Conway (2002), SHIMMER: A Spatial Heterodyne Spectrometer for Remote Sensing of Earth's Middle Atmosphere, *Applied Optics*, 41, 1343-1352.
- Harlander J.M., F.L. Roesler, C.R. Englert, J.G. Cardon, R.R. Conway, C.M. Brown, J. Wimperis (2003), Robust monolithic ultraviolet interferometer for the SHIMMER instrument on STPSat-1, *Applied Optics*, 42, 2829-2834, 2003.
- Harlander J.M., F.L. Roesler, C.R. Englert, J.G. Cardon, J. Wimperis (2004), Spatial Heterodyne Spectroscopy for high resolution space based remote sensing, *Optics & Photonics News*, January 2004, Optical Society of America, 1047-6938/04/01/0046/6.
- Harlander J.M. C.R. Englert, D.D. Babcock, F.L. Roesler, Design and laboratory tests of a Doppler Asymmetric Spatial Heterodyne (DASH) interferometer for upper atmospheric wind and temperature observations (2010), submitted to *Optics Express*.
- Holton, J.R., The influence of gravity wave breaking on the general circulation of the middle atmosphere (1983). *J. Atmos. Sci.*, 40, 2497-2507.
- Lawlor J.E., Z.E. Labby, J.M. Harlander, and F.L. Roesler (2008), Broadband, high-resolution spatial heterodyne spectrometer, *Appl. Opt.* 47, 6371–6384.
- NASA (2009), Heliophysics, The solar and space physics of a new era, Recommended Roadmap for Science and Technology 2009-2030.
- NASA (2010) 2010 Science Plan for NASA’s Science Mission Directorate, NASA.
- Stephan SG, Chakrabarti S, Vickers J, Cook T, Cotton D. Interplanetary H Ly α observations from a sounding rocket. *Astrophys. J.* 2001; 559: 491-500.
- Watchorn S., Roesler F.L., Harlander J.M., Jaehnig K.P., Reynolds R.J., Sandersa W.T. (2001), Development of the Spatial Heterodyne Spectrometer for VUV remote sensing of the interstellar medium. *Proc. of SPIE*, 4498, 284-295.

Investigations of Plasmasphere Boundary Layer Processes in the Coupled Earth-Sun Geospace System

P. J. Erickson, A. J. Coster, D. Oberoi
MIT Haystack Observatory

November 12, 2010

Abstract

The paper proposes a focused attack on advancing knowledge of ionospheric structuring driven by M/I coupling in the critically important subauroral plasmasphere boundary layer (PBL), through use of multipoint measurement networks in both the American and Australian sectors. PBL processes are critical for atmosphere - magnetosphere - ionosphere understanding, require a micro to mesoscale observational perspective, and have important space weather consequences. The new Murchison Widefield Array and other observational networks in western Australia provide unprecedented opportunities for advancement of PBL layer physics and prediction of its space weather effects on total electron content, density gradients, and ionospheric irregularities. Support for PBL science analysis has a large leverage potential and would be multidisciplinary.

1 Introduction: Effects of Subauroral Magnetosphere - Ionosphere Coupling on Geospace

The coupled geospace system which governs plasma behavior in near-Earth space contains many examples of boundary layer phenomena where plasmas of distinctly different properties overlap. In particular, as pointed out by Carpenter and Lemaire (2004), these layers are sites of significant wave-particle energy and momentum exchange, cross-boundary energy transfer, and particle precipitation. Dynamics of these processes can be especially intense in the mid-latitude plasmasphere boundary layer (PBL), where the cool ($< 1\text{eV}$) dense ($400/\text{cm}^3$ or higher) plasma of the closed plasmasphere overlaps or lies in close proximity to hot (100 eV to 100 keV), tenuous ($1 / \text{cm}^3$ or less) plasmas characteristic of the plasmasheet and ring current. The regions characterized by these overlaps have multiple processes which significantly affect the energy and transport parameters of the plasma, creating conditions ideal for the development of plasma instabilities, severe density gradients, and large bulk and microscale flow channels driven by polarization electric fields. All of these processes can have a wide variety of spatial and temporal time scales which make observations and interpretations challenging, but which also shape the ultimate effects of PBL coupling processes on the ionospheric plasma state.

2 PBL Coupling Example: SAPS and SED

PBL processes are compelling scientifically, as they have ramifications for the dynamics of the entire coupled geospace system especially during disturbed times. As a specific example, Foster and Burke (2002) introduced the term Sub-Auroral Polarization Stream (SAPS) to describe a

persistent PBL region westward ionospheric convection velocity peak noted by several previous studies. This velocity peak is equatorwards and clearly separated from the auroral convection flow pattern during dusk sector geomagnetically disturbed intervals. The mid-latitude ionospheric flow is broad in general (3-5 degrees wide) with typical velocity magnitudes of several hundred meters per second or larger, associated with poleward electric field strengths > 20 mV/m. Within this broader SAPS channel, narrow regions ($< 1^\circ$) of intense, highly structured, variable velocities up to 2-3 km/s in L shell aligned directions have been reported as subauroral ion drifts (e.g. Anderson et al. (1991, 1993)), associated with electric fields > 100 mV/m (Foster and Erickson, 2000; Erickson et al., 2002). SAPS velocities are a key signature of magnetosphere/ionosphere coupling, and are magnetically associated with region 2 current closure systems as low energy ions flow downwards and close through the low conductivity subauroral ionosphere (Rich et al 1980; Senior and Blanc, 1984; Gkioulidou et al, 2009). The effects of SAPS on the PBL ionosphere and beyond are profound and include intense Farley-Buneman irregularities Erickson et al. (2002) along with mesoscale plasma density reconfigurations in latitude, longitude, and altitude [Foster et al, 2005]. Perhaps most importantly, the SAPS flows participate in the creation of severe electron density gradients associated with storm enhanced density (SED) plumes (Foster, 1993; Foster and Vo, 2002; Foster et al., 2007).

3 PBL Process Connections to Overlying Magnetospheric Configurations

During disturbance times, closure currents within the PBL are intimately connected with the asymmetric enhanced O^+ ring current in the Sun-Earth equatorial plane, driven by inner magnetosphere hot particle pressure gradients (Ridley and Liemohn, 2002). The M-I coupling processes driven by this configuration have important implications for both the ionosphere and plasmasphere. In particular, ionospheric SED features are connected magnetically to large dusk sector plasmasphere plumes observed by spacecraft such as IMAGE (Goldstein, 2006). Additionally, during moderate to severe geomagnetically disturbed intervals ($K_p \geq 3$), SAPS flows overlap the edge of the inner plasmasphere where SED features reside, and deliver significant levels of ionospheric flux westward to the cusp region (Foster et al., 2004). This flux provides an important stormtime source population for cusp ion upwelling (Andre and Yau, 1997), which has been identified as an important source of heavy, cold ions populating the radiation belts and mass loading the plasma sheet and ring current (e.g. Jordanova (2003)). The question of how much these ions contribute to overall magnetospheric configurations, polar cap plasma patches, and cross-polar cap transport is cross-cutting and of intense current interest, as the topic overlaps with several GEM focus groups and was the subject of a recent NASA Living With A Star Targeted Research and Technology science focus group.

Furthermore, an accurate inner magnetospheric specification has been shown recently to be crucial in estimating such fundamental parameters as the rate of change of the open/closed field line boundary, and indeed the overall 3D magnetospheric configuration during southward IMF turnings (Rae et al., 2010).

4 Space Weather Consequences

Understanding PBL effects has large consequences not only for advancing fundamentals of M-I coupling, but also for the larger technological community and the general public in the area of adverse space weather prediction. Since the beginning of this decade, mid-latitude space weather has been recognized to pose serious threats to space-based and land-based technological systems. An example of this was described by Doherty et al. (2004), when, during the 2003 October 29th and October 30th storms, there was no vertical navigation service in the FAA's Wide

Area Augmentation Service (WAAS) coverage area for 15 hours and 11.3 hours respectively. Ionospheric phenomena that can introduce space weather effects include exceptionally large values of total electron content (TEC), large spatial and temporal gradients in the TEC, and small-scale density irregularities. These features are all present during disturbance intervals within the PBL.

PBL effects also interfere with radio signals used for communication and navigation. For example, ionospheric irregularities can produce scintillation effects that interfere with a receiver's ability to track signals, and large spatial and temporal gradients in TEC can produce range errors in navigation systems (e.g. the Global Positioning System, GPS). Of particular importance are range errors that are introduced into differential GPS (DGPS) positioning solutions required by high precision real-time positioning applications such as in-vehicle navigation, marine positioning, emergency response, and railway control. A compelling need exists to understand PBL processes sufficiently well to incorporate them into next generation predictive, operationally focused models.

5 PBL Hemispheric Asymmetries: The Missing Observational Case

Advancing our knowledge of the important coupling processes mentioned above (and in particular including them in future accurate global models) requires examination of their morphologies in both hemispheres in order to correctly model the electrodynamic situation along the entire magnetic field line. To date, this has been quite difficult to accomplish using ground sites, which complement satellite data well by avoiding time/space ambiguities. In fact, almost all of our knowledge on mid-latitude M/I coupling effects in the ionosphere (such as SED) is derived from northern hemisphere observations. For instance, many of the significant mid-latitude space weather events that have been reported are from North American observations of SED (Foster, 1993; Foster and Vo, 2002) along with associated large TEC gradients (Foster and Coster, 2007) and small-scale irregularities (Ledvina et al., 2002). One reason for the focus on North America is that during the last solar maximum (2000-2003) only the US had a dense enough network of GPS receivers to fully view SED plumes in a mesoscale context.

During the past few years, however, the community has begun to address ground based studies in both hemispheres. These studies are of relatively few events and are hampered by the much smaller data coverage for observing such effects in the southern hemisphere. Specifically, there have been isolated GPS TEC studies focused on the southern hemisphere (e.g. Yizengaw et al. (2005)). These efforts have already yielded interesting PBL morphologies. For instance, from analysis of current (and very limited) event-based data sets, several aspects of the SED feature appear to be magnetically conjugate (Foster and Rideout, 2007). In particular, the entry of the SED plume into the polar cap near noon is observed in both hemispheres. Foster and Rideout (2007) interpreted this conjugacy as an indication that storm-time electric fields are at least partially responsible for the generation of SED. However, a prominent feature of SED that does not appear to be conjugate is the magnitude of TEC enhancement at the base of the SED plume. This TEC enhancement appears to exhibit localized and longitude-dependent features, perhaps due to the offset of the geographic and geomagnetic latitudes in the American sector. These topics to date are just beginning to be explored.

A tantalizing satellite-based glimpse into hemispheric differences within the PBL has very recently been provided by preliminary magnetometer data from the AMPERE satellite constellation, showing field-aligned subauroral currents which seem to have significant differences between hemispheres (B. Anderson, private communication). These differences would lead to large differences in closure currents and associated ionospheric structuring electric fields. However, very little is known about the effects of these current asymmetries on the ionosphere, or

how they interact with changing integrated electron densities in each hemisphere in a feedback manner.

6 Developing An Integrated, Dual Hemisphere Observational Capability

These large uncertainties in our knowledge of PBL effects in conjugate hemispheres suggest that significant advances could be made if monitoring were available of key ionospheric parameters during mid-latitude space weather events at a ground location in the southern hemisphere that has an offset ($\geq 10^\circ$) between the geographic and geomagnetic latitudes similar to that found in the eastern part of North America. The western portion of Australia has these properties, and within the last decade ground based instrumentation efforts are underway. Figure 1 shows a map of current or planned instrument deployments. These stations naturally complement more global satellite overflight data. The emerging network of multiple, independent science instruments in both hemispheres are a natural starting point for a distributed array of scientific instruments (DASI) approach to providing both broad coverage and fine-scale spatial/temporal resolution of upper atmospheric processes in the PBL. This distributed approach was called out specifically in the 2002 NRC decadal survey report as crucial to understanding the coupled AIM system.



Figure 1: Murchison Widefield Array location and future distribution of GPS receivers in Australia.

6.1 Ionospheric Science Using Low Frequency Arrays in Australia

As a centerpiece of Australian sector efforts, a new observational system, the Murchison Widefield Array (-26.7 latitude, 116.7 longitude, and -38 magnetic latitude) (Lonsdale et al., 2009) provides a particularly strong opportunity. The MWA is a new low frequency array being built in western Australia as a collaborative international project between the U. S., Australia, and India with independent funds from NSF, AFOSR, and others to develop powerful new capabilities for radio astronomy and heliospheric science. The MWA is fundamentally a radio interferometer that will image large fields of view with unparalleled fidelity and precision at frequencies in the VHF range (meter scale wavelengths). The MWA will consist of 8192 dual-polarization dipole

antennas arranged as 512 "tiles", each being a 4x4 array of dual-polarization dipoles. Currently, 32 tiles are fielded with full array buildout anticipated in mid 2012. Figure 6.1 shows the MWA aerial photograph and some deployed antennas.



Figure 2: Aerial photograph of the MWA site in Western Australia, showing the first 32 deployed tiles and (inset) a close up of an antenna tile consisting of 16 dipoles placed on a ground screen (Lonsdale et al., 2009).

Low frequency arrays have recently received a strong endorsement from the Astro-2010 decadal survey, which bodes well for robust and continued community support. Most importantly, the MWA is a cross-disciplinary instrument and offers powerful new opportunities for ionospheric science. In particular, the low frequency heliospheric and radio astronomy missions require a very accurate removal of ionospheric effects, which as a byproduct specifies aspects of the electron density state extremely well. For instance, Faraday rotation measurements in the heliosphere and beyond are first and second order effects compared with the order-unity effects of the ionosphere, and this implies accurate relative ionospheric/plasmaspheric TEC for successful removal. This will be accomplished through continuous monitoring of ~ 300 calibrator radio sources by the wide field array. This technique for the MWA ionospheric calibration product will produce maps of TEC spatial gradients at an average spatial resolution of ~ 5 km across an ionospheric region of an approximate 175 km square at 350 km altitude. These measurements will be made on short temporal (~ 10 sec) scales and with very fine amplitude resolution, < 0.01 TEC units. For reference, GPS can only measure relative TEC to ~ 0.03 TEC units. Figure 3 shows an example of the precise differential TEC product which will be available from the MWA calibration output.

MWA observations are also uniquely suited for the exploration of ionospheric irregularities and scintillation. When fully built out, the 512 tile MWA system will provide multifrequency VHF scintillation measurements at multiple positions over this same 175 km square region in the ionosphere. Fine scale investigations of ionospheric structure during selected VHF satellite beacon overflights are also possible, along with full Faraday rotation capability in the case of the DMSP F15 linearly polarized beacon.

6.2 Regional Context

Regional synergistic, DASI-style observational networks are becoming available at the same time to complement the MWAs ionospheric capabilities. In particular, during the upcoming

Differential TEC derived from Measured Offsets of Radio Sources

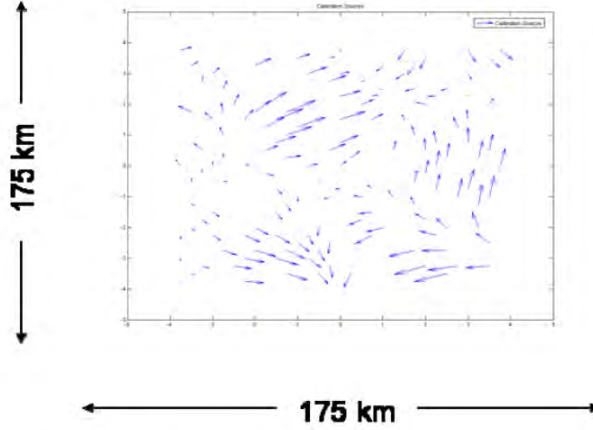


Figure 3: Example of differential TEC map (relative scale) estimated from measured offset in position of a large number of radio sources. Arrows represent magnitude and direction of the TEC gradient. The field of view of the full MWA at 350 km altitude is approximately 175 km x 175 km. In this example 300 radio sources (the approximate number for MWA operations) were used as calibrators. (S. Doleman, private communication).

rise of the next solar cycle, we will be able to monitor mid-latitude space weather effects at this southern hemisphere location at a very fine level of detail through the combination of ionospheric observations from a new network of GPS receivers in Australia, from three specially equipped colocated GPS scintillation receivers at the MWA site, and from the MWA itself. These can be combined with detailed snapshot passes from the DMSP constellation (840 km altitude) and DEMETER (710 km altitude), which provide electric field/ion drift data and particle precipitation information at a range of appropriate energies.

7 New Opportunities for M/I Coupling and Irregularity Science

These capabilities, when taken as a whole, provide the ionospheric and magnetospheric community with an completely unprecedented and unique opportunity to flesh out knowledge of crucially important M-I coupling effects. By comparing and contrasting mid-latitude space weather events between the two hemispheres at the longitudes of maximum pole offset, the seasonal and geomagnetic dependences of SED can be characterized and separated from other factors that may impact the formation of SED; for example, in the North American sector, the influence of the South Atlantic Anomaly (SAA).

Science possibilities for Australian/North American coordinated studies include:

1. compare and contrast the development of SED in two hemispheres at locations where the geomagnetic and geographic offset is maximum;
2. observe TEC gradients associated with SED in Australia with considerable fidelity, enhancing our understanding of this phenomenon both locally and globally;

3. associate plasma irregularities with features such as large TEC gradients associated with SED;
4. examine how energy cascades from larger to smaller scale sizes within the plasma;
5. monitor the speed, direction, and amplitude of ionospheric irregularities as they cross the subauroral hemisphere;
6. identify background irregularity conditions associated with VHF and L-band scintillation.

All of these capabilities will allow us to significantly advance the state of knowledge of both ionospheric structuring and space weather.

8 Cost-Effective, Ground Based Synergistic Measurements

The M-I science community can take advantage of the ongoing and future Australian sector measurements with the MWA and associated instruments in an extremely cost effective manner, since deployments of ground based hardware and data systems are already underway. For instance, an extensive array of geodetic quality receivers is in the process of being installed by the Australian government through Geoscience Australia. This data is freely available on the web. Three GPS scintillation receivers capable of measuring L-band scintillation statistics have been provided by Dr. Keith Groves (AFRL) and are currently in operation at the MWA site. Observations from other ground-based (e.g. ionosondes) and satellite based platforms, including COSMIC, DMSP, and JASON, can be utilized to provide additional input to the characterization of the background ionospheric plasma and electrodynamic conditions in Western Australia.

9 Needed Support

Since instruments in the Australian sector are currently being fielded, fully realizing the capabilities available for PBL science as described above most properly centers on vigorous support for dedicated scientific analysis. This is a very minor part (or nonexistent) in current observational network operations and management plans. Increased community support for science activities has the advantage of a large leverage potential.

There is also a pressing need to have PBL data from diverse sources easily available in a virtual observatory or other distributed database environment in order to enable the kind of synthesis and system level analysis which is crucial to identifying fundamental PBL processes. Common user interfaces are also highly desirable. These kind of capabilities exist in systems such as Madrigal (developed at MIT), but need additional effort and support. These community investments would lead to an ideal coordinated science tool, when combined with advanced visualization environments such as the widely available Google Earth interface.

References

- Anderson, P. C., Hanson, W. B., Heelis, R. A., Craven, J. D., Baker, D. N., and Frank, L. A. (1993). A proposed production model of rapid subauroral ion drifts and their relationship to substorm evolution. *Journal of Geophysical Research*, 98:6069.
- Anderson, P. C., Heelis, R. A., and Hanson, W. B. (1991). The ionospheric signatures of rapid subauroral ion drifts. *Journal of Geophysical Research*, 96:5785.
- Andre, M. and Yau, A. (1997). Theories and observations of ion energization and outflow in the high latitude magnetosphere. *Space Science Reviews*, 80:27.

- Carpenter, D. and Lemaire, J. (2004). The plasmasphere boundary layer. *Annales Geophysicae*, 22:4291–4298.
- Doherty, P., Coster, A., and Murtagh, W. (2004). Space weather effects of October–November 2003. *GPS Solutions*, 8:267–271.
- Erickson, P., Foster, J., and Holt, J. (2002). Inferred electric field variability in the polarization jet from Millstone Hill E region coherent scatter observations. *Radio Science*, 37(2):1027.
- Foster, J. (1993). Storm time plasma transport at middle and high latitudes. *Journal of Geophysical Research*, 98(A2):1675–1689.
- Foster, J. and Burke, W. (2002). SAPS: A new characterization for sub-auroral electric fields. *EOS*, 83:393–394.
- Foster, J., Coster, A., Erickson, P., Rich, F., and Sandel, B. (2004). Stormtime observations of the flux of plasmaspheric ions to the dayside cusp/magnetopause. *Geophys. Res. Lett.*, 31:L08809.
- Foster, J. and Rideout, W. (2007). Storm enhanced density: magnetic conjugacy effects. *Annales Geophysicae*, 25:1791–1799.
- Foster, J., Rideout, W., Sandel, B., and Forrester, W. (2007). On the relationship of saps to storm-enhanced density. *Journal of Atmospheric and Solar-Terrestrial Physics*, 69:303–313.
- Foster, J. C. and Coster, A. J. (2007). Conjugate localized enhancement of total electron content at low latitudes in the american sector. *Journal of Atmospheric and Solar-Terrestrial Physics*, 69:1241.
- Foster, J. C. and Erickson, P. J. (2000). Simultaneous observations of E-region coherent backscatter and electric field amplitude at F-region heights with the Millstone Hill UHF radar. *Geophysical Research Letters*, 27:3177.
- Foster, J. C. and Vo, H. B. (2002). Average characteristics and activity dependence of the subauroral polarization stream. *Journal of Geophysical Research*, 107:1475.
- Goldstein, J. (2006). Plasmasphere response: Tutorial and review of recent imaging results. *Space Science Reviews*, 124:203–216.
- Jordanova, V. K. (2003). New insights on geomagnetic storms from model simulations using multi-spacecraft data. *Space Science Reviews*, 107:157.
- Ledvina, B. M., Makela, J. J., and Kintner, P. M. (2002). First observations of intense GPS L1 amplitude scintillations at midlatitude. *Geophysical Research Letters*, 29:4.
- Lonsdale, C. J., Cappallo, R. J., Morales, M. F., Briggs, F. H., Benkevitch, L., Bowman, J. D., Bunton, J. D., Burns, S., Corey, B. E., Desouza, L., Doleman, S. S., Derome, M., Deshpande, A., Gopala, M. R., Greenhill, L. J., Herne, D. E., Hewitt, J. N., Kamini, P. A., Kasper, J. C., Kincaid, B. B., Kocz, J., Kowald, E., Kratzenberg, E., Kumar, D., Lynch, M. J., Madhavi, S., Matejek, M., Mitchell, D. A., Morgan, E., Oberoi, D., Ord, S., Pathikulangara, J., Prabu, T., Rogers, A., Roshi, A., Salah, J. E., Sault, R. J., Shankar, N. U., Srivani, K. S., Stevens, J., Tingay, S., Vaccarella, A., Waterson, M., Wayth, R. B., Webster, R. L., Whitney, A. R., Williams, A., and Williams, C. (2009). The Murchison Widefield Array: Design overview. *Proceedings of the IEEE*, 97:1497.

- Rae, I. J., Kabin, K., Lu, J. Y., Rankin, R., Milan, S. E., Fenrich, F. R., Watt, C. E. J., Zhang, J.-C., Ridley, A. J., Gombosi, T. I., Clauer, C. R., Tóth, G., and DeZeeuw, D. L. (2010). Comparison of the open-closed separatrix in a global magnetospheric simulation with observations: The role of the ring current. *Journal of Geophysical Research*, 115:08216.
- Ridley, A. J. and Liemohn, M. W. (2002). A model-derived storm time asymmetric ring current driven electric field description. *Journal of Geophysical Research*, 107:1151.
- Yizengaw, E., Dyson, P. L., Essex, E. A., and Moldwin, M. B. (2005). Ionosphere dynamics over the southern hemisphere during the 31 March 2001 severe magnetic storm using multi-instrument measurement data. *Annales Geophysicae*, 23:707.

1. Science Concept

There is growing realization of the extent to which the sub-systems (the spheres) of terrestrial space science and the processes therein are interconnected, such that the composite of these regions, Geospace, acts as a coupled, complex system whose pieces and properties should not and cannot be studied in isolation. Whole-atmosphere models extending from interactions with the oceans and land masses below to the magnetosphere and solar drivers above are examples of the effort to understand the workings of the system as a whole. Individual features and research topics contribute to the composite picture, but a larger-perspective point of view is needed to begin to appreciate how these components work together to shape our space environment, its dynamics, and variability. At present, this larger-picture view of Geospace is extremely data deficient. Individual projects and missions address separate regions, topics and processes. Flexible, coordinated, cost-effective observations of Earth's space environment are needed to provide the wide range of measurements with the spatial distribution and temporal resolution needed to support Geospace monitoring, research, modeling, and predictive activities. An initiative to undertake a phased implementation of a program to field and coordinate Distributed Arrays of Scientific Instruments (DASI) is a much needed component of the Space Science research and development plan for the coming decade.

This white paper does not attempt to describe the final global network of next-generation instruments and data infrastructure which is the goal of the DASI initiative, but strives to express the need to take the first steps to structure and coordinate activities in the planning and utilization of space research instrument arrays in support of the broad needs of Geospace system research and to provide the real-time data and analysis capabilities needed for Space Weather modeling and predictive activities. DASI is already underway, with many regional arrays in place and operational. An initial effort to coordinate activities, facilitate data exchange and utilization, and to expand the data-user community within the North American region is suggested as a Phase I (developmental) implementation of DASI. Development of miniaturized instruments and efficient deployment strategies would lead to extended array coverage and capabilities while addressing cost efficiency - potentially during a DASI Phase II which would take increased funding and commitment. Incorporation of observations from small-satellite constellations, from extended ocean and remote-location instrument platforms, and strong international participation would be a goal for the implementation of DASI on the global scale needed to address the requirements of studying, understanding, and predicting Earth's coupled Geospace system.

2. DASI: Advancing space science and operational capabilities

The enhanced observational capability and ready data availability provided by DASI will enhance and extend the scientific activities of the owners and users of the contributing arrays by providing access to correlated data and contributing phenomena observed over an extended (global) spatial scale, providing contextual information with which to interpret localized phenomena (e.g. the system history and context of specific RBSP observations). Wide-ranging, readily-available data will support system-level research projects, extending the scope of research programs to incorporate Geospace system aspects and influences. For operational capabilities, the DASI concept addresses directly the needs of Space Weather research and operational products to have access to

global real-time observations of atmospheric and space characteristics and phenomena. DASI is a part of providing an integrated space situational awareness capability.

3. Cost estimates

Adopting a phased approach to DASI implementation would avoid the unrealistic task of trying to design, cost, and coordinate the needed global observational capability before further developing concepts, strategies, and instrument anointment based on the arrays and infrastructure already in place. Phase I - enhancing coordination and communication among existing regional arrays and users - would come at flexible, lower cost (and be classified as a Small Project) since the size and pace of this prototype regional program could be scaled to the amount of support available. A second Phase of enhanced instrument development and deployment would take appropriate levels of targeted funding. As the need and utility for DASI instrumentation becomes apparent as the program gains momentum, the scope of a global DASI Phase III and its final costs would scale with the priority given to such a program by the community.

It is worth noting that the NSF CEDAR research program began 25 years ago as a no-new-funding Phase 0 effort to coordinate separate research activities in a way which demonstrated the value-added having a common focus and goal. CEDAR Phases I, II, and III followed, each with enhanced levels of funding, community participation, and scientific contribution.

4. DASI: Specific Review Criteria:

a. Identified as a high priority in previous studies

A small instrument distributed ground-based network (DASI) was recommended as a Priority small program in the previous Solar and Space Physics Decadal Survey. Following the Survey recommendation, an NRC community workshop and report were organized to further explore the DASI concept (ref). The DASI report emphasized three recurrent themes: 1) the need to address Geospace as a system, 2) the need for real-time observations, and 3) the insufficiency of current observations (observational space physics is data starved).

b. Contribution to Panel themes;

The Scope of the Survey includes the physics associated with the magnetospheres, ionospheres, thermospheres, mesospheres, and upper atmospheres of the Earth and other planets. In many instances, these areas have been treated as separate 'spheres' and missions and programs have embraced a focus too narrow to explore the inter-sphere connections which are inherent in the Geospace system. By providing a growing observational framework and access to data from instruments addressing a wide spectrum of Geospace characteristics, DASI contributes to the better understanding of both the sub-system processes and the working of the Geospace system as a whole.

c. Contributes to important scientific questions

DASI by providing for a wide spectrum of inter-related geospace observations with good spatial and temporal resolution is a key element in addressing the coupled-system aspects of Geospace. The inter-relationship of thermosphere effects and variability to the characteristics and processes in Earth's magnetosphere are largely unexplored and thus not incorporated in descriptive or pre-

dictive models because of the lack of coordinated observations which could be used to address such system effects. An adequate observational base is needed to advance our understanding of all aspects of Geospace connections.

d. Contributes to applications and/or policy making

The development of coordinated, readily-accessible network of distributed geospace observations directly supports the needs of Space Weather real-time monitoring and prediction activities. An operational DASI network will provide a robust framework for supporting the Space Weather requirements. This is particularly evident in that the specific processes and effects (e.g. ionospheric scintillations) of interest in the Space Weather arena are manifestations of the complex, coupled system which is Geospace, requiring distributed real-time observations for both monitoring and understanding.

e. Complements other observational systems or programs

The DASI concept- facilitating the acquisition, exchange, and multi-user access to distributed real-time geospace observations - directly complements existing and planned ground and space-based missions exploring the processes and interconnections of Earth's space and atmospheric environment. By combining resources addressing characteristics and processes in the different spheres of the geospace system, DASI supports the extension of the scientific value of such systems and missions by providing a ready means for exploring the linkages among the sub-systems of geospace. For example, RBSP (Radiation Belt Storm Probes) will provide in-situ probes of Earth's energetic particle environment. Processes investigated there depend on ambient conditions (e.g. cold plasma often of ionospheric origin and M-I coupling effects on the plasma waves which energize and scatter the radiation belt particles), which require a system understanding of conditions and processes occurring in the underlying ionosphere and thermosphere.

f. Is affordable (cost-benefit)

Initial aspects of DASI instruments and infrastructure already exist and are in operation supporting regional studies of aspects of the Geospace system. A cost-effective beginning for a more-extensive DASI program will be in supporting the scientific coordination and development of the infrastructure which will provide ready access to such existing data resources to an extended user community. Implementation of such efforts in a region which already has considerable observational resources and data dissemination networks (e.g. the North American sector) would make infrastructure development, testing, and initial coordination activities less costly. Major space research instruments including incoherent scatter radar facilities, portions of the high-latitude and new mid-latitude SuperDARN HF radar networks, THEMIS GBO optical and magnetic observatories, as well as numerous small -instrument arrays are in place in this region. As a coordinated system perspective of our space science activities gains emphasis, development of our existing instrument and coordination infrastructure capabilities will produce highly cost-effective returns.

g. Readiness

For a North American regional DASI, major space research instruments including incoherent scatter radar facilities, portions of the high-latitude and new mid-latitude SuperDARN HF radar networks, THEMIS GBO optical and magnetic observatories, as well as numerous small -instrument arrays are in place in this region. Joint US and Canadian activities (AMISR radars at Reso-

lute Bay, THEMIS GBO networks as well as a large body of active researchers addressing aspects of ground and space-based geospace research in the region speak to the readiness of this region to undertake an initial deployment of a regional DASI program.

The new AMISR ISRS (supported in past Survey) have been put into operation and their numbers are increasing with phased-array radars in place in Alaska and Resolute Bay with others underway in Argentina and a 2nd installation at Resolute Bay and proposed for Antarctica. Combined with existing dish ISRs at Sondrestrom, Millstone Hill, and Arecibo, the radars, whose operations are coordinated and data distributed to a broad user community via MADRIGAL, form a large-instrument backbone for a regional distributed-instrument capability for obtaining detailed real-time observations needed to address geospace science and space weather objectives.

h. National and international plans and activities.

Within the NSF research community, the GEO Division has issued a GeoVision plan which promotes the investigation of the Earth environment as a coupled system, with both the understanding of system processes and sub-system interconnections, and human-system interactions and effects (e.g. Space Weather and aspects of Climate Change) being important goals for GEO research. The DASI initiative addresses these cross-coupled objectives nicely in its objective to make space science observations readily available to the more-broad research, applications, and education communities. On a more-focused scale, a significant portion of ground-based space science research activities are supported by the Upper Atmosphere Research Section of the GEO Directorate's Division of Atmospheric and Geospace Sciences (AGS). Here, the CEDAR (Aeronomy/ITM) research community has put forward a new Strategic Plan which puts an emphasis on addressing the broader aspects of CEDAR research within the framework of Geospace system science. DASI, both with its instrument arrays and coordination and dissemination of data, can be envisioned as a major observational capability supporting the implementation of the new CEDAR plan.

References

National Research Council (NRC). 2003. *The Sun to the Earth - and Beyond: A Research Strategy in Solar and Space Physics*, The National Academies Press, Washington, D. C.

National Research Council (NRC). 2006. *DASI - Distributed Arrays of Small Instruments for Solar-Terrestrial Research: A Workshop Report*, The National Academies Press, Washington, D. C.

Solar Forcing of the Thermosphere and Ionosphere from below: Coupling via Neutral Wave Dynamics

Dave Fritts and Ruth Lieberman, NorthWest Research Assoc., and Bela Fejer, Utah State Univ.
Executive Summary

Neutral atmosphere – ionosphere coupling represents a key link in the response of the terrestrial environment to solar forcing. Neutral motions, primarily waves arising at lower altitudes or in the auroral zone, enter the thermosphere and ionosphere, where they achieve large amplitudes and play major roles in accounting for the neutral and plasma structure and variability. Key questions include:

1. How does the global spectrum of thermospheric waves evolve with altitude, time, and solar forcing?
2. How do waves contribute to the mean structure, dynamics, and electrodynamics of the thermosphere and ionosphere?
3. How do waves drive thermospheric and ionospheric variability?

Understanding these coupling processes in the terrestrial environment is essential to improved Space Weather prediction, mitigating communication and navigation impacts of ionospheric scintillations and plasma bubbles, predicting orbits and orbital decay for spacecraft and space debris, and gaining insights into similar coupled dynamics in other planetary atmospheres.

Such a mission requires global measurements and could be addressed, depending on instrumentation, with one or several satellite platforms. The instrument capabilities exist and such a mission could be performed within the scope of an EX or MIDEX Explorer budget.

1. Introduction

Observations, modeling, and theoretical studies over ~50 years have suggested that neutral atmosphere influences account for many of the most prominent features of the ionosphere and its spatial and temporal variability. These include the E and F region dynamos, large-scale polarization electric fields, currents, plasma drifts, and the equatorial ionospheric anomaly (*e.g.*, Crain et al. 1993; Heelis 2004; England et al. 2006; Immel et al. 2006). Neutral atmosphere influences have also been implicated in smaller-scale dynamics, *i.e.*, traveling ionospheric disturbances, equatorial plasma bubble seeding, and mid-latitude spread F, both under geomagnetically quiet and active conditions (McClure et al. 1998; Nicolls et al. 2004). Indeed, recent studies have provided compelling evidence that many of these ionospheric dynamics are influenced directly by neutral dynamics and “weather” arising in the lower atmosphere.

These ionospheric influences accompany various wave motions that arise due to solar energy inputs in the lower and middle atmosphere (~0-10 km and ~10-100 km, respectively) and in the thermosphere. Waves represent indirect rather than direct solar forcing of the thermosphere and ionosphere (TI) at higher altitudes. They induce very large neutral wind and density variations extending well into the TI due to the approximately exponential growth of wave amplitudes with decreasing mean density. These induce, in turn, large plasma density perturbations, plasma drifts, currents, and polarization electric fields that account for many of the major features of the ionosphere. They also appear to contribute to plasma instabilities and their finite-amplitude manifestations that can have significant and adverse effects on human activities. Indeed, these indirect effects of solar forcing may contribute enormous neutral and plasma variability in the TI specifically ***because*** their sources are at lower altitudes, where neutral densities are as much as 10^{10} larger than in the F layer. This allows waves having apparently insignificant amplitudes at their source altitudes to nevertheless achieve large amplitudes and effects in the rarified TI, especially under quiet geomagnetic conditions (*e.g.*, Fejer and Schierless 2001).

The neutral waves that account for the transport of solar energy from lower altitudes into the TI include solar tides, planetary waves (PWs), and internal gravity waves (GWs). Migrating and non-migrating solar tides are global-scale waves that are excited by the diurnal cycle of solar radiation absorption in the troposphere and stratosphere, the longitudinal variations of this absorption accompanying the distributions of deep tropical convection and stratospheric ozone, and in situ forcing in the thermosphere (*e.g.*, Hagan and Forbes 2002; Oberheide and Forbes 2008a). Nonlinear interactions among the tides, PWs, and GWs where wave amplitudes are large contribute additional modes and variability (*e.g.*, Teitelbaum and Vial 1991; Ortland and Alexander 2006; Fritts et al. 2010). These tidal forcing mechanisms yield a complex superposition of tidal modes that are apparent in the mesosphere and lower thermosphere (MLT) and which also extend into the E and F regions (*e.g.*, Forbes et al. 2009; Hagan et al. 2009; Oberheide et al. 2009; Talaat and Lieberman 2010). Tides in the MLT and above also appear to provide the major modulation of GW propagation and fluxes of energy and momentum into the TI, the dominant forcing of the E and F region dynamos, large perturbations of minor species and plasma, and the pre-conditioning of the ionosphere for various plasma instabilities (*e.g.*, Fritts and Vincent 1987; Immel et al. 2006; Häusler et al. 2007, 2010; Lühr et al. 2007, 2008; Kudeki et al. 2007; Vadas 2007; Fritts and Vadas 2008; Oberheide and Forbes 2008b).

PWs arise in the lower and middle atmosphere through a variety of processes. Large-scale topography and land-sea temperature contrasts excite global-scale, quasi-stationary PWs that typically extend well into the stratosphere. Occasionally these lead to strong PW/mean-flow interactions during winter known as sudden stratospheric warmings (SSWs), primarily in the northern hemisphere, which have influences extending into the MLT and TI (*e.g.*, Liu and Roble 2002; Goncharenko and Zhang 2008; Fejer et al. 2010). Other PWs arise due to barotropic and/or baroclinic instabilities of the large-scale circulation and thermal structure and exhibit spatial structures and periods resembling those expected for the various normal modes of the atmosphere. These motions are also seen to impose variability in the TI and modulation of plasma dynamics, though the mechanisms by which this coupling occurs remain unknown at present (*e.g.*, Forbes and Leveroni 1992; Pancheva et al. 2006, 2008; Fejer et al. 2010).

GWs have many sources in the lower and middle atmosphere. These include topography, convection, jet streams and frontal systems, nonlinear interactions and instability processes, and local body forces arising due to GW dissipation and momentum deposition. Their importance in the MLT and TI derives from their amplitude growth with altitude, their strong interactions with, and modulation by, mean winds, tides, and PWs, and their significant transport of energy and momentum (*e.g.*, Fritts and Alexander 2003; Vadas 2007; Fritts and Vadas 2008; Fritts et al. 2008). Their large vertical group velocities also allow them to achieve high altitudes despite their relatively smaller spatial scales. GW effects in the TI appear to include large neutral density and wind perturbations, corresponding large plasma densities, gradients, drifts, and polarization electric fields, significant neutral accelerations accompanying their momentum transport and deposition, and potential seeding of plasma dynamics and instabilities, especially equatorial plasma bubbles that may penetrate to altitudes of ~ 1000 km or higher.

Indirect solar forcing is recognized to be critical in defining TI structure, energetics, and variability, but our understanding of these dynamics and their effects is very primitive at present. New global observations are paramount to advance the understanding of TI dynamics and their implications for human activities. The remainder of this white paper provides examples of neutral wave influences in the TI and a brief overview of satellite mission requirements that would enable a quantitative understanding of these neutral atmosphere-ionosphere dynamics.

2. Examples of Neutral Atmosphere Influences in the TI

a. Tidal influences

Both non-migrating and migrating tides are expected to penetrate to high altitudes in the TI and to have significant dynamic and electrodynamic effects. The diurnal eastward-propagating wavenumber 3 (DE3) is seen to make the dominant non-migrating component contribution at high altitudes, where it is seen to dominate the neutral wind field near the dip equator (see Figure 1) and to induce significant longitudinal perturbations in the equatorial anomaly (Immel et al. 2006). Importantly, however, the maximum magnitude of this component is only ~ 10 m/s. Non-migrating tide signatures are also seen in exospheric temperatures inferred from CHAMP accelerometer data (see Figure 2) and suggest significant neutral and associated plasma densities at these altitudes.

Migrating tides were anticipated to be significant in the TI, based on TIME GCM simulations and limited ground-based measurements, but less attention has been paid to defining their component amplitudes to date. These modes may also be estimated from CHAMP wind data averaged over a yaw cycle, however (see Figure 3), and the results of such an assessment show diurnal and semidiurnal amplitudes that are substantially larger than the corresponding non-migrating modes (despite averaging over a yaw cycle), with terdiurnal components having comparable amplitudes.

Thus we might expect that the migrating tides in the E and F layers and above may have comparable or larger effects to those already highlighted for the non-migrating modes in previous studies. New, comprehensive, global observations of these tides in the TI are needed to quantify these tidal contribution to TI dynamics and energetics.

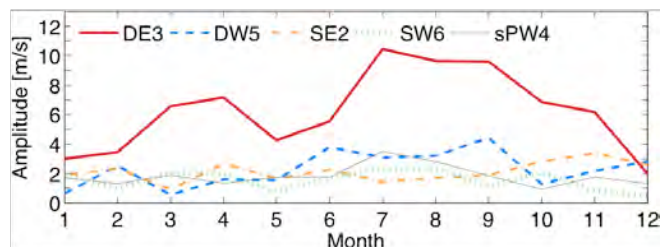


Figure 1. Amplitudes of the diurnal (DE3, DW5), semidiurnal (SE2, SW6), and stationary (sPW4) components contributing to the wave-4 structure in the zonal wind along the dip equator (Häusler and Lühr 2009).

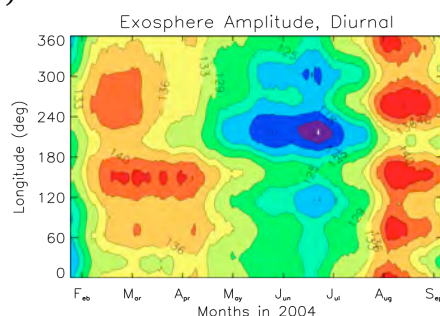


Figure 2. Exospheric diurnal tidal temperatures from the CHAMP accelerometer (Forbes et al. 2009).

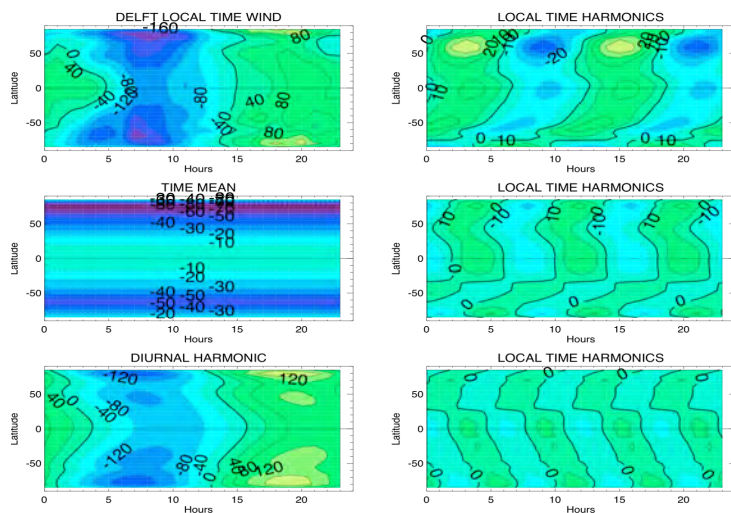


Figure 3. Total zonal wind with LT and latitude (top left), mean zonal wind (middle left), and diurnal, semidiurnal, terdiurnal, and 6-hour migrating tides inferred from DELFT CHAMP winds averaged over a yaw cycle. Note that the three dominant migrating modes are all larger than the non-migrating modes (image by R. Lieberman).

b. PW influences

PWs likewise exhibit interesting, but not understood, influences in the ionosphere in various data sets. Figure 4 shows an example of vertical plasma drifts measured by the Jicamarca radar that exhibit 2-day and 16-day wave periodicities that become much more pronounced during a northern hemisphere SSW.

A second example shown in Figure 5 reveals a correlation between electrojet current and North Pole temperature at 32 km also occurring during the 2002-3 SSW event. These results suggest that even slowly-varying, global-scale motions have pronounced responses in the E and F layers that represent coupling from the neutral atmosphere into the ionosphere accompanying waves propagating upward from lower altitudes. The physical processes accounting for these responses, and their implications for ionospheric behavior and variability, are completely unknown at present, however, and new global measurements are required to define these influences quantitatively.

c. GW influences

As seen above for tides, GW perturbations in winds, temperatures, and electron densities are directly observable in many data sets. As a result, it has been possible to compare observed GW spatial scales and frequencies with those predicted based simply on viscous propagation theory (Vadas and Fritts 2005). The results of such comparisons show generally good agreement, despite the linear description of GW structure and propagation inherent in ray tracing theory (*e.g.*, Earle et al. 2008; Vadas and Nicolls 2008, 2009). An example of the variations of GW vertical wavelengths with altitude seen in electron densities and observed with the MU radar in Japan is shown in Figure 6. This reveals an approximately exponential increase in vertical wavelength with altitude that is largely consistent with theory (Vadas 2007). A second example using the Poker Flat Incoherent Scatter Radar (PFISR) in Alaska for a multiple-beam experiment is shown in Figure 7 and

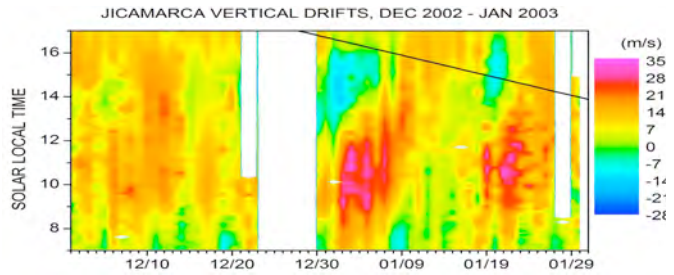


Figure 4. Vertical plasma drifts measured by the JRO ISR during the 2002-3 SSW (after 12/28) indicating strong plasma motions modulated by PWs at ~2- & 16-day periods (Fejer et al. 2010).

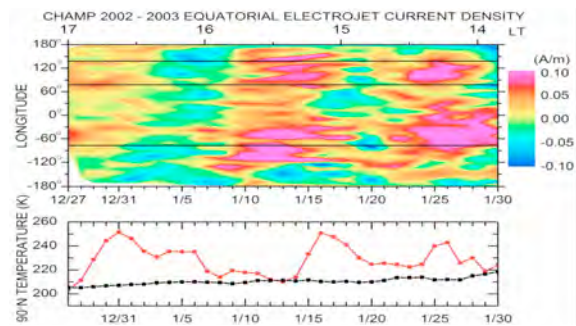


Figure 5. Electrojet current measured by CHAMP (top) and stratospheric temperature over the North Pole at 32 km (bottom, red). Note the strong ~16-day modulation in both data (Fejer et al. 2010).

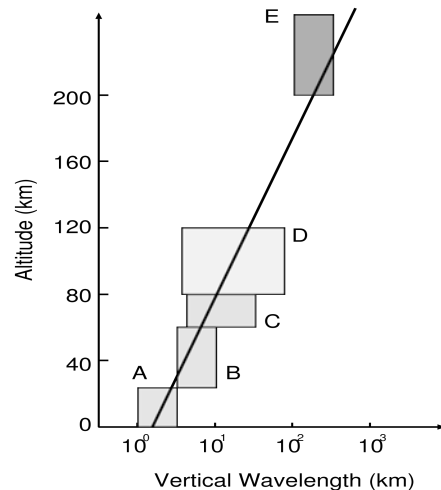


Figure 6. Variations of GW vertical wavelength seen by the MU radar under quiet solar conditions (Oliver et al. 1997).

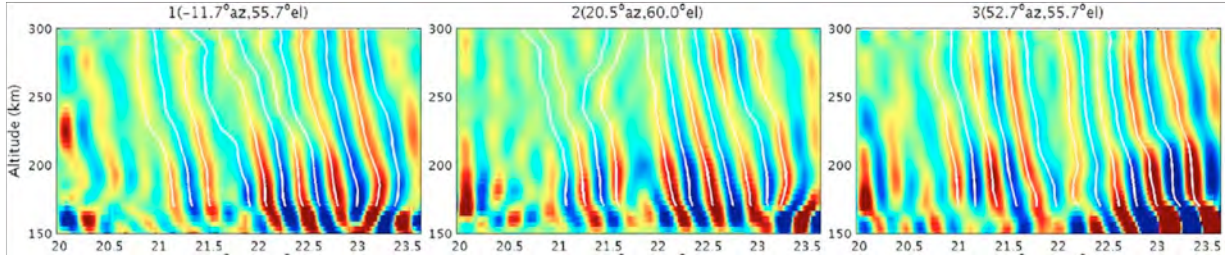


Figure 7. GW vertical structure seen by the PFISR in a multi-beam experiment (Vadas and Nicolls 2009). Maximum electron density perturbations at the lowest altitudes exceed 20%.

reveals fractional electron density perturbations exceeding 20%, with GW structures clearly penetrating to altitudes above ~ 300 km. Also seen here is clear downward phase progression, which is indicative of upward energy propagation for GWs, clearly pointing to a source in the lower atmosphere.

Finally, there is evidence of GW perturbations in electron densities also extending to significantly higher altitudes, despite the dramatic increases in viscosity with altitude. Two examples are shown in Figure 8. The left image is from measurements using the Arecibo ISR during strong auroral energy inputs at high northern latitudes; the right image is digisonde data obtained at Fortaleza, Brazil during the Spread F Experiment (SpreadFEx) (Fritts et al. 2009). The Arecibo measurements indicate penetration of aurorally-generated GWs to altitudes well above 500 km and apparent initiation of plasma instabilities above ~ 300 km. The electron density perturbations at right exhibit large vertical motions and correlate closely with the seeding of equatorial plasma bubbles, suggesting their role in the seeding process (Abdu et al. 2009). In both cases, GW phase structures exhibit downward propagation, confirming upward propagation from sources at lower altitudes.

3. A Canonical Neutral Atmosphere – Ionosphere Coupling Mission

a. Satellite parameters

Neutral waves obviously influence the ionosphere in many ways, their dynamics are expected to be functions of the wave sources, the environments through which they propagate, and the ionospheric conditions that they encounter (*e.g.*, magnetic field orientation), and some of

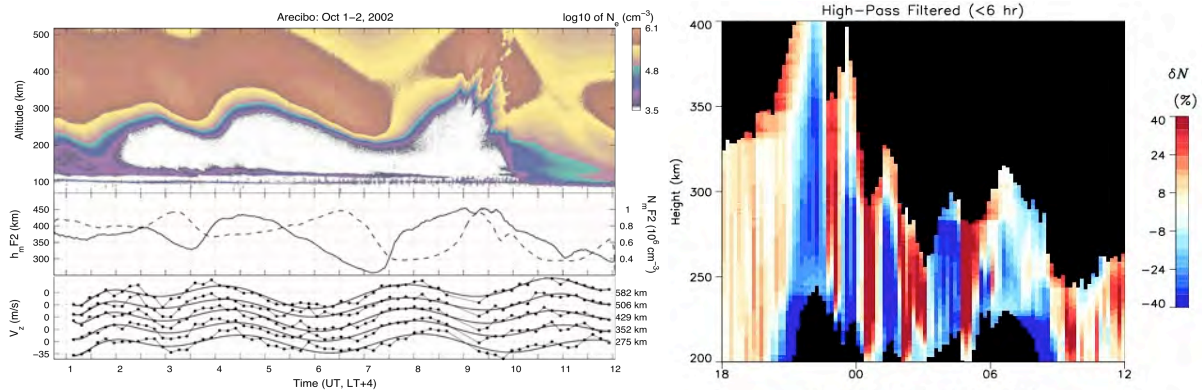


Figure 8. Electron density fluctuations observed by the Arecibo ISR (Nicolls et al. 2004) and by a digisonde at Fortaleza, Brazil (Abdu et al. 2009). These apparently arose due to auroral heating and deep convection, respectively, the latter either directly or as secondary radiation. Note the downward phase progression in each case indicative of upward GW propagation.

the neutral dynamics impacting large- and small-scale wave responses extend to polar latitudes. Thus, a mission to explore these coupling dynamics in a comprehensive manner must perform global measurements of the neutral fields defining the various waves and the environments through which they propagate. Such a mission must also provide sensitivity to the spatial scales (horizontal and vertical) of the dominant wave motions and to their seasonal variations and geographic variability. Spatial scales are simple to define for mean structures, tides, and PWs, as limb measurements of horizontal winds and temperatures are sufficient for this purpose. Limb measurements are not sufficient to define GWs arising from lower atmosphere sources below ~150 km altitude, however, so these motions must be viewed in the nadir or sublimb in order to capture the important horizontal scales. These depend on the altitudes from which it is desired to track GW propagation, as the dominant GW scales increase strongly with altitude into the TI.

The desire for near-global coverage and a reasonable precession rate drives a trade-off between a high inclination (needed for global coverage with limb viewing) and a lower inclination (required for more rapid precession and tidal definition). This choice will also depend on the latitudes over which nadir or sublimb viewing of GWs is desired in order to provide sensitivity to the major source regions. Alternatively, these needs could be separately addressed with two satellites in different inclinations, and likely having different instrumentation.

b. Instrument measurement requirements

Horizontal winds should be defined in the TI, as well as at lower altitudes in order to define the environment through which the various wave motions must propagate en route to the TI.

Limb temperatures are required to define the mean thermal structure and the tides and PWs. However, nadir or sublimb viewing of suitable emissions will be necessary to provide the higher horizontal resolution required to define the GWs arising in the lower atmosphere, depending on the TI dynamics that are intended to be addressed.

References

- Abdu, M. A., E. A. Kherani, I. S. Batista, E. R. de Paula, D. C. Fritts, 2009: An evaluation of the ESF/bubble irregularity growth conditions under gravity wave influences based on observational data from the SpreadFEx campaign, *Ann. Geophys.*, SpreadFEx special issue, 27, 2607-2622.
- Crain, D. J., R. A. Heelis, G. J. Bailey, and A. D. Richmond (1993), Low-Latitude Plasma Drifts From a Simulation of the Global Atmospheric Dynamo, *J. Geophys. Res.*, 98(A4), 6039–6046, doi:10.1029/92JA02196.
- Earle, G. D., A. M. Musumba, and S. L. Vadas, 2008: Satellite-based measurements of gravity wave-induced midlatitude plasma density perturbations, *J. Geophys. Res.*, 113, A03303, doi:10.1029/2007JA012766.
- England, S. L., S. Maus, T. J. Immel, and S. B. Mende (2006), Longitudinal variation of the E-region electric fields caused by atmospheric tides, *Geophys. Res. Lett.*, 33, L21105, doi:10.1029/2006GL027465.
- Fejer, B. G., M. Olson, J. Chau, C. Stolle, H. Luhr, L. P. Goncharenko, K. Yumoto, and T. Nagatsuma, 2010: Lunar dependent equatorial ionospheric electrodynamic effects during sudden stratospheric warming events, *JGR*, 15, doi:10.1029/2010JA015273, published.
- Fejer, B. G., and L. Scherliess, 1981: On the variability of equatorial vertical plasma drifts, *JASTP*, 63, 893-897.
- Forbes, J. M., and S. Leveroni, 1992: Quasi-16 day oscillation in the ionosphere, *Geophys. Res. Lett.*, 19, 981-983.
- Forbes, J. M., S. L. Bruinsma, X. Zhang, and J. Oberheide (2009), Surface-exosphere coupling due to thermal tides, *Geophys. Res. Lett.*, 36, L15812, doi:10.1029/2009GL038748.

- Fritts, D. C., M. A. Abdu, B. R. Batista, I. S. Batista, P. P. Batista, R. Buriti, B. R. Clemesha, T. Dautermann, E. de Paula, B. J. Fechine, B. Fejer, D. Gobbi, J. Haase, F. Kamalabadi, E. A. Kherani, B. Laughman, P. P. Lima, H.-L. Liu, A. Medeiros, D. Pautet, F. Sao Sabbas, J. H. A. Sobral, P. Stamus, H. Takahashi, M. J. Taylor, S. L. Vadas, F. Vargas, and C. Wrasse, 2009: Overview and summary of the Spread F Experiment (SpreadFEx), *Ann. Geophys.*, 27, 2141-2155.
- Fritts, D. C., and M. J. Alexander, 2003: Gravity dynamics and effects in the middle atmosphere, *Rev. Geophys.*, 41, doi:10.1029/2001RG000106.
- Fritts, D. C., D. Janches, H. Iimura, W. K. Hocking, N. J. Mitchell, B. Fuller, B. Vandepeer, J. Hormaechea, C. Brunini, and H. Levato, 2010: Southern Argentina agile meteor radar (SAAMER): System design and initial measurements of large-scale winds and tides, *J. Geophys. Res.*, 115, D18112, doi:10.1029/2010JD013850.
- Fritts, D. C., and S. L. Vadas, 2008: Gravity wave penetration into the thermosphere: Sensitivity to solar cycle variations and mean winds, *Annals. Geophys.*, 26, 3841-3861.
- Fritts, D. C., S. L. Vadas, D. M. Riggan, M. A. Abdu, I. S. Batista, H. Takahashi, A. Medeiros, F. Kamalabadi, H.-L. Liu, B. J. Fejer, and M. J. Taylor, 2008: Gravity wave and tidal influences on equatorial spread F based on observations during the Spread F Experiment (SpreadFEx), *Ann. Geophys.*, 26, SpreadFEx special issue, 3235-3252.
- Fritts, D. C., and R. A. Vincent, 1987: Mesospheric momentum flux studies at Adelaide, Australia: Observations and a gravity wave/tidal interaction model, *J. Atmos. Sci.*, 44, 605-619.
- Goncharenko, L., and S.-R. Zhang (2008), Ionospheric signatures of sudden stratospheric warming: Ion temperature at middle latitude, *Geophys. Res. Lett.*, 35, L21103, doi:10.1029/2008GL035684.
- Hagan, M. E., and J. M. Forbes, 2002: Migrating and nonmigrating diurnal tides in the middle and upper atmosphere excited by tropospheric latent heat release, *J. Geophys. Res.*, 107, 4751, doi:10.1029/2001JD001236.
- Hagan, M. E., A. Maute, and R. G. Roble (2009), Tropospheric tidal effects on the middle and upper atmosphere, *J. Geophys. Res.*, 114, A01302, doi:10.1029/2008JA013637.
- Häusler, K., H. Lühr, M. E. Hagan, A. Maute, and R. G. Roble, 2010: Comparison of CHAMP and TIME-GCM nonmigrating tidal signals in the thermospheric zonal wind, *J. Geophys. Res.*, 115, D00I08, doi:10.1029/2009JD012394.
- Häusler, K., and H. Lühr, 2009: Nonmigrating tidal signals in the upper thermospheric zonal wind at equatorial latitudes as observed by CHAMP, *Ann. Geophys.*, 27, 2643-2652.
- Häusler, K., H. Lühr, S. Rentz, and W. Köhler, 2007: A statistical analysis of longitudinal dependences of upper thermospheric zonal winds at dip equator latitudes derived from CHAMP, *JASTP*, 69 (12), 1419-1430.
- Heelis, R. A., 2004: Electrodynamics in the low and middle latitude ionosphere: a tutorial, *J. Atmos. Solar-Terres. Phys.*, 66(10), 825-838, doi:10.1016/j.astp.2004.01.034.
- Immel, T. J., E. Sagawa, S. L. England, S. B. Herderson, M. A. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton, 2006: Control of equatorial ionospheric morphology by atmospheric tides, *GRL*, 33, doi:10.1029/2006GL026161.
- Kudeki, E., A. Akgiray, M. Milla, J. L. Chau, and D. L. Hysell: Equatorial spread-F initiation: post-sunset vortex, thermospheric winds, gravity waves, *J. Atmos. Solar-Terres. Phys.*, 69, 2416-2427, 2007.
- Liu, H.-L., and R. G. Roble, 2002: A study of a self-generated stratospheric sudden warming and its mesospheric-lower thermospheric impacts using the coupled TIME-GCM/CCM3, *J. Geophys. Res.*, 107(D23), 4695, doi:10.1029/2001JD001533.

- Lühr, H., K. Häusler, and C. Stolle, 2007: Longitudinal variation of F region electron density and thermospheric zonal wind caused by atmospheric tides, *Geophys. Res. Lett.*, 34, L16102, doi:10.1029/2007GL030639.
- Lühr, H., R. Rother, K. Hausler, P. Alken, and S. Maus, 2008: The influence of nonmigrating tides on the longitudinal variation of the equatorial electrojet, *JGR*, 113, doi:10.1029/2008JA0123064.
- McClure, J. P., S. Singh, D. K. Bamgboye, F. S. Johnson, and H. Kil: Occurrence of equatorial F region irregularities: Evidence for tropospheric seeding, *J. Geophys. Res.*, 103, 29,119-29,135, 1998.
- Nicolls, M. J., M. C. Kelley, A. J. Coster, S. A. Gonzalez, and J. J. Makela, 2004: Imaging the structure of a large-scale TID using ISR and TEC data, *Geophys., Res., Lett.*, 31, L09812, doi:10.1029/2004GL019797.
- Oberheide, J., and J. M. Forbes (2008a), Tidal propagation of deep tropical cloud signatures into the thermosphere from TIMED observations, *Geophys. Res. Lett.*, 35, L04816, doi:10.1029/2007GL032397.
- Oberheide, J., and J. M. Forbes (2008b), Thermospheric nitric oxide variability induced by nonmigrating tides, *Geophys. Res. Lett.*, 35, L16814, doi:10.1029/2008GL034825.
- Oberheide, J., J. M. Forbes, K. Häusler, Q. Wu, and S. L. Bruinsma, 2009: Tropospheric tides from 80 to 400 km: Propagation, interannual variability, and solar cycle effects, *J. Geophys. Res.*, 114, D00I05, doi:10.1029/2009JD012388.
- Oliver, W. L., Y. Otsuka, M. Sato, T. Takami, and S. Fukao, 1997: A climatology of F region gravity wave propagation over the middle and upper atmosphere radar, *J. Geophys. Res.*, 102, 14,499-14,512.
- Ortland, D. A., and M. J. Alexander (2006), Gravity wave influence on the global structure of the diurnal tide in the mesosphere and lower thermosphere, *J. Geophys. Res.*, 111, A10S10, doi:10.1029/2005JA011467.
- Pancheva, D. V., P.J. Mukhtarov, N.J. Mitchell, D.C. Fritts, D.M. Riggin, H. Takahashi, P.P. Batista, B.R. Clemesha, S. Gurubaran, and G. Ramkumar, 2008: Planetary wave coupling (5–6-day waves) in the low-latitude atmosphere–ionosphere system, *J. Atmos. Solar-Terres. Phys.*, 70 (1), 101-122.
- Pancheva, D. V., P. J. Mukhtarov, M. G. Shepherd, N. J. Mitchell, D. C. Fritts, D. M. Riggin, S. J. Franke, P. P. Batista, M. A. Abdu, I. S. Batista, B. R. Clemesha, and T. Kikuchi, 2006: Two-day wave coupling of the low-latitude atmosphere-ionosphere system, *J. Geophys. Res.*, 111, A07313, doi:10.1029/2005JA011562.
- Talaat, E. R., and R. S. Lieberman, 2010: Direct observations of nonmigrating diurnal tides in the equatorial thermosphere, *Geophys. Res. Lett.*, 37, L04803, doi:10.1029/2009GL041845.
- Teitelbaum, H., and F. Vial, 1991: On Tidal Variability Induced by Nonlinear Interaction With Planetary Waves, *J. Geophys. Res.*, 96(A8), 14,169–14,178, doi:10.1029/91JA01019.
- Vadas, S. L.: Horizontal and vertical propagation, and dissipation of gravity waves in the thermosphere from lower atmospheric and thermospheric sources, *J. Geophys. Res.*, 112, A06305, doi:10.1029/2006JA011845, 2007.
- Vadas, S. L., and D. C. Fritts: Thermospheric responses to gravity waves: Influences of increasing viscosity and thermal diffusivity, *J. Geophys. Res.*, 110, D15103, doi:10.1029/2004JD005574, 2005.
- Vadas, S. L., and M. Nicolls: Using PFISR measurements and gravity wave dissipative theory to determine the neutral, background thermospheric winds, *Geophys. Res. Lett.*, 35, L02105, doi:10.1029/2007GL031522, 2008.
- Vadas, S. L., and M. Nicolls: Temporal evolution of neutral, thermospheric winds and plasma response using PFISR measurements of gravity waves, *J. Atmos.-Terres. Phys.*, 71, 740-770, 2009.

Solar Proton Event Risk Modeling for Variable Duration Human Spaceflight

D. Fry¹, Anne Adamczyk², S. Blattnig², M. Cloudsley², K. T. Lee¹, J. Norbury², N. Zapp¹

¹Space Radiation Analysis Group
Lyndon B. Johnson Space Center
Houston, TX

²Langley Research Center
Hampton, VA

Introduction

The purpose of this white paper is to outline a strategy for near-term development of modeling tools needed to mitigate Solar Proton Event (SPE) impact for future long-duration human exploration missions. Such a strategy would facilitate greater use of numerous past and current satellite missions, possibly identify new directions for physics-based modeling efforts and foster increased collaboration between the research and operations community. Moreover, many of these same tools would be directly amenable to the needs of National technology infrastructure such as GPS telecommunications, commercial air flight and power grids.

There is a pressing short-term need to develop tools that assess SPE risk on multiple time scales. Identified are the following four areas that should be pursued, with relatively short lead-time, in parallel to physics-based model development:

1. Probabilistic models of SPE occurrence covering mission durations of 1 week to 1 year – useful to mission planners, flight control teams and design trades.
2. Probabilistic models of exceeding SPE energy-dependent flux covering mission durations of 1 week to 1 year – useful as input to transport codes to assess expected exposure (design trades).
3. Probabilistic models of exceeding a threshold dose for a select time period during an event – useful for modifying crew activities to adhere to dose limits and design trades.
4. Probabilistic models of SPE spectral characteristics capable of generating risk estimates over a 12-hour period prior to and during an EVA.

Of the various sources of radiation exposure humans experience during exploration missions, SPEs represent the largest challenge operationally. SPEs, high intensity bursts of energetic ions (primarily protons) with energies exceeding 500 MeV can persist for weeks. There is no known method to predict during any given 24-hour time frame when an event will occur, it's duration, intensity and time profile, or it's spectral

characteristics. SPE mission impact is not known at any point other than in real-time, minute-by-minute as the event commences and evolves. As a result, risk to human crews is frequently addressed by defining an upper bound on SPE fluence using several of the recorded “large” events in the historical record - a worst case analysis that typically incurs additional parasitic mass for vehicle designs, constrains planning of mission profiles, reduces flexibility and increases complexity in both planning and real-time modification of crew activities. Although useful for short-duration mission, this may not be valid for long-duration missions during which crews will spend a substantial fraction of time away from mass shelters. For these missions, cumulative effects of the “smaller” events that occur with higher frequency may be non-negligible. Models that relax this constraint and incorporate the entire distribution of historical events as well as arbitrary mission duration and potential precursors to SPE onset are needed.

Historically, events occur with a mix of rise and decay times. Impulsive events for instance frequently show a 6-order of magnitude increase in proton fluence in the first several hours after event onset. In general, time profiles show variability from relatively smooth exponential decay over the course of several days to a monotonic decay accompanied by a series of additional bursts anywhere from one to two orders of magnitude in fluence intensity.

Physics-based modeling of solar dynamics has continuously progressed over the last 15 years, largely due to the success of numerous NASA Science Mission Directorate (SMD) missions. These same missions supply a great deal of daily situational awareness to space weather operations. In addition, our understanding of fundamental physical processes driving solar activity has drastically increased. From an operational perspective, these models are highly desired, but utility depends on transitioning from research to operations, which is complex and time-consuming. Apart from the efforts of the NASA Community Coordinated Modeling Center (CCMC), there has been limited dedicated community involvement in tool development. Tools to forecast SPE occurrence and characteristics remain elusive.

Mitigation of exposure risk from SPEs for any human exploration mission occurs on the ground pre-mission through design trades to assess vehicle shielding, and in real-time during mission operations. It is important to understand what is needed both by mission planners and real-time operations to mitigate SPE impact. For example, mission planners require the ability to assess risk of SPE occurrence over the duration of the mission, when planning EVA activities, and manipulating daily crew activities. In real-time, operators need the ability to assess risk on time frames from 24 hours down to nominal 6 hour EVA duration, with advanced knowledge of event particle flux spectra and time evolution. In all, tools are needed to inform crews and flight teams over a large dynamic time range of potential high-risk conditions. Also, architecture and system designers need tools to perform trade studies.

Given near-term development timelines, augmentation of past risk mitigation strategies is of utility to both mission design and operations. Tools are needed to allow for more comprehensive design trades to be conducted in radiation shielding and to help mission

operators make more informed decisions in assessing real-time impact to crews. Physics-based models are clearly the preference, but other short-term methods need to be investigated in the interim. Past work and gaps are first identified below and then a brief outline of how to augment these efforts in the future efforts is given.

Past Work and Gaps

Several gaps exist in pursuing a direct translation from hardware assessment models to those applied to human spaceflight.

- There is limited spectral information contained in risk estimates developed with any of the existing models [1-7] covering the range (>1 MeV to > 500 MeV) needed for human space flight.
- A second issue is the statistical quality of the data. Detector design has evolved over the last several decades resulting in increased energy resolution capability and dynamic range, but there are gaps. Piecing together proton fluence spectra from different satellites that have different orbital tracks is time consuming, and has potential to introduce uncertainties that skew model results. Observations over different time spans have typically been conducted over different regions of the proton spectral energy range needed for quantitative dose calculations. There has not been any study to date that systematically compares fluence spectra from the early 1960s to the present day that ascertains whether or not data gaps or assumption made to fill gaps induces negligible uncertainty in calculated crew dose.
- There is also variance in which time span is used to calculate integral fluence values. Of the known models, each utilized annual proton fluence, daily fluence and/or event-integrated fluence. In general, for human exploration missions models need to be constructed from integral fluence values where the limits of integration δt are limited to $\delta t \leq T$ the mission duration. Daily fluence should be sufficient for any human mission durations considered, but a full comparison of available models with the same integration range are lacking.
- Dedicated space-based observation platforms for measuring SPE spectra out to >500 MeV are lacking.

There has been a fair amount of high quality efforts aimed at assessing risk of SPEs for robotic missions as well as design trades for human rated vehicles. Unfortunately, the majority of risk models are not directly amenable as tools for human spaceflight and those that have been attempted are limiting when considering long-duration missions. There are multiple factors of influence. Much more shielding is needed for human rated vehicles, which make large safety factors less of an option, and thus more detailed and accurate modeling is needed. Mission durations for robotic and human space flight missions are not commensurate – robotic missions are typically more than a year in duration whereas long-duration human missions are being considered at most a year in length. Considering the growth in observational platforms over the last three decades

and inherent variance in spectral range of proton fluence measured, the dataset of historical events is sparse and has numerous gaps for which spectral measurements are not available. Past missions have been conducted with the majority of crew time spent inside spacecraft, where parasitic shielding mass is designed to standards set by considering an upper bound to proton fluence determined by several historically intense events. Long-duration space flight that includes landing on other planetary/Near Earth Object (NEO) surfaces will likely include more time outside of well-shielded environments, and thus tools are needed that address SPE risk over time frames of hours to total mission duration.

In general, risk mitigation strategies can be grouped into two categories: (1) Physics-based models aimed at forecasting SPEs, and; (2) Historical data-driven statistical models. Physics-based models are capable of producing the most complete picture of solar dynamics. However, they are highly complex, typically have long lead times to develop, require extensive data sets for verification and validation, and as a result are costly.

Conversely, data-driven models, built upon statistical analysis of what has occurred during past events, can be developed in ways analogous to other risk mitigation and forecasting strategies. This approach is less time consuming, has a tried-and-tested history of use in conventional terrestrial weather forecasting, but is highly dependent on the size and quality of historical observations. A clear advantage however is that the detailed underlying physical processes generating SPEs and propagation throughout the heliosphere are not *a priori* required. These methods could provide a way to develop valuable tools at relatively low cost, building on observations already made, making use of an extensive history of use in other fields such as economics, engineering design, real estate market prediction and epidemiology, and possibly identifying new direction for development of physics-based modeling.

Past statistical approaches have several commonalities. All models make use of historical event data casting probability in terms of distributions of proton fluence. The models of Feynman *et al.* [1-2] consolidated proton flux measurements made between 1963 and 1991 (solar cycles 19-21, IMP and OGO spacecraft) and assess distributions of daily fluence values for integral proton energies from >1 MeV to >60 MeV. Results showed that over a cumulative probability range from 25% - 99% the distribution of fluence at fixed integral energy is nearly log-normal, facilitating sampling from a known distribution. Other studies have found similar results for integral fluences out to >100 MeV [3-6]. Risk estimates for exceeding a preset fluence for mission durations ≥ 1 year are then determined by assuming individual events are Poisson-distributed in time over a solar active period centered on the peak in annual sunspot number, and sampling the log-normal distribution of fluence. Modifications to this approach have also been conducted [7] but with incomplete spectral information.

An alternate methodology is to use Maximum Entropy Theory (MET). Particularly suited for dealing with incomplete data sets and analysis of rare events, Xapsos *et al.* applied MET to produce a model of probability of exceeding a given fluence for missions ≥ 1

year in duration at a select confidence level [3-6]. Considering that roughly 100 out of 6386 (1986-2010) days resulted in SPE onset (total of ~4.2% of days at elevated proton fluence levels), it is feasible that MET might work quite well for projections of exceeding a pre-determined cumulative fluence threshold. How this can be mapped into risk assessment for crew dose is unclear at the present time.

Design trades are critical for assessing the level of shielding any vehicle will provide, feeding input in real-time during the design cycle. Due to a lack of sufficient event data and available tools, the vast majority of past work has dealt with SPE impact by assuming an upper bound set by either the August 1972 (King), the November 1960, the February 1956, or the September 1989 events, four characteristically large events recorded during the modern space flight era [10-18]. In addition, combinations of these events have been used to represent an upper bound in particle fluence covering the largest spectral proton energy range available [15]. Neglecting the remaining events in the historical records however explicitly constrains the trade space to worst-case scenarios, which typically result in extra parasitic mass.

Additional efforts have been made that address crew risk from models that incorporate biological effects [19-21]. The relatively large uncertainty in biological endpoint, limited data sets covering particle type and energy similar to the interplanetary environment, and high variability of genetic factors between crewmembers present a large challenge towards effective utilization of these models. In general assessment has been limited to integral proton energy of >30 MeV. This is due to a tradeoff between expanding the number of historical events observed to decrease statistical error and the cost of working with gaps in the proton energy range measured over the last 60 years. Attempts have been made to circumvent the issue of limited observations by assuming a particular spectral form and fitting observations to extend to higher proton energy [22].

Next Steps

It is expected that exploration of solar dynamics will continue with new and advanced missions being launched in upcoming years. However, other avenues are potentially available to bring to fruition in the near term tools of utility to both operations and vehicle design by incorporating analysis techniques similar to terrestrial weather forecasting and risk mitigation in other disciplines. Risk modeling is not new science. There are numerous risk modeling strategies available that should be carefully looked at, with the potential to produce within the next several years tools that will assist operators and mission designers in assessing SPE risk [23-30].

Physics-based modeling is currently a focus of heliophysics. **It is proposed here that efforts also be dedicated to probabilistic risk-based modeling that address multi-variate correlations and trending to identify characteristics such as event size, evolution, and occurrence that have direct overlap with operational timelines of mission planning, mission duration, EVAs, daily crew activities and design**

trades. To accomplish this task there needs to be a concerted effort from all agency stakeholders in SPE mitigation (SMD/SOMD/ESMD) to:

- 1.) Ensure dedicated measurement of SPE fluence spectra out to >500 MeV (preferably 1 GeV), at a maximum cadence of 5-min, by leveraging future robotic missions.
- 2.) Glean as much information as possible from the historical event record to support development of near-term tools characterizing event probability of occurrence, size and spectral characteristics over proton energies from >1 MeV to >500 MeV, time evolution and duration.
- 3.) Develop near-term SPE forecasting tools based entirely on SPE statistics, and feedback to the long-term development of physics-based forecasting models
- 4.) Maximally leverage risk model development in other disciplines. Risk assessment and forecasting of extreme/rare events is a specific area of research in terrestrial weather, financial markets and geology (earthquakes) that have direct application to SPEs.
- 5.) Develop dedicated funding to support the transition from research to operations fostering collaboration between the research and operations communities extending beyond agency efforts.

There are essentially two mutually exclusive paths to deal with any risk – limit the risk, and mitigate a specific endpoint of the risk. It is widely accepted that the largest source of uncertainty in assessing risk of exposure is uncertainty in biological endpoint [19-21]. Shielding, mission planning, and real-time operational contingencies are used to limit exposure risk. Efforts to minimize biological uncertainty *and* develop better tools should be worked in parallel. This will allow agency science to directly feed operations with tools based on the current state of knowledge and build strategies for transition from research to operations when more mature physics-based models are available. Current space weather operations are centered on minimizing crew exposure levels, i.e. ALARA, or As Low As Reasonably Achievable. This will likely remain an overarching principle for future long-duration missions.

Operators need quick, relatively short outlook tools to better inform both the crew and mission planners of impending SPE impact. Architecture design trades need statistical tools that allow for a probabilistic sampling over larger range of proton energy and fluence. Future efforts aimed at long-term development of physics-based models and near-term development of probabilistic tools should be a key focus of the next decade of heliophysics. Ensuring a continual measurement of needed spectral ranges of proton fluence and fostering a robust collaboration between research and operations will support both long-term and near-term goals of the agency to protect crews on long-duration missions.

Bibliography

1. J. Feynman, T. P. Armstrong, L. Dao-Gibner and S. Silverman, New Interplanetary Proton Fluence Model, *Journal of Spacecraft*, **27**, 403 (1990).
2. J. Feynman, G. Spitale and J. Wang, Interplanetary Proton Fluence Model: JPL 1991, *Journal. of Geophysical Research: Space Physics*, **98**, (1993).
3. M. A. Xapsos, G. P. Summers, P. Shapiro and E. A. Burke, New Techniques for Predicting Solar Proton Fluences for Radiation Effects Applications, *IEEE Transactions on Nuclear Science*, **43** (1996).
4. M. A. Xapsos, G. P. Summers and E. A. Burke, Extreme Value Analysis of Solar Energetic Proton Peak Fluxes, *Solar Physics*, **183**, 157-164 (1998).
5. M. A. Xapsos, J. L. Barth, E. G. Stassinopoulos, E. A. Burke and G. B. Gee, Space Environmental Effects: Model for Emission of Solar Protons (ESP) - Cumulative and Worst-Case Event Fluences, *NASA/TP-1999-209763* (1999).
6. M. A. Xapsos, J. L. Barth, E. G. Stassinopoulos, S. R. Messenger, R. J. Walters, G. P. Summers and E. A. Burke, *IEEE Transactions on Nuclear Science*, **27** (2000).
7. R. A. Nymmik, Probabilistic Model for Fluences and Peak Fluxes of Solar Energetic Particles, *Radiation Measurements*, **30**, 287-296 (1999).
8. I. Jun, R. T. Swimm, A. Ruzmaikin, J. Feynman, A. J. Tylka and W. F. Dietrich, Statistics of Solar Energetic Particle Events: Fluences, Durations, and Time Intervals, *Advances In Space Research*, **40**, 304-312 (2007).
9. M. Y. Kim, J. W. Wilson, F. A. Cucinotta, L. C. Simonsen, W. Atwell, F. F. Badavi and J. Miller, Contribution of High Charge and Energy (HZE) Ions During Solar-Particle Event of September 29, 1989, *NASA/TP-1999-209320* (1999).
10. NASA's Exploration Systems Architecture Study Final Report, *NASA/TM-214065* (2005).
11. T. T. Pham and M. S. El-Genk, Dose Estimates in a Lunar Shelter with Regolith Shielding, *Acta Astronauta*, **64**, 697-713 (2009).
12. L. C. Simonsen, J. W. Wilson, M. H. Kim, and F. A. Cucinotta, Radiation Exposure for Human Mars Exploration, *Health Physics*, **79**, 515-525 (2000).
13. M. S. Cloudsley, B. M. Anderson, N. Luetke, and J. W. Wilson, "Calculation of Radiation Protection Quantities and Analysis of Astronaut Orientation

Dependence,” in *Proceedings of AIAA Space 2006 Conference*, San Jose, California, September 19-21, 2006, AIAA 2006-7441.

14. M. S. Cloudsley, J. E. Nealy, J. W. Wilson, B. M. Anderson, M. S. Anderson, and S. A. Krizan, “Radiation Protection for Lunar Mission Scenarios,” in *Proceeding of AIAA Space 2005 Conference*, Long Beach, California, August 30 – September 1, 2005.
15. Shielding Strategies for Human Space Exploration, Edited by J. W. Wilson, J. Miller, A. Konradi, and F. A. Cucinotta, *NASA Conference Publication 3360* (1997).
16. L. C. Simonsen, J. E. Nealy, L. W. Townsend and J. W. Wilson, Radiation Exposure for Manned Mars Surface Missions, *NASA/TP-2979* (1990).
17. L. C. Simonsen and J. E. Nealy, Radiation Protection for Human Missions to the Moon and Mars, *NASA/TP-3019* (1991).
18. W. D. Hypes, A. J. Butterfield, C. B. King, G. D. Qualls, W. T. Davis, M. J. Gould, J. E. Nealy and L. C. Simonsen, Concepts for Manned Lunar Habitats, *NASA/TP-104114* (1991).
19. J. W. Wilson, J. E. Nealy, W. Schimmerling, F. A. Cucinotta and J. S. Wood, Effects of Radiobiological Uncertainty on Vehicle and Habitat Shield Design for Missions to the Moon and Mars, *NASA/TP-3312* (1993).
20. F. Cucinotta, W. Schimmerling, J. Wilson, L. Peterson and J. Dicello, Space Radiation Cancer Risk Projections for Exploration Missions: Uncertainty Reduction and Mitigation, *NASA/TP-2002-210777* (2002).
21. M. Y. Kim, G. De Angelis and F. A. Cucinotta, Probabilistic Assessment of Radiation Risk for Astronauts in Space Missions, *Acta Astronautica*, in press (2010).
22. M. Y. Kim, M. J. Hayat, A. H. Feiveson and F. A. Cucinotta, Prediction of Frequency and Exposure Level of Solar Particle Events, *Health Physics* (2009).
23. D. Vose, Risk Analysis: A Quantitative Guide, 3rd Edition, John Wiley and Sons, Ltd. (2008).
24. C. A. Doswell III, R. Davies-Jones, and D. L. Keller, On Summary Measures of Skill in Rare Event Forecasting Based on Contingency Tables, *Weather and Forecasting*, **5**, 576-585 (1990).
25. E. J. Gumbel, Statistics of Extremes, Dover Publications (2004).

26. G. King, L. Zeng, Logistic Regression in Rare Events Data, *Society for Political Methodology* (2001).
27. G. King and L. Zeng, Explaining Rare Events in International Relations, *International Organization*, **55**, 693-715 (2001).
28. R. W. Katz, M. B. Parlange and P. Naveau, Statistics of Extremes in Hydrology, *Advances in Water Resources*, **25**, 1287-1304 (2002).
29. T. Bedford and R. M. Cooke, Probabilistic Risk Analysis: Foundations and Methods, Cambridge University Press (2001).
30. G. R. Dargahi-Noubary, The Use of Modern Statistical Theories in the Assessment of Earthquake Hazard, with Application to Quiet Regions of Eastern North America, *Soil Dynamics and Earthquake Engineering*, **22**, 361-369 (2002).

Forecasting Ionospheric Irregularities

Tim Fuller-Rowell, Arthur Richmond, Joe Huba, Rashid Akmaev,
Mihail Codrescu, David Anderson, John Retterer

There is a pressing need to develop a capability to forecast the likely occurrence of ionospheric irregularities and their detrimental impact on communications and navigation. Forecasting the day-to-day variability of ionospheric irregularities is one of the long-standing challenges in space physics. Recent observations, and development of whole atmosphere models, have now revealed that large-scale, near-global atmospheric processes are directly impacting the ionosphere in a predictable way. For instance, the large-scale meteorological planetary scale atmospheric process, known as a sudden stratospheric warming (SSW), can impose a coherent predictable response in the upper atmosphere and ionosphere. The simulations also indicated that the dynamic response can be forecast a few days ahead, indicating, for the first time, the possibility of ionospheric irregularity prediction. The speculation is that the evolution of the background wind, wave, and electrodynamics fields provide the drivers, or the precursors, for the onset of irregularities. The capability to model the coupled atmosphere-ionosphere system is required to determine if this speculation can be turned into a reality. This suggests that space weather also has a terrestrial driver. To be able to predict the impact of this natural hazard on operational systems is of huge societal benefit and is potentially life saving.

Although low to mid latitude irregularities have been modeled for several years the capability to forecast their occurrence and day-to-day variability is still elusive and remains a challenge. We understand many of the causes for the structure; one of the important mechanisms being the generalized Rayleigh Taylor instability. If the ionosphere has been raised to higher altitude by large-scale background electric fields or winds, the medium can become gravitationally unstable. Any perturbation or ripple can provide the seed that can generate a localized polarization field that allows the plasma-rich regions to fall in patches and the plasma-poor regions to float upwards. The result is the picture seen in Figure 1, with extreme gradients at the boundaries of the plasma bubbles. The GRT is not the only mechanism. Wind shear on the bottom-side ionosphere can also create structure by a collisional shear instability (CSI). Modeling showed recently that the two mechanisms can combine to produce apparently emergent phenomena, such that together the instabilities grow faster, and with characteristics that are more aligned with observations.

Recent measurements and modeling indicate that meteorology and atmospheric science may provide the key for predicting these irregularities. It has recently been revealed that large-scale, near-global atmospheric processes are directly impacting the ionosphere in a predictable way. For instance, the large-scale meteorological planetary scale atmospheric process, known as a sudden stratospheric warming (SSW), which can last for several days and sometimes weeks can impose a coherent predictable response in the upper atmosphere and ionosphere.

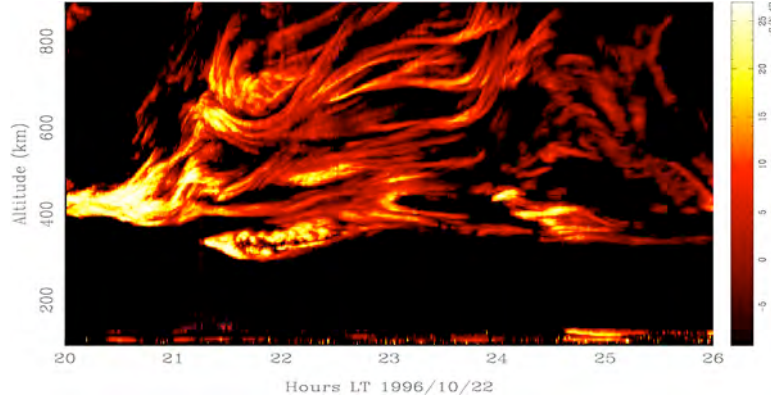


Figure 1. Spectacular observation from the JULIA coherent scatter radar of ionospheric irregularities above Jicamarca, Peru, which cause the severe disruption of satellite communication and navigation signals (courtesy David Hysell).

Under normal conditions a strong circumpolar vortex develops in the winter polar stratosphere. During an SSW this vortex breaks down; the zonal winds reduce and sometimes even reverse from eastward to westward. The change in circulation is often associated with a build up in the amplitude of either planetary wave number 1 or 2, or both, in the troposphere and stratosphere (see Figure 2 from Whole Atmosphere Model simulations).

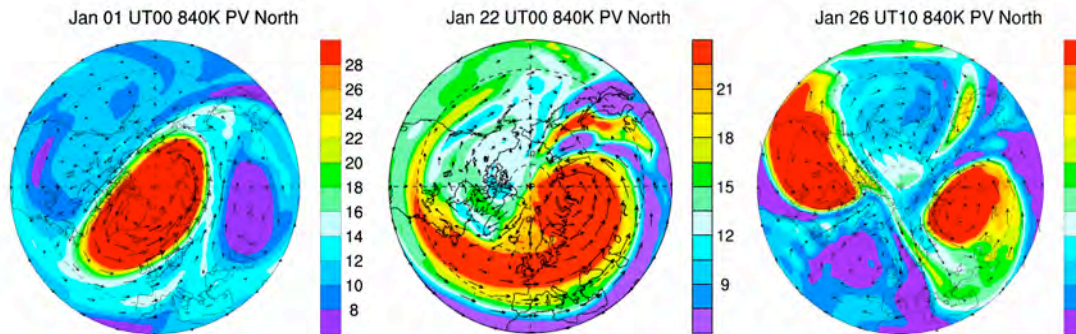


Figure 2. Illustration of planetary wave zero, 1, and 2 on the isentropic potential vorticity surface at 840K, at an approximate altitude of 30 km in the stratosphere, from WAM. On the left is a fairly symmetric vortex (wave zero), middle a displaced vortex (wave 1), and on the right a split vortex (wave 2).

Our current hypothesis is that the development of the higher order planetary waves seen in Figure 2, and changes in the lower and middle atmosphere circulation and temperature structure, modify the vertical propagation of tidal modes generated in the troposphere and stratosphere by absorption of solar radiation by water vapor and ozone, respectively. These tides drive wind and wave fields and electrodynamics in the conducting upper atmosphere, which impact the ionospheric structure and the likelihood of triggering irregularities.

This challenge should ideally be addressed through a combined theory, modeling, and observational thrust. The theory and modeling have made some important advances recently, both in the large-scale whole atmosphere models with interactive electrodynamics, and the small-scale

simulations of ionospheric irregularities. These modeling advances need to be integrated, tested, and quantified, and supported by a comprehensive observation program from ground and space of both the large-scale drivers of the system – neutral atmosphere dynamics, electrodynamics, and waves – and the characteristics of the ionospheric irregularities themselves – gradients, scale-sizes, drift, lifetime, and evolution.

MAGnetosphere-Ionosphere Connector (MAGIC): Investigation of Magnetosphere-Ionosphere Coupling from High-to-Low Latitudes

A White Paper submitted to 2013 Solar and Space Physics Decadal Survey

Shing F. Fung (shing.f.fung@nasa.gov)

Robert F. Benson

Joseph M. Grebowsky

George V. Khazanov

NASA/Goddard Space Flight Center, Greenbelt, Maryland

Donald L. Carpenter (Emeritus)

STAR Laboratory, Stanford University, Stanford, California

Abstract

An outstanding challenge in Heliophysics is to understand the physics of systemic coupling between the magnetosphere and ionosphere from hemisphere to hemisphere, and the roles played by the intervening plasmasphere. Measurements from past missions like the dual *Dynamics Explorers (DE)*, *Polar*, and *IMAGE* have shown the complexities of the inner magnetosphere. Yet, we still do not have a coherent understanding of the multi-scale processes and their interrelationships that couple the ionosphere and magnetosphere from high to low latitudes.

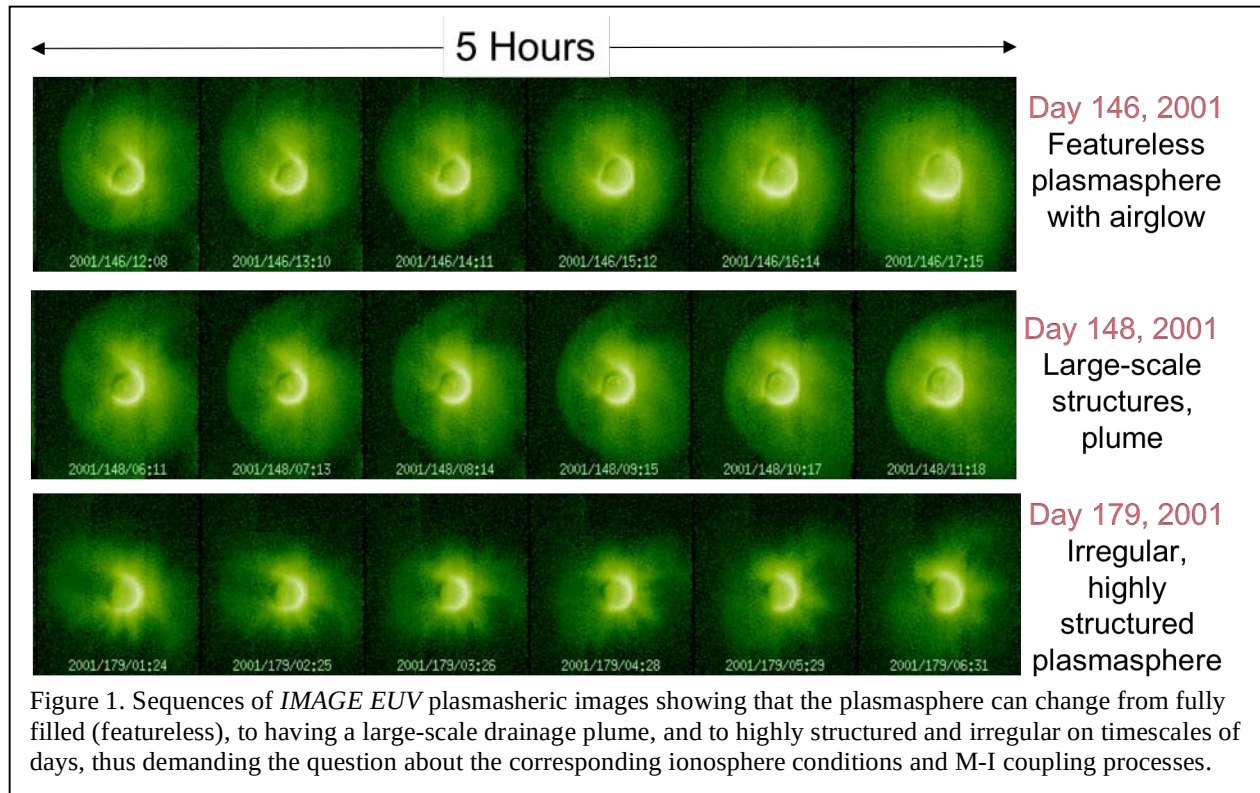
DE, *Polar* and *IMAGE* clearly demonstrated the importance of global-scale imaging in providing the context for understanding local measurements. Except for *DE* limited to high latitudes, imaging observations from *Polar* and *IMAGE* were seldom accompanied by coordinated *in situ* observations of targeted regions. For example, plasmasphere imaging by *IMAGE EUV* were performed at $L > 4$ and near apogee ($\sim 8 R_E$), but *in situ* or radio-sounding measurements of the plasmasphere were best when the *IMAGE Radio Plasma Imager* was nearer the plasmasphere ($L < 4$).

Plasmaspheric mass loss through large-scale plumes has been confirmed by *IMAGE*, although quantitative knowledge of such losses and their impact on the ionosphere and magnetosphere are still lacking. *IMAGE* also revealed numerous previously unknown plasmopause structures, such as notches, fingers, and shoulders, indicating much unknown about the coupling processes of ionosphere-plasmasphere-magnetosphere system. Other topics that need investigation in the context of magnetosphere-ionosphere coupling include the plasmasphere boundary layer, plasmaspheric thermal structure, ring current interaction, plasmaspheric refilling as revealed from field-aligned density profiles, field-aligned irregularities, ionospheric outflows, equatorial *spread-F* plumes, etc.

We propose the *MAGnetosphere-Ionosphere Connector (MAGIC)* mission consisting of a high-altitude, high-inclination satellite for auroral and plasmaspheric imaging and multiple lower-orbiting spacecraft for simultaneous *in situ* and radio sounding measurements. Complemented by ground-based observations, *MAGIC* will investigate magnetosphere-plasmasphere-ionosphere coupling processes from high-to-low latitudes.

Introduction

Magnetosphere-ionosphere coupling is an age-old problem in heliophysics. Except for the direct influence of solar irradiance on the underlying ionosphere and atmosphere, magnetosphere-ionosphere (M-I) coupling processes mediate almost all effects of solar wind input into the ionosphere and its



feedback on the magnetosphere above through the plasmasphere. Understanding M-I coupling processes is critical to understanding the solar-terrestrial relationship.

Since its discovery in the 1950's, the plasmasphere has been regarded as a passive body of cold plasma with little activity or consequences. The predication in the 1970's of a large-scale drainage plume as a result of a preponderance of magnetospheric convection electric field over the plasmaspheric co-rotational electric field [Grebowsky, 1970] was only verified after the 2000-launch of the *Imager for Magnetopause-to-Aurora Global Exploration (IMAGE)* satellite. Figure 1 shows different 5-hour sequences of global-scale images of the plasmasphere obtained by observing the resonantly scattered 304-nm solar radiation by He^+ ions. The highly variable plasmaspheric structures clearly imply the presence of correspondingly significant changes in ionospheric conditions mediated by M-I coupling processes. To fully comprehend M-I coupling we will need to investigate not only the physics of individual coupling processes, but also their inter-relationships, so that we can arrive at an understanding of M-I coupling at the Sun-Earth system level.

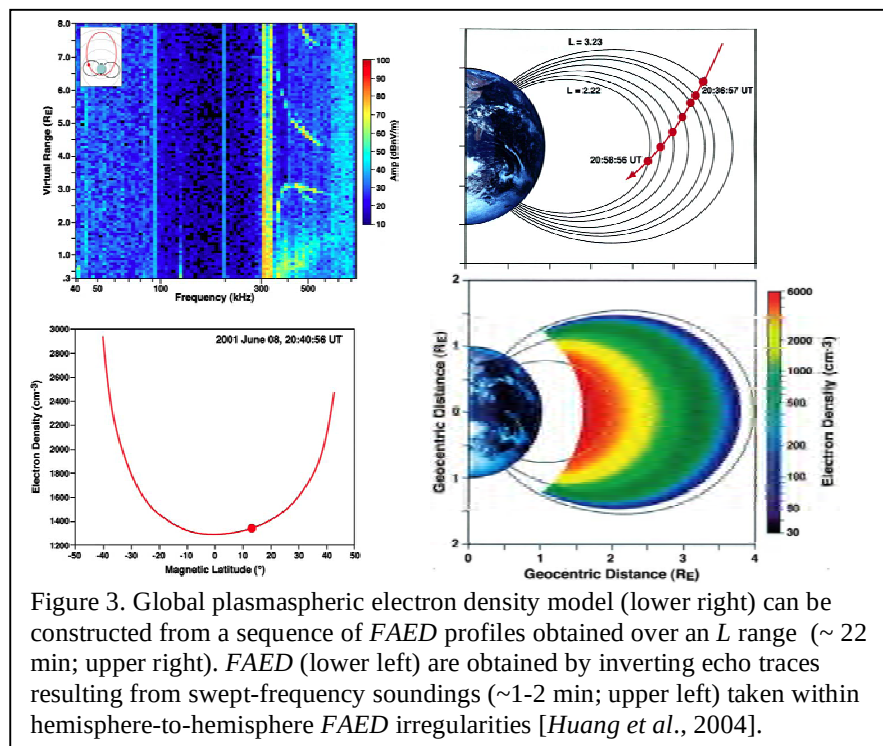
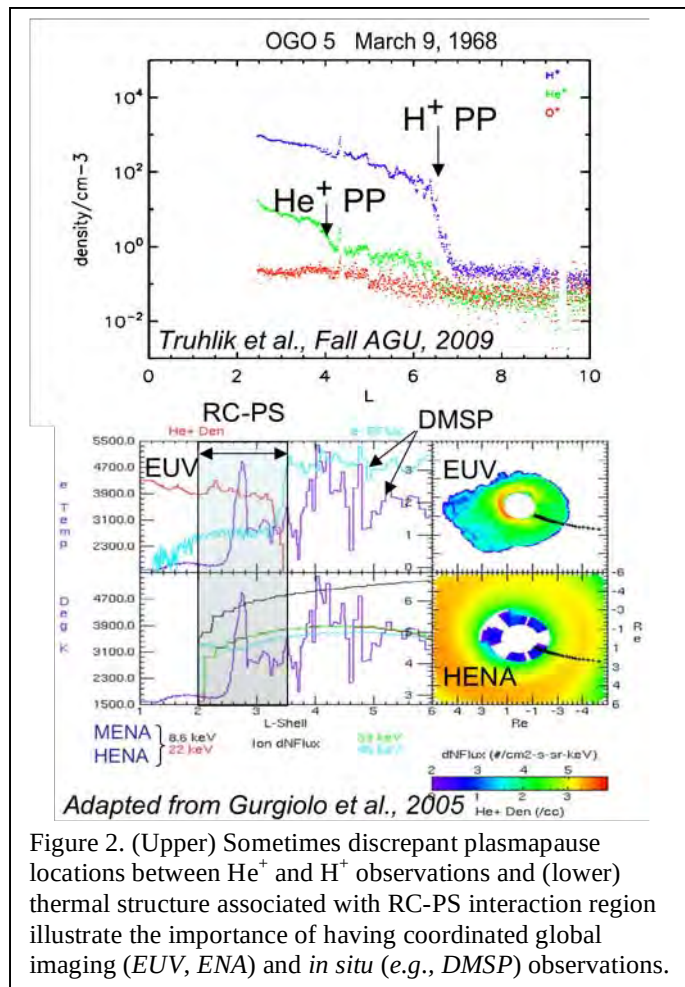
M-I Coupling: Plasmaspheric Perspective

The plasmasphere boundary layer can change from featureless to appearing highly structured on timescales of hours to days (Figure 1). Such dynamic variability must be accompanied by changes in plasmaspheric and ionospheric structures. Figure 2 upper panel shows the ion composition measurements by the *OGO 5* satellite on March 9, 1968, revealing significant differences at times in the H^+ and He^+ plasmopause locations. Since *EUV* observation depends on He^+ distributions, had *EUV* imaging been available during the time of *OGO 5* observations it would have led to the conclusion of a much smaller plasmasphere than that indicated by the H^+ observations. *In situ* ion composition measurements are thus critical for relating internal plasmaspheric structures with magnetospheric dynamics and interpreting global *EUV* images (Figure 1).

Figure 2 lower panel shows the overlapping region (shaded) between the plasmasphere (PS) at $L \leq 3.5$ (represented by the red He^+ -density curve in the upper left portion) and the ring current (RC) at $L > 2$

[demarcated by the smooth energetic neutral atom (*HENA* and *MENA*) profiles in the lower left portion]. While electron precipitation fluxes (light blue curve in upper portion) are generally low inside the PS as shown by the simultaneous *DMSP* magnetic conjunction measurements, electron precipitation can be more significant in the higher latitude plasma trough and subauroral regions, where the electron temperature (purple curve) can also be higher. Inside the PS-RC interaction region, however, localized electron heating occurred apparently due to plasma instabilities or wave-particle interaction processes. Understanding the local wave and wave-particle processes may also hold a key to understanding the plasmapause irregularities [Carpenter *et al.*, 2002] and their roles in mediating systemic (larger-scale) M-I coupling.

In order to properly investigate wave propagation and wave-particle interaction processes in the presence of thermal structure in the plasmasphere and in the PS-RC interaction region as shown in Figure 2 lower panel, we must have adequate knowledge of plasmaspheric electron and ion distributions. While ion distributions can be obtained from composition measurements, *nearly instantaneous* field-aligned electron distribution, critical for understanding plasmasphere erosion and refilling processes [Reinisch *et al.*, 2004], can *only* be obtained by radio sounding technique [Huang *et al.*, 2004]. Figure 3 shows an example of ducted echo traces (upper left) and the corresponding inverted field-aligned electron density (*FAED*) profile (lower left) from one set of swept-frequency soundings observation near $L = 3$ in the plasmasphere. By combining successive soundings through the plasmasphere (Figure 3 upper right), it is possible to construct a plasmaspheric electron model distribution (Figure 3 lower right) on half-



hour time scale [Reinisch *et al.*, 2009], sufficient to resolve storm or substorm dynamics.

The existence of an annual density variation in the plasmasphere has been known for many years [e.g., pp. 89-92 in Lemaire and Gringauz, 1998]. Past work on the annual density variation found a peak density in December a factor of ~ 3 above levels in June over a broad L range inside $L \sim 4$. Limited modeling work that has been done on this phenomenon has suggested that the annual density variation may be related to the asymmetry in the conjugate ionospheres (associated with the large offset in Earth's spin and magnetic axes near the 75° -west longitude) and thus is an issue of ionosphere-plasmasphere coupling. It seems possible that the same conditions that favor the occurrence of the annual density variation could also affect the subauroral ionosphere and contribute to locally deeper or more efficient penetration of convection electric fields to lower latitudes, thus giving rise to certain temporal and/or longitudinal preference in plume formation. It is therefore quite possible that there are previously unappreciated plasmasphere-ionosphere coupling processes that may lead to both temporal (seasonal) as well as longitudinal variations in the erosion of the plasmasphere and associated plume formation.

M-I Coupling: Ionospheric Perspective

With myriad plasma processes having a wide range of spatial and temporal scales, the ionosphere is a complex system in its own right. Many of the processes involving Earth's magnetic field extend to altitudes well beyond the ionosphere, coupling the ionosphere to the thermosphere below and magnetosphere above. Figure 4 depicts many of the processes that occur from the equator to the poles. In

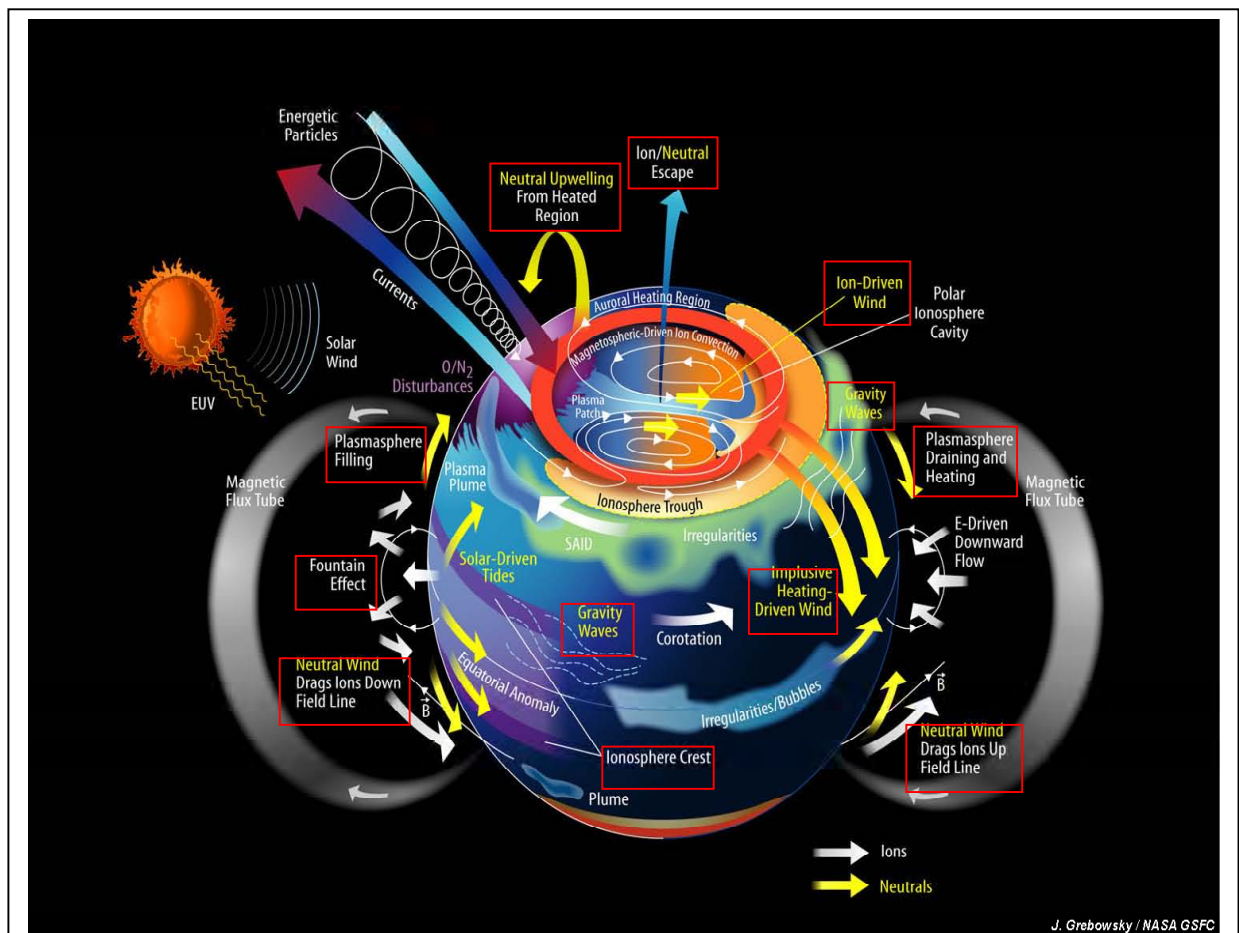
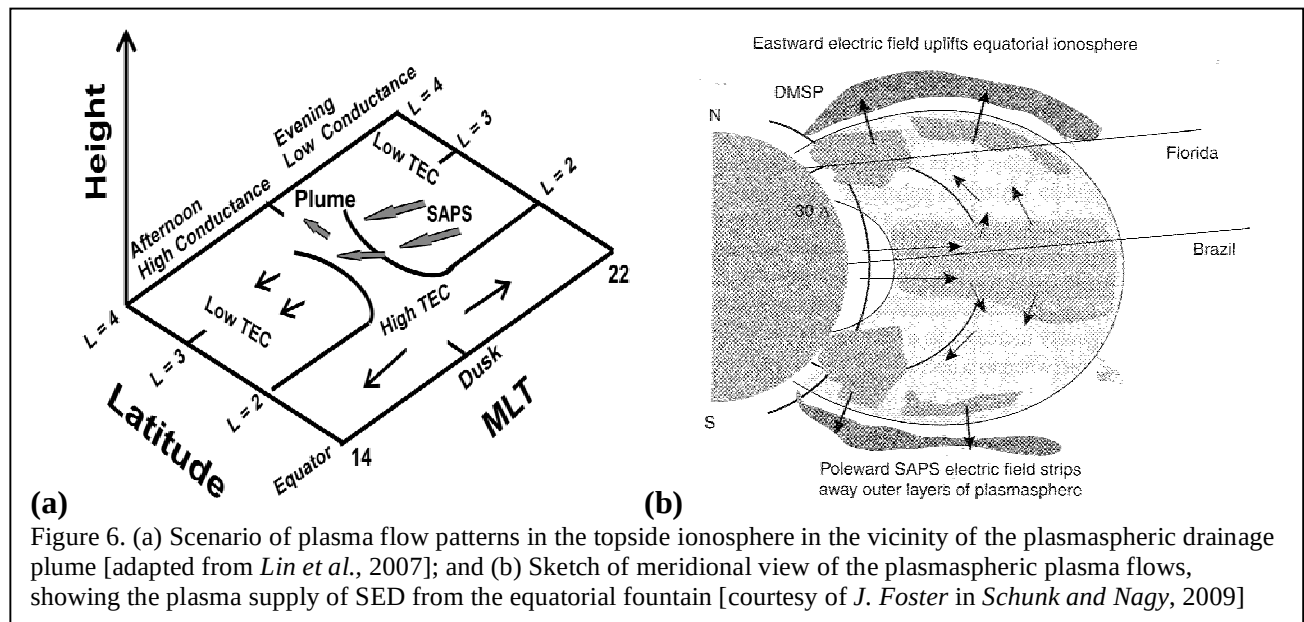
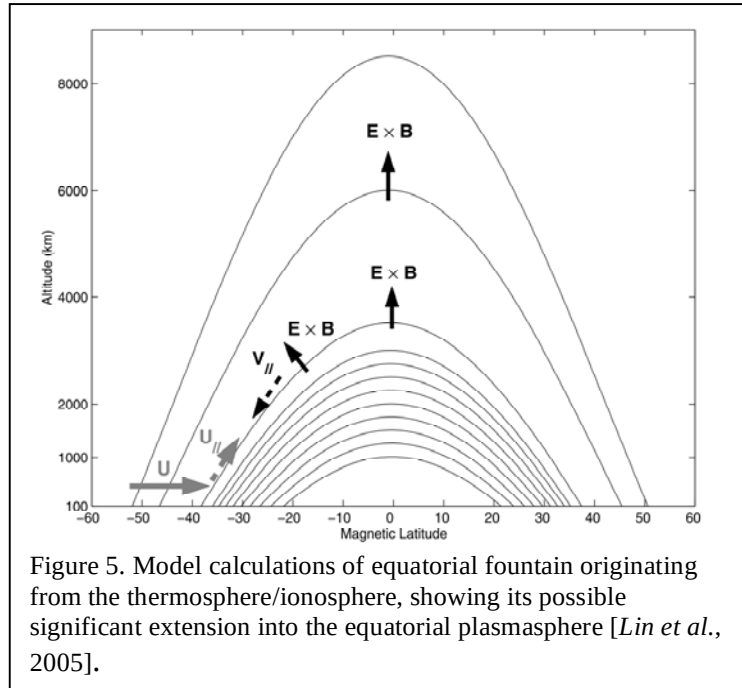


Figure 4. Complex ionospheric/thermospheric processes occurring from equator to poles, many of which (highlighted by red boxes) extend and couple to the plasmasphere and magnetosphere above, contributing to systemic magnetosphere-ionosphere coupling.

recognizing their potentially important contributions to systemic M-I coupling, the red boxes in Figure 4 highlight the pertinent processes that collectively link the ionosphere/thermosphere to the plasmasphere.

As noted in the last section and in Figure 4, plasmaspheric density structure and plume formation are closely affected by plasmasphere-ionosphere coupling processes. Shortly after the plasmaspheric plume observation by *IMAGE EUV*, it was quickly realized that plumes are high-altitude extensions of storm-enhanced density (SED) signatures often seen in ionospheric TEC observations during storms [Foster *et al.*, 2002]. The equatorial fountain (Figure 4), long recognized to result from the *E*- and *F*-region dynamo (eastward electric field) driven by thermospheric neutral winds, may help supply the plume plasmas [Schunk and Nagy, 2009]. This is supported by recent model calculations showing that the equatorial fountain may reach nearly 1.3 *R_E* in altitude at the equator (as shown in Figure 5), well into the plasmasphere. After analyzing the “Bastille Day storm” using multiple satellite data, Lin *et al.* [2007] used Figure 6a to describe the plasma flows on the Earth’s corotating frame. The solid line near *L* = 2 is the plasmaspheric boundary in the afternoon and evening sectors. The boundary extends to higher latitudes near the plume. Arrows indicate subauroral polarization stream (SAPS) flows in the evening sector, initially westward, but diverted poleward due to the reduction of poleward electric fields inside the plasmasphere at the dusk terminator. The proposed flow pattern is consistent with Figure 6(b) that depicts the meridional view of the large-scale flows in which the equatorial fountain can be seen as a possible source of the plume plasmas.



MAGnetosphere-Ionosphere Connector (MAGIC) Mission Concept

Global *EUV* and *ENA* imaging carried out by the *IMAGE* satellite (see Figures 1 and 2) have revived interest in plasmasphere research and revealed the important roles the plasmasphere plays in multiple processes [Figures 4, 5, 6] that contribute to systemic M-I coupling. While global imaging can clearly provide the overall context of magnetospheric dynamics, it cannot replace the ability of *in situ* observations in yielding accurate details of plasmaspheric core structures (Figures 2 and 3) that are required for understanding the physics of individual coupling processes.

The *MAGIC* mission will employ coordinated global imaging, *in situ* space observations and ground-based measurements to provide the empirical data needed to investigate M-I coupling processes. Data assimilation, theory and modeling techniques will be used to validate the processes and integrate the knowledge to arrive at an understanding of systemic M-I coupling. *MAGIC* will consist of:

MAGIC Components	Instrumentations	Observational variables
1) High-altitude (≤ 9 Re), high-inclination ($\sim 90^\circ$) satellite; primarily for global imaging of aurora, plasmasphere and ring current	Imaging: EUV, FUV, ENA; Vector magnetic field; Plasma wave instrument Relaxation sounder	Global imaging of plasmasphere, auroral and ring current dynamics; magnetic field and ULF waves, plasma wave and electron density environment
2) Two or more lower-altitude (≤ 4.5 Re), lower-inclination ($\leq 50^\circ$) satellites, phased in time and local time; primarily for plasmaspheric “ground-truth” observations	Long-range radio sounder; Plasma wave instruments; Vector electric and magnetic fields; Ion mass spectrometer; Retarding potential analyzer; Langmuir probe	<i>Field-aligned</i> electron density profiles; Passive wave observations; Vector E & B field; Plasma electron density and temp Plasma ion density, composition, temp Energetic ring-current particles
3) Ground-based ionospheric “ground-truth” observations and extended local-time coverage	Ground magnetometer; All-sky camera; Incoherent scatter radar; Ionosonde/digisonde	ULF waves; Ionospheric density structures and profiles, temperature, flows, electric field, SAPS, SED, TIDs, TEC
4) Data assimilation, theory and modeling for science closure	CCMC, Virtual Modeling Repository (VRM), Guest Investigator Program	Model development and validation with data

Temporal resolutions of satellite observational variables:

Imaging: FUV & ENA ~ 1 -2 min, EUV ~ 5 -10 min (may be improved with technology)

Vector magnetic and electric fields: 0.1 s

Plasma wave: Swept-frequency receiver, 0.1 – 800 kHz, 2 min; waveform capture, 0.1-1 ms

Electron plasma density and temperature: 0.1 s

Ion plasma density and temperature: 0.1 s

Field-aligned electron density profile: 1-2 min

While *IMAGE* has confirmed plasmaspheric mass loss through large-scale plumes, it has also revealed numerous previously unknown structures in the plasmasphere boundary layer, *i.e.*, the plasmopause, such as notches, fingers and shoulders, and has clearly demonstrated the fibrous irregular nature of this boundary through the scattering and ducting of radio waves [Figure 3] and the ability to quantify plasmaspheric refilling rates based on *FAED* profiles obtained from radio sounding in hemisphere-to-hemisphere *FAED* irregularities [Reinisch *et al.*, 2004]. These observations indicate that much is unknown concerning the magnetosphere-ionosphere coupling processes [Lemaire and Gringauz, 1998; Darrouzet *et al.*, 2009]. Understanding these processes, and the important plasmasphere/ring-current interaction, will require *IMAGE*-type measurements in addition to *in situ* measurements of plasmaspheric thermal structure and ion composition.

A major advantage of *MAGIC* is that it will offer campaign-like simultaneous, multi-point observations throughout the mission. *In situ* measurements will always be analyzed within the context of

corresponding inner magnetospheric dynamic conditions from global images and in terms of ionosphere-plasmasphere-ring current interaction processes. This capability will address a major limitation of the *IMAGE* mission where, with only one satellite, the plasmaspheric hemisphere-to-hemisphere field-aligned electron-density profiles were separated in time from the plasmaspheric *EUV* and *ENA* images by many hours and no in-situ measurements, such as ion-composition, were available. *MAGIC* will provide simultaneous multi-point observations, including global *EUV*, *FUV* and *ENA* images, *FAED* profiles at multiple locations in the plasmasphere/plasmapause, ground-based observations, and will enable quantitative studies to be made of wave-particle interaction processes, plasmaspheric erosion and refilling, and other ionosphere-plasmasphere-ring current interactions. Data collected by *MAGIC* will allow proper posing of boundary conditions for developing and validating physical models.

Given that all proposed *MAGIC* instrumentations have well documented flight heritage (*OGO-5*, *Alouette-ISIS*, *DE*, *Polar*, *IMAGE*, *TWINS*, *Cluster*, etc.), there should be little or no cost associated with new instrument technology development. In addition, since only a relaxation sounder is needed for the high-altitude satellite and radio sounding on the lower-altitude satellites will only be operating near and within the plasmasphere, much shorter antennas than those on *IMAGE* will be needed for *MAGIC*. With the lower-altitude satellites being identically built, the rough estimated cost of *MAGIC* may be ~\$750M. If launched within the next 10 years, *MAGIC* can become a significant part of the next Heliophysics System Observatory and enhance the scientific returns of other *NASA* and international space missions such as *RBSP* (*NASA*), *GEC* (*NASA*), *ORBITALS* (Canadian), and *ERG* (Japanese).

The last *Solar and Space Physics Decadal Survey* (2003) already recognized the significant contributions and potential of global *EUV* and *ENA* imaging in addressing the system-coupling complexity confronting Heliophysics today. By combining multi-satellite observations, ground-based measurements and a robust theory and modeling program, *MAGIC* will implement effectively the recommendations of the *Panels on Solar Wind-Magnetosphere Interactions* (*SWM*), *Atmosphere-Ionosphere-Magnetosphere Interactions* (*AIM*), as well as *Theory, Modeling and Data Exploration* (*TMDE*). *MAGIC* will be well positioned to directly address the outstanding science questions enumerated in the *SWMI* and *AIM* Panel Reports. It will help mobilize the efforts and integrate the results of national (*ROSES*, *LWS*, *SHINE*, *GEM*, *CEDAR*, *Heliophysics Virtual Observatories-VxOs*) and international (e.g., *SuperDARN*, *ILWS*, *ISWI* [Davila et al., 2010]) research and instrumentation programs, and contribute directly to satisfying the national needs to advance the knowledge of Space Sciences and promoting a technically trained and increasingly diverse workforce, as well as a scientifically literate citizenry (*Education and Society Panel Report*).

References

- Carpenter, D. L., et al., Small-scale field-aligned plasmaspheric density structures inferred from RPI on *IMAGE*, *J. Geophys. Res.*, vol. 107, No. A9, SMP-22, 2002.
- Darrrouzet, F., J. De Keyser, and V. Pierrard, editors, *The Earth's Plasmasphere: A Cluster and IMAGE Perspective*, Springer, 2009.
- Davila, J. M., S. F. Fung, D. E. Rowland, R. F. Pfaff, T. Fuller-Rowell, *International Space Weather Initiative (ISWI)*, a white paper submitted to *Solar and Space Physics Decadal Survey*, November 12, 2010.
- Grebowsky, J. M. (1970), Model study of plasmaspheric motion, *J. Geophys. Res.*, 75, 4329.
- Gurgiolo, C., et al., (2005), Overlap of the plasmasphere and ring current: Relation to subauroral ionospheric heating, *J. Geophys. Res.*, 110, A12217, doi:10.1029/2004JA010986.
- Foster, J. C., P. J. Erickson, A. J. Coster, J. Goldstein, and F. J. Rich (2002), Ionospheric signatures of plasmaspheric tails, *Geophys. Res. Lett.*, 29(13), 1623, doi:10.1029/2002GL015067.
- Huang, X., B. W. Reinisch, P. Song, J. L. Green, D. L. Gallagher, Developing an empirical density model of the plasmasphere using *IMAGE*/RPI observations, *Adv. Space Res.*, 33, 6, 829-832, 2004.
- Lemaire, J. F., and K. I. Gringauz, *The Earth's Plasmasphere*, Cambridge University Press, New York, 1998.

- Lin, C. H., *et al.*, (2005), Theoretical study of the low- and midlatitude ionospheric electron density enhancement during the October 2003 superstorm: Relative importance of the neutral wind and the electric field, *J. Geophys. Res.*, 110, A12312, doi:10.1029/2005JA011304.
- Lin, C. S., *et al.*, (2007), Magnetospheric Convection near a Drainage Plume, *J. Geophys. Res.*, 112, A05216, doi:10.1029/2006JA011819.
- Reinisch, B. W., *et al.*, Plasmaspheric mass loss and refilling as a result of a magnetic storm, *J. Geophys. Res.*, Vol. 109, No. A1, A01202, 10.1029/2003JA009948, 2004.
- Reinisch, B. W., *et al.*, Augmented empirical models of plasmaspheric density and electric field using IMAGE and Cluster data, *Space Sci. Rev.*, 145, pp. 231-261, 2009; also appeared in *The Earth's Plasmasphere: A Cluster and IMAGE Perspective*, edited by F. Darrouzet, J. De Keyser, and V. Pierrard, Springer, 2009.
- Schunk, R., and A. Nagy, *Ionospheres*, 2nd edition, Cambridge University Press, New York, p. 388, 2009.
- Truhlik, V, et al., (2009), Relating OGO-5 H⁺ Plasmopause Transitions to Mid-Latitude Topside-Ionospheric Signatures, *Eos Trans. AGU*, 90(52), Fall Meet. Suppl., Abstract SA33A-1436.



SCINTILLATION AND ENERGY INPUT FOR SPACE SITUATIONAL AWARENESS AND MONITORING THE ENVIRONMENT (SESSAME)



Objective: The primary science objective of this new instrument suite, SESSAME, is to measure high latitude energy input and scintillation at both high and equatorial latitudes. Current capabilities are provided by the Defense Meteorological Satellite Program (DMSP) and Communication/Navigation Outage Forecasting System (C/NOFS) programs. For the future, we recommend development of a compact, modularized, and flexible suite of sensors that will measure the environmental characteristics required to forecast scintillation and determine energy input into the ionosphere with the necessary precision.

The proposed SESSAME architecture fits easily into the National Science Foundation CubeSat specifications. All sensors and the central control unit will be designed to occupy 1U or 2U cubes. The suite includes four key instruments that are fundamental for any Space Weather application in a new modularized design: a body-mounted no-boom magnetometer, a planar Langmuir probe (PLP), an ion velocity meter (IVM), and a precipitating particle sensor (LEESA2). The design is compact with a single interface to the spacecraft for all four sensors. Each instrument could also be flown separately on rides of opportunity. The overall SESSAME objective is to provide a new generation of space environmental monitoring instruments that measure key performance parameters with sufficient flexibility of design to meet ever-evolving Space Weather requirements.

Description: The SESSAME suite measures energy input at high latitudes and scintillation at both high and equatorial latitudes. Three of the instruments have strong and successful flight heritage, but the CubeSat versions are at different stages of maturity. The C/NOFS PLP was built in house at AFRL; a CubeSat version has been designed but not yet built. The C/NOFS IVM was built by the University of Texas at Dallas which is currently working on a CubeSat version. The proposed particle sensor LEESA2 is the CubeSat version of LEESA, a cousin of the DMSP SSJ5, that will soon fly on DSX. LEESA2, to be designed and built in-house at AFRL, is an electrostatic analyser with sufficient energy sensitivity and resolution to provide accurate determination of the auroral boundaries.

The magnetometer will be a revolutionary no-boom system that reduces integration costs significantly. AFRL has invested in this new technology through an SBIR which is currently in phase 1. Three contractors are developing designs for an inexpensive magnetometer that can yield scientifically useful data with no booms. The new system includes a highly sophisticated software package that will extract the geophysical signal from all the spacecraft magnetic noise.

Background: Key space environment measurements are (1) ionospheric scintillation, which can disrupt communication, navigation, and surveillance and (2) the electron density profile (EDP), which is a critical input for specification and forecast models. C/NOFS has taught us that *in situ* density and electric field are the two most important measurements for the low latitude ionosphere. These, combined with a model like AFRL's physics-based model PMod, contribute significantly to specification of the EDP and identification of scintillation regions in the equatorial ionosphere. Such complex global models also need accurate measurements of the

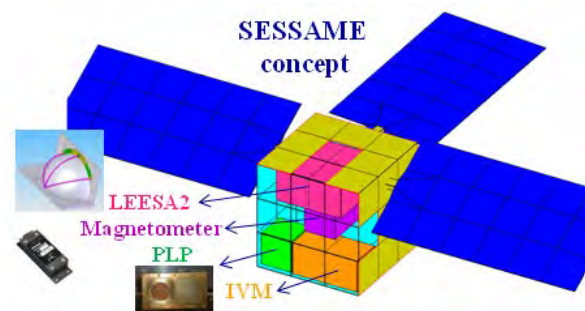
energy input from the magnetosphere to the ionosphere via a range of parameters, such as auroral boundaries and Poynting flux.

Poynting flux (determined from measurements of electric and magnetic fields) is the largest component of energy input into the ionosphere-thermosphere (IT) system during disturbed times by at least a factor of 2-5. A measure of the Poynting flux is thus essential for any meaningful forecast of space weather effects during storms or substorms. The thermosphere responds rapidly to energy input as Joule heating increases thermospheric neutral densities in the F layer. Knowledge of incoming Poynting flux is also required to predict neutral density changes that present significant challenges for satellite drag and orbital prediction models.

Advantages: SESSAME promotes and tests next-generation sensors with compact innovative designs, requires no booms eliminating associated high integration costs for electric or magnetic field measurements, and has a superior particle instrument. SESSAME's modular design allows installation in various configurations on rides of opportunity. SESSAME therefore:

- Provides an opportunity to raise the Technology Readiness Levels of next generation space environmental monitoring instruments that are critical to US Space Weather capability.
- Allows flexible deployment of either the full suite or individual sensors on CubeSats.
- Provides and tests a single interface, modular design for any combination of instruments.
- Includes a revolutionary body-mounted magnetometer system to eliminate costly integration and simplify bus design. We will use new technology, including sophisticated software packages under development through an SBIR (already in phase I), to integrate a magnetometer onto a spacecraft and separate useful geophysical signals from all the satellite magnetic noise.

While scintillation can be measured from the ground as it is happening, the measurements critically necessary to forecast it, namely the EDP and its spatial structure as well as the energy input boundary parameters, can be uniquely done only from space. Rocket or balloon experiments have measured auroral magnetic and electric fields and even particle precipitation, however, such experiments fly through the aurora too briefly to provide any meaningful data for Space Weather specification and forecasting. Although the Earth is reasonably well covered with ground-based magnetometers that can remotely sense ionospheric magnetic fields, they cannot sense magnetospheric field-aligned currents, and thus cannot determine the energy input to the ionosphere because of the ionospheric conducting layer. SESSAME has tremendous potential to contribute significantly to the future US space environment monitoring program.





EQUATORIAL FORECASTING OF SCINTILLATION

DESCRIPTION: The **CHIEFS** satellites are designed (1) to monitor irregularities that cause communication and navigation outages in the night-time low-latitude ionosphere and (2) to specify and forecast the ionosphere and neutral atmosphere to support orbital drag models. This capability is vital for every satellite in orbit to support mission development and planning, accurate tracking, maneuver planning, reentry prediction, and collision avoidance. Sensors include a thermal plasma suite, mass spectrometer, *in situ* neutral wind monitor, accelerometer, and vector electric and magnetic field instruments (Table 3). The four Equator spacecraft would fly in low inclination ($\sim 15^\circ$) circular orbits near ~ 400 km altitude, spaced $\sim 60^\circ$ apart. (C/NOFS orbit is elliptical at 13° inclination.) This configuration optimizes global coverage of ionospheric conditions and provides consistent measurements to support space weather models that are currently starved for data. The estimated on-orbit lifetime is $\sim 3+$ years with small-scale propulsion capability to counteract the high drag at low altitudes.

BACKGROUND: In this age of rapid growth in the capabilities of hand-held electronic devices that depend on satellite technologies, real-time characterization of the near-Earth space environment as well as reliable and timely space weather forecasts are more critical than ever. A constellation of small dedicated satellites orbiting the equator combined with data input from ground-based instruments would greatly advance our understanding of ionospheric effects on communication and navigation systems. This constellation of small equatorial satellites represents a low cost solution that meets many current space weather objectives and provides global measurements for ionospheric forecast models still starved for data.

Although initially planned for solar maximum, the Communication/Navigation Outage Forecasting System (C/NOFS) mission has proved its worth as a faithful observer of ionospheric dynamics at solar minimum. C/NOFS discoveries include: (1) deep plasma density depletions found from post-midnight to dawn, and (2) measured ionospheric parameters that are significantly different from any pre-mission model predictions. The implications of these discoveries have far-reaching effects on our efforts to understand the global dynamics of the magnetosphere-ionosphere-thermosphere system. Equatorial plasma bubbles (EPBs), usually observed after sunset and during magnetic storms when penetration electric fields are present, are frequently observed after midnight when overshielding electric fields are dominant. Even the altitude and range parameters of the ionosphere are different. The peak of the ionospheric F layer is 200 km lower than predicted, plasma density is quite sparse ($< 10^4 \text{ cm}^{-3}$), and the peak daytime ion temperature is 1000 K cooler than expected. These results demonstrate that very different dynamics are at work during solar minimum. Although C/NOFS results have led to vast improvements in ionospheric models, there is still much to be learned, especially throughout the upcoming transition toward solar maximum. With C/NOFS well into its third year of operations, planning a follow on mission is long overdue. This proposed constellation of four small equatorial satellites will provide data to specify and forecast the ionosphere, including the presence of irregularities that cause scintillation, and to predict the flow of energy into the ionosphere and thermosphere for more accurate satellite drag and neutral atmosphere models.



**CONSTELLATION FOR HELIOSPHERIC AND IONOSPHERIC
EQUATORIAL FORECASTING OF SCINTILLATION**



Instruments for Equator	Measurement
Thermal Plasma Suite (including IDM, RPA, PLP)	Drift velocity, Temperatures of ions & electrons, Density and density fluctuations
Mass Spectrometer	<i>In situ</i> total neutral density, neutral composition for NDP
Neutral Wind Monitor	3 components of the thermospheric wind
Accelerometer	Drag, total neutral density, winds
DORIS Receiver	TEC, EDP, scintillation
Vector Electric Field Instrument with Magnetometer	High sensitivity electric fields and waves

Table 1. Recommended instruments for Equator spacecraft.



POLAR AND EQUATORIAL COMMUNICATION OUTAGE SATELLITES (PECOS)



FOR MONITORING THE NEAR-EARTH SPACE ENVIRONMENT

INTRODUCTION: As we become more dependent on technologies that operate in space, real-time characterization of the near-Earth space environment as well as reliable and timely space weather forecasts are more critical than ever. Global networks of communication and navigation satellites provide much of the information we rely on to keep in touch with our families, communities, and the world. No one satellite, even one carrying a full complement of space environment sensors, can satisfy every need. However, a constellation of small dedicated satellites combined with data input from ground-based instruments and other space-based sensors on rides of opportunity could go a long way toward meeting these requirements. Small satellites are cost-effective platforms for innovative sensor designs and new technologies to respond to changing space weather needs. These three paired constellations of small satellites flying in multiple orbits represent a low cost solution that meets many of our current space weather objectives with specialized arrays of sensors on each spacecraft. PECOS is a flagship mission that enhances current capabilities to meet future needs.

RATIONALE: The US is facing an imminent loss in critical space environment monitoring capability with the approaching end of the DMSP mission. For more than thirty years, we have had multiple DMSP spacecraft in orbit providing key space environment measurements. Almost 700 scientific papers using DMSP data have been published in the past 20 years. To date, there is no existing program to replace the full DMSP capability. The C/NOFS mission is 2.5 years into a projected 3-year mission. We have been quite fortunate that conditions have been relatively benign in the past few years. How far can we push our luck? The PECOS constellations put spacecraft in the right orbits with optimal complements of sensors to maintain and enhance our space environment capabilities.

The **POLAR-HIGH** satellites will provide measurements to specify and forecast the ionosphere, including the presence of irregularities that cause scintillation of radio signals, and to predict the flow of energy into the ionosphere and thermosphere for more accurate satellite drag and neutral atmosphere models. Sensors include a thermal plasma suite, vector magnetometer, precipitating auroral particle spectrometer, DORIS receiver and GPS occultation system (Table 1). Two spacecraft are suggested for optimal value: (1) one in a circular, Sun-synchronous 98° inclined orbit at 650 km in a LT plane near the dawn/dusk meridian, and (2) one in a circular prograde 82° inclined orbit at 650 km whose ascending node precesses through all local times twice per year. This configuration provides consistent measurements to optimize global coverage of ionospheric conditions for space weather models that are currently starved for data. Projected mission lifetimes for each spacecraft are 3 to 6 years.

The **POLAR-LOW** satellites will provide thermospheric data needed to specify and predict satellite drag. This capability is vital for mission planning, accurate tracking, maneuver planning, reentry prediction, and collision avoidance. Sensors include a thermal plasma suite, mass spectrometer, *in situ* neutral wind monitor, accelerometer, and DORIS receiver (Table 2). The POLAR-Low spacecraft would both fly near ~ 400 km altitude in high inclination (60° - 80°) circular orbits that precess rapidly to sample all LTs. Estimated useful lifetime of each spacecraft is $\sim 3+$ years with small-scale propulsion capability to counteract the high drag at low altitudes. Once proof-of-concept is demonstrated, a minimum of two Polar-Low spacecraft should be in orbit at all times.

The **EQUATOR** satellites are designed to monitor irregularities that cause communication and navigation outages in the night-time low-latitude ionosphere and to specify and forecast the ionosphere and neutral atmosphere. Primary sensors include a thermal plasma suite, mass spectrometer, *in situ* neutral wind monitor, and vector electric and magnetic field instruments (Table 3) along with a DORIS receiver and

GPS occultation system. The Equator spacecraft should fly at ~400 km altitude in a low inclination (~15°) circular orbit spaced ~ 180° apart. (The C/NOFS orbit is elliptical at 13° inclination). Estimated on-orbit lifetime is ~3+ years with small-scale propulsion.

Instruments for Polar-High	Measurement
Thermal Plasma Suite	Drift velocity, temperatures of ions & electrons, density and density fluctuations
Vector Magnetometer with Star Camera on Boom	Magnetic field, field-aligned currents, Dst index
Precipitating Auroral Particle Spectrometer (SSJ6+)	30 eV - 30 keV, 50 keV - 1 MeV, dosimeter
DORIS Receiver	TEC, EDP, scintillation
GPS Occultation System	TEC, EDP, L-band scintillation

Table 1. Recommended instruments and critical measurements for the Polar-High constellation.

Instruments for Polar-Low	Measurement
Thermal Plasma Suite	Drift velocity, Temperatures of ions & electrons, Density and density fluctuations
Mass Spectrometer	<i>In situ</i> total neutral density, neutral composition for NDP
Neutral Wind Monitor	Three components of the thermospheric wind
Accelerometer	Drag, total neutral density, winds
DORIS Receiver	TEC, EDP, scintillation

Table 2. Recommended instruments and critical measurements for the Polar-Low constellation.

Instruments for Equator	Measurement
Thermal Plasma Suite	Drift velocity, Temperatures of ions & electrons, Density and density fluctuations
Mass Spectrometer	<i>In situ</i> total neutral density, neutral composition for NDP
Neutral Wind Monitor	3 components of the thermospheric wind
Vector Electric Field Instrument with Magnetometer	High sensitivity electric fields and waves
DORIS Receiver	TEC, EDP, UHF, VHF, L-band scintillation
GPS Occultation System	TEC, EDP, L-band scintillation

Table 3. Recommended instruments and critical measurements for the Equator constellation.

AURORAL FORMS AND THEIR ROLE IN THE DUNGEY CONVECTION CYCLE

Jesper W Gjerloev

Johns Hopkins University - Applied Physics Laboratory

Co-Authors/Supporters:

Shin Ohtani, Johns Hopkins University - Applied Physics Laboratory

Scott Bounds, University of Iowa

Craig Kletzing, University of Iowa

John Sigwarth, Goddard Space Flight Center

Robert Pfaff, Goddard Space Flight Center

Douglas Rowland, Goddard Space Flight Center

Robert A Hoffman, Goddard Space Flight Center

Contents:

Science Objective

Science Motivation

Required Measurements

Summary

1. Science Objective

Closing the Dungey global convection cycle (Dungey, 1961) is a fundamental space physics problem that all driven magnetospheres have to resolve. At Earth this cycle appears to be intimately tied to the electrodynamics of distinct meso-scale auroral forms seen in the nighttime sector. Remarkably, these forms are recurrent indicating that they reflect similarly recurring underlying electrodynamics of fundamental importance for the magnetosphere-ionosphere coupling. Auroras are also associated with field-aligned currents and plasma flows/convection, which are ultimately responsible for the auroral distributions and return flow from the nightside toward the dayside.

It has been 50 years since the Dungey model was first proposed yet the physics of how flux is compressed on the nightside and the flow initiated to return the flux to the dayside remains a fundamental outstanding question in magnetospheric physics. Moreover, it applies universally to any convecting magnetospheric system. This presents one of the last large gaps in our understanding of the solar wind-magnetosphere-ionosphere interaction and we encourage the Decadal Report Committee to emphasize the importance of this fundamental topic in the forthcoming report:

What is the role of meso-scale auroral forms in the Dungey global convection cycle?

In this brief paper we argue for the importance of the problem and show that while past missions have led to great discoveries they were not designed to provide the observations required to attack the problem.

2. Science Motivation:

Active aurora signal fundamental magnetospheric reconfigurations and intense localized energy transport to the ionosphere. However, we do not understand the physics that these signatures reflect because we do not know where field-aligned currents flow, where plasma flows occur, or where electromagnetic (EM) energy is conveyed relative to the auroral forms.

Global convection requires that nightside ionospheric (and magnetospheric) flux circulation be completed, but how it happens and whether it is always done the same way is not known. Because induced fields do not map from the magnetosphere to the ionosphere, there is not a one-to-one correspondence between ionospheric and magnetospheric flows. Nonetheless, nightside low-altitude flows reflect potential flows in the magnetosphere and, together with currents and auroras, reflect signatures of ways that the M-I system solves this basic problem.

Establishing where the ionospheric nightside plasma flows occur relative to auroral forms is of great significance because nighttime active auroras link to the magnetotail where Earthward convection occurs. Because the Earthward flow requires that flux tubes compress, the system must do work, and the flow is energetically unfavorable. Any astrophysical plasma in which a driven convection process requires a compressive flow may lead to the so-called “pressure catastrophe”.

Previous missions that imaged aurora (e.g. DE, Freja, Polar, and IMAGE) made major discoveries. From their observations we now realize that there are persistently recurring meso-scale (hundreds of kilometers) auroral forms with characteristic patterns and development. This in itself is remarkable but the consistent behavior from case to case must be a testimony to a corresponding consistency in the electrodynamics. As the auroral forms repeat, even in varying locations, the

underlying electrodynamics also repeats (Fujii *et al.*, 1994; Gjerloev *et al.*, 2000; Gjerloev *et al.*, 2007). Thus, the consistent behavior from case to case is evidence of a corresponding consistency in the electrodynamics.

Birkeland currents are particularly important because they provide the linkage between high and low altitudes. They also transfer energy and momentum between the magnetosphere and the ionosphere. Because the currents are conserved between high and low altitudes, they provide insight into the behavior of the magnetospheric generator and are important for quantitative comparisons with models. Unfortunately, one cannot determine the current density from magnetic field observations along a single track without making assumptions that are typically violated (e.g. Hoffman *et al.*, 1994). As a result, we currently have little if any firm knowledge of the currents.

3. Required Measurements:

Despite decades of auroral research, our limited understanding of auroral electrodynamics indicates that in order to unlock the secrets of these processes we need a new approach. There are two reasons why our understanding of electrodynamics in major auroral regions is limited: (i) The field-aligned currents are too structured to be determined from single-satellite measurements, and (ii) the range of locations at which auroral forms occur is comparable to or larger than the forms themselves. Consequently, we have been unable to determine the electrodynamics associated with these forms because the currents cannot be measured from single satellite tracks. Moreover, compilations of multiple satellite crossings in the absence of auroral imagery to use as an instantaneous reference frame smear out the electrodynamics. Fortunately, from the limited joint *in situ*/imaging analyses that have been performed, we know that auroral emissions provide a reference system for the electrodynamics (Fujii *et al.*, 1994; Gjerloev *et al.*, 2007). Thus, advancing our scientific knowledge hinges on our ability to unambiguously measure the currents and flows while simultaneously imaging the aurora to place the *in situ* observations in their auroral contextual reference.

Magnetic field data from individual LEO satellites have been used for decades to estimate Birkeland current intensities, but the analysis of single-satellite observations requires assumptions that break down spectacularly in the regions of interest – active meso-scale auroral forms. These assumptions on which the current density is based are well-known although often ignored. For data acquired from a single satellite, the currents must be assumed to be time-stationary and to extend infinitely in longitude. Considering the structure and dynamics of regions of interest, this approach is woefully inadequate. As a result, we have little if any firm knowledge of the currents in the vicinity of these repeatable auroral forms.

4. Summary:

In this brief paper we urge the Decadal Report Committee to include the fundamental science objective: What is the electrodynamics of convection transport associated with meso-scale auroral forms on the nightside? How do these electrodynamics contribute to closing the Dungey global convection cycle? We do so because of the universal importance of the science and argue that while past missions have led to great discoveries they have not been designed to provide the observations required to address this fundamental problem.

Advancing our understanding hinges on our community realizing the importance of the problem as well as obtaining the required observations. A key objective is the definitive,

instantaneous measurement - not estimation - of Birkeland current densities without the usual crippling assumptions. We must further establish the relationship of ionospheric plasma flows down to a scale of a few kilometers with the auroral emissions and Birkeland currents. Finally we need to characterize the electromagnetic energy input in active auroral regions and its relationship to the particle energy flux and ion outflows (e.g Strangeway *et al.*, 2005; Chaston *et al.*, 2007).

We finally list a few specific objectives that may be helpful for the committee.

- Where do the upward and downward field-aligned currents flow relative to the aurora?
- How intense are the currents, and how do they change in time?
- Do N-S aligned arcs have a N-S aligned up-down Birkeland current pair?
- Do N-S arc events carry the bulk of the ionospheric return flow?
- Are intense return convection throats associated with specific auroral forms?
- What is the EM energy flux for different auroral forms?
- Is EM energy flux dominated by large or small scales?
- Does intense EM energy input drive ion outflows?

References

- Chaston, C. C., C. W. Carlson, J. P. McFadden, R. E. Ergun, and R. J. Strangeway (2007), How important are dispersive Alfvén waves for auroral particle acceleration?, *Geophys. Res. Lett.*, 34, L07101, doi:10.1029/2006GL029144.
- Dungey, J. W., (1961), Interplanetary magnetic field and the auroral zones, *Phys.Rev.Lett.*, 6, 47.
- Frank, L. A., J. D. Craven, K. L. Ackerson, M. R. English, R. H. Eather, and R. L. Carovillano, Global auroral imaging instrumentation for the Dynamic Explorer mission, *Space.Sci.Instr.*, 5, 369-393, 1981.
- Fujii, R., R. A. Hoffman, P. C. Anderson, J. D. Craven, M. Sugiura, L. A. Frank, and N. Maynard, Electrodynamic parameters in the nighttime sector during auroral substorms, *J.Geophys.Res.*, 99, 6093,1994.
- Gjerloev, J. W., and R. A. Hoffman, Height-integrated conductivity in auroral substorms, 1. Data, *J.Geophys.Res.*, 105, 215, 2000.
- Gjerloev J. W., R. A. Hoffman, J. B. Sigwarth, L. A. Frank (2007), Statistical description of the bulge-type auroral substorm in the far ultraviolet, *J. Geophys. Res.*, 112, A07213, doi:10.1029/2006JA012189.
- Gjerloev, J. W., R. A. Hoffman J. B. Sigwarth , L. A. Frank, and J. Baker (2008), The typical auroral substorm: A bifurcated oval, *J.Geophys.Res.*,113, A03211, DOI: 10.1029/2007JA012431.
- Hoffman, R.A., R.Fujii and M.Sugiura, Characteristics of the field-aligned current system in the nighttime sector during auroral substorms, *J.Geophys.Res.*, 99, 21,303, 1994.
- R. J. Strangeway, R. E. Ergun, Y. Su, C. W. Carlson, and R. C. Elphic (2005), *J. Geophys. Res.*, 110, A03221, doi:10.1029/2004JA010829.

SUPERMAG: THE GLOBAL GROUND BASED MAGNETOMETER INITIATIVE

Jesper W Gjerloev

Johns Hopkins University - Applied Physics Laboratory, USA

Co-Authors/Supporters (from 9 countries):

Aaron Ridley, U. of Michigan, USA	Mervyn P. Freeman, BAS, UK
Alexander Zaitsev, IZMIRAN, Russia	Nicola J. Fox, JHU-APL, USA
Alexandre Koustov, U. Saskatchewan, Canada	Olaf Amm, FMI, Finland
Anthony Lui, JHU-APL, USA	Patrick Newell, JHU-APL, USA
Chris Paranicas, JHU-APL, USA	Peter Chi, UCLA, USA
Claudia Stolle, DTU Space, Denmark	Robert A Hoffman, NASA/GSFC, USA
Dave Sibeck, NASA/GSFC, USA	Robin Barnes, JHU-APL, USA
Donald H. Fairfield, NASA/GSFC, USA	Thomas Sotirellis, JHU-APL, USA
James Carbary, JHU-APL, USA	Toshihiko Iyemori, Kyoto University, Japan
Jian Yang, Rice University, USA	William K. Peterson, CU/LASP, USA
Kan Liou, JHU-APL, USA	Whitham D. Reeve, SAM Project, USA
Lars P Dyrud, JHU-APL, USA	William Denig, NOAA, USA
Marc Lessard, U. of New Hampshire, USA	Xiaoyan Zhou, JPL, USA
Mark Engebretson, Augsburg College, USA	Yudian Ma, Buaa University, China
Martin Connors, Athabasca University, Canada	Jurgen Watermann, jfwConsult, France
Matthew Fillingim, UCLA, USA	

Contents:

- Executive Summary
- What is SuperMAG
- Why SuperMAG
- Science Motivation
- Summary

1. Executive Summary

For decades ground based magnetometers have proven to be the workhorse of the magnetosphere-ionosphere (M-I) physics community and their importance is indisputable. The SuperMAG initiative (Gjerloev, 2009) is a collaboration between virtually all the worlds magnetometers thereby providing a powerful tool for continuous and global monitoring of the ionospheric and magnetospheric currents. In this paper we urge the Decadal Report Committee to acknowledge the strength of this data set and the need for global collaborations such as SuperMAG since they:

- a) provide a cost effective way to maximize the utilization of this comprehensive data set;
- b) enable studies of the structure and dynamics of the global ionospheric and magnetospheric current systems;
- c) can support current and future missions (e.g. the upcoming RBSP mission).

While investigations based on single stations and/or chains have led to great discoveries the next logical step is to combine all stations into a truly global network allowing us to further advance our understanding of the M-I system.

2. What is SuperMAG

SuperMAG is a worldwide collaboration of organizations and national agencies (Table 1) that currently operate more than 300 ground-based magnetometers. It provides easy access to validated measurements of magnetic field perturbations in the same coordinate system, with identical time resolution and a common baseline removal approach. Studies of the variations caused by electric currents flowing in the ionosphere and magnetosphere require a subtraction of the dominant and slowly varying Earth main field. Hence, both absolute and variometer data (data with unknown baselines) are included in SuperMAG.

SuperMAG utilizes vector measurements of the magnetic field, which represent a variety of file formats, temporal resolutions, units, and coordinate systems and are provided with or without baseline subtracted. SuperMAG resamples the raw data to 1-minute temporal resolution and converts all units into nanoteslas (nT). Artifacts and errors are removed by automated as well as manual correction routines. Data are then rotated into a local geomagnetic coordinate system and finally the baseline is subtracted by an automated technique.

3. Why SuperMAG

Before SuperMAG, global or even local studies required painstaking and labor-intensive data-handling, which effectively limited research. Analysts faced several inherent complications: confusing or even unknown coordinate systems, a multitude of data artifacts and errors, unknown baselines, and even difficulties obtaining data. These problems have resulted in a serious underutilization of data from magnetometers. With the introduction of SuperMAG, researchers, teachers, students and the public are provided with a practical and time effective means to study magnetic data (Figure 1).

Chains / Agencies / Organizations	
Intermagnet	Penguin
Greenland	Izmiran
Carisma	Kyoto WDC
Image	Autumn
SAMBA	ICESTAR
SAMNET	Athabasca U.
MEASURE	BAS
MACCS	RapidMAG
210 Chain	GIMA
AARI	USGS
DTU Space	DMI

Table 1. Current members of the constantly expanding SuperMAG family.

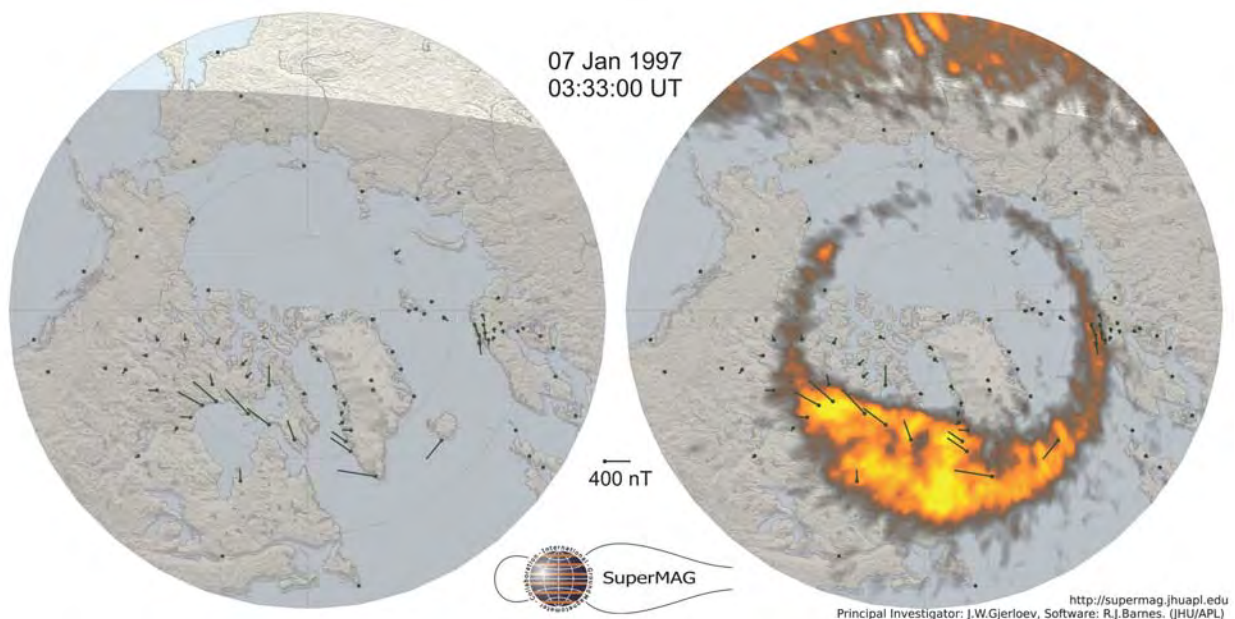


Figure 1. Ground based magnetometers provides global and continuous monitoring of the currents flowing in the ionosphere and magnetosphere. Powerful user-friendly data products are now available.

Beyond the research community, SuperMAG targets the general public—in particular teachers and students. This puts additional requirements on the site because these groups cannot be assumed to have extensive knowledge of either the dataset or the underlying physics. Consequently, the SuperMAG's Web Site is based on an intuitive interface with easily accessible tools and products.

The recent eGY provided an argument for the philosophy on which SuperMAG is based. The eGY declaration stated “*providing ready and open access to the vast and growing collections of cross-disciplinary digital information is the key to understanding and responding to complex Earth system phenomena that influence human survival*” (Baker et al., 2008). Although, humankind is likely to survive without magnetic field data, the philosophy behind the SuperMAG initiative certainly is in line with eGY's focus on data availability, access, and utilization.

Virtual observatories provide access to a wide variety of data acquired in the solar wind, in the magnetosphere, and on the ground. However, one of the fundamental assumptions of virtual observatories is that data and data description are available online. Most often this is not the case for ground based magnetometer data and as a consequence SuperMAG operates as a data repository obtaining the raw data from collaborators and performing all required data handling prior to the inclusion on the web-service.

4. Science Motivation

Ground magnetometers have three observational strengths that make their application in monitoring and understanding the M-I system particularly useful:

- a) Continuous uninterrupted monitoring;
- b) Nearly global coverage (see Figure 2);
- c) Decades of observations.

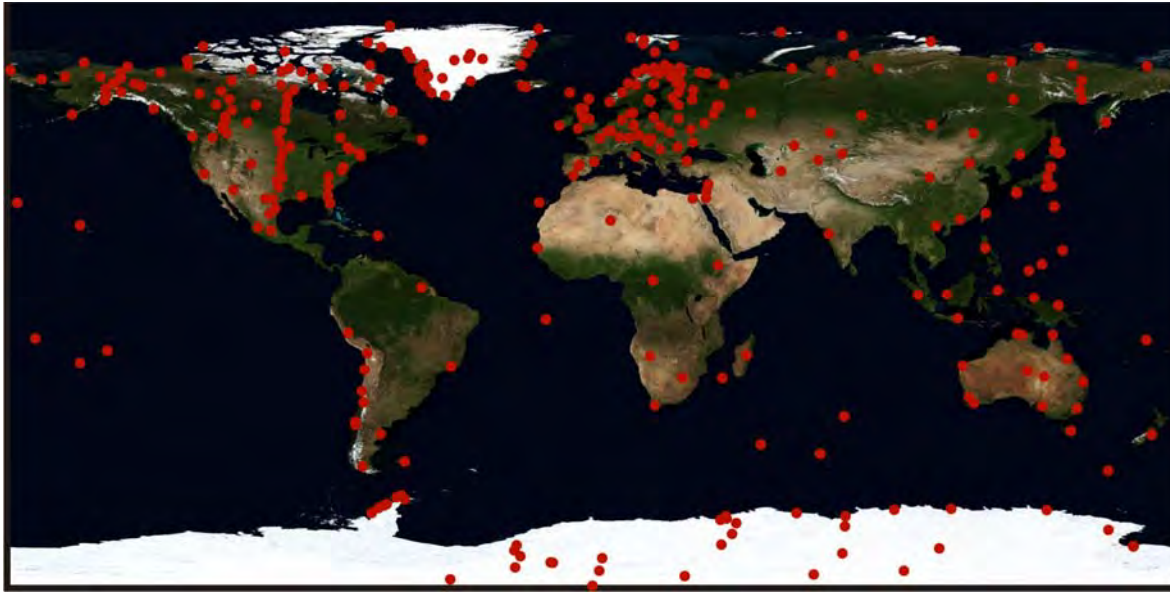


Figure 2. The world's ground based magnetometer community provides continuous and global monitoring.

For more than a century ground magnetometers have been the work-horse of M-I research. Popular indices such as Kp, AE and Dst are all derived from ground based magnetometer measurements and within the last two decades the use of ground magnetometer observations from sites worldwide has found application in estimating global scale electrodynamics by means of assimilation techniques (e.g. Richmond and Kamide 1988; Lu et al., 1996). Magnetometers have played a central role in a long list of published studies for example: the structure and behavior of the auroral electrojets system (e.g. DP2 current system, Nishida, 1971; Gjerloev et al., 2010; substorm growth phase, McPherron, 1970; two-component electrojets system, Kamide and Kokubun, 1996); traveling convection vortices (Lanzerotti et al., 1986; Friis-Christensen et al., 1988; Glassmeier et al., 1989); characteristic response time of the magnetosphere-ionosphere system to relatively discontinuous changes in the IMF (e.g. Murr and Hughes, 2001); and long-term behavior of the current systems (e.g. Kihn and Ridley, 2006; Ridley and Kihn, 2004). The list of discoveries and other scientific advances for which ground based magnetometer data played a key role is very extensive. In fact the above short list of publications and topics is by no means intended to be complete but simply to illustrate the historic success of magnetometers and the fact they continue to serve as an indispensable measurement.

We have no intention of providing an all-inclusive list of research topics and discoveries enabled by ground based magnetometer data but the Decadal Report Committee may find the following list helpful:

- “What are the differences and similarities of the ionospheric currents in the conjugate hemispheres and what is the cause of non-conjugate events?”
- “What is the spatiotemporal behavior of the auroral electrojets system during magnetic storms?”
- “What are the solar cycle variations of the ionospheric current system and its coupling to the magnetosphere?”
- “What is the role of the terminator in the magnetosphere-ionosphere current system?”
- “What storm-time dynamics do physics based models predict well and what do they miss?”
- “What is the spatiotemporal behavior of the equatorial electrojets?”

5. Summary

In this brief paper we urge the Decadal Report Committee to acknowledge the need for infrastructure initiatives such as SuperMAG. Ground based magnetometers have been the workhorse of M-I research for decades and combining all stations into a global continuous monitoring system was the logical next step. SuperMAG is truly an excellent example of the saying that “*the whole is greater than the sum of all parts*”. Investigations based on single stations and/or chains have led to great discoveries but further advancing our understanding of the M-I system dynamics and structure hinges on global collaborations.

References

- Baker, D. N., C. E. Barton, W. K. Peterson, and P. Fox (2008), Informatics and the 2007–2008 Electronic Geophysical Year, *Eos Trans. AGU*, 89(48), 485–486.
- Friis-Christensen, E., M. A. McHenry, C. R. Clauer, and S. Vennerstrom, Ionospheric traveling convection vortices observed near the polar cleft: A triggered response to changes in the solar wind, *Geophys. Res. Lett.*, 15, 253, 1988.
- Gjerloev, J. W., A Global Ground-Based Magnetometer Initiative, *EOS*, 90, 27, 230-231, 2009.
- Gjerloev, J. W., R. A. Hoffman, S. Ohtani, J. Weygand, and R. Barnes, Response of the Auroral Electrojet Indices to Abrupt Southward IMF Turnings, *Annales Geophysicae*, 28, 1167-1182, 2010.
- Glassmeier, K. H., M. Honisch, and J. Untiedt, Ground-based and satellite observations of traveling magnetospheric convection twin vortices, *J. Geophys. Res.*, 94, 2520, 1989.
- Kamide, Y. and S. Kokubun, Two-component auroral electrojet: Importance for substorm studies, *J. Geophys. Res.*, 101, 13,027, 1996.
- Kihn, E. A. and A. J. Ridley, A statistical analysis of the AMIE auroral specification, *J. Geophys. Res.*, 110, A07225, 10.1029/2004JA010775, 2005.
- Lanzerotti, L. J., L. C. Lee, C. G. MacLennan, A. Wolfe, L. V. Medford, Possible evidence of flux transfer events in the polar ionosphere, *Geophys. Res. Lett.*, 13, 1089, 1986.
- Lu, G., B. A. Emery, A. S. Roger, M. Lester, J. R. Taylor, D. S. Evans, J. M. Ruohoniemi, W. G. Denig, O. de la Beaujardiere, R. A. Frahm, J. D. Winningham, and D. L. Chenette, High-latitude ionosphere electrodynamics as determined by the assimilative mapping of ionospheric electrodynamics procedure for the conjunctive SUNDIAL/ATLAS 1/GEM period of March 28-29, 1992, *J. Geophys. Res.*, 101, 26697, 1996.
- McPherron, R. L., Growth phase of magnetospheric substorms, *J. Geophys. Res.*, 75, 5592, 1970.
- Murr, D. L. and W. J. Hughes, Reconfiguration Timescales of Ionospheric Convection, *Geophys. Res. Lett.*, 28, 2145, 2001.
- Nishida, A., DP 2 and polar substorms, *Planet. Space. Sci.*, 19, 205, 1971.
- Richmond, A. D. and Y. Kamide, Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique, *J. Geophys. Res.*, 93, 5741, 1988.
- Ridley, A. J. and E. A. Kihn, Polar cap index comparisons with AMIE cross polar cap potential, electric field, and polar cap area, *Geophys. Res. Lett.*, 31, doi:10.1029/2003GL019113, 2004.

Coupling through planetary waves: from the stratosphere to ionospheric irregularities

Larisa P. Goncharenko, Anthea J. Coster,
Haystack Observatory, MIT Haystack Observatory
Jorge C. Chau,
Jicamarca Radio Observatory

Summary

The presence of ionospheric irregularities adversely impacts communication and navigation systems. We propose observational strategy that would investigate potential effects of planetary waves on irregularities. If the concept is proven, it can serve as a path towards future capability of multi-day ionospheric forecast.

1. Summary of the science concept

Recent studies have demonstrated a relationship between the quasi-stationary planetary waves, equatorial $E \times B$ drift, and variations in the equatorial ionization anomaly (EIA), shown as Link A in Figure 1. Numerous studies also indicate a relationship between EIA gradients and ionospheric irregularities, shown as Link B in Figure 1. We are proposing to investigate Link C, e.g. the connection between the formation or suppression of irregularities and generation or suppression of the EIA related to the enhancement of quasi-stationary planetary waves.

Quasi-stationary planetary waves reach their maximum amplitudes prior to the specific wintertime meteorological phenomenon called stratospheric sudden warming (SSW) [O'Neill, 2003]. Currently, SSW events can be predicted up to ten days in

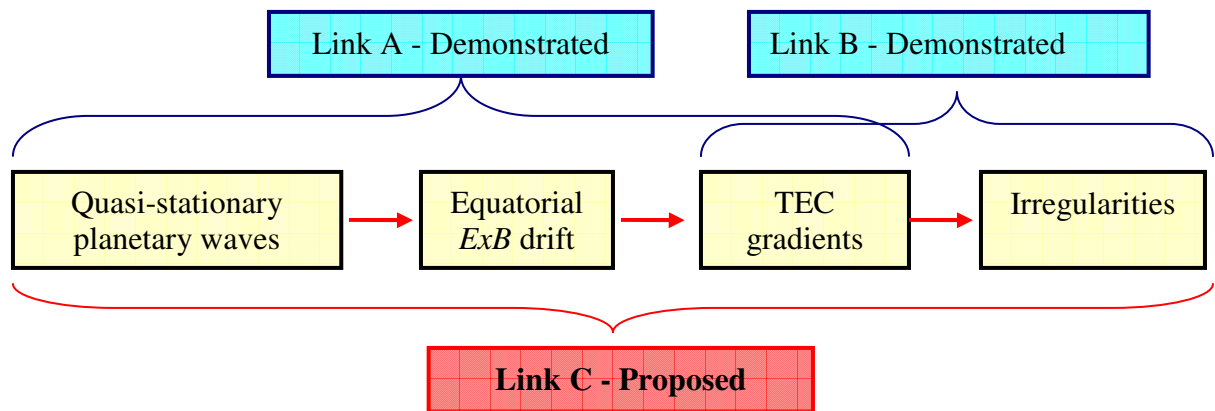


Figure 1. Block diagram of the proposed program. The relationship between quasi-stationary planetary waves, equatorial $E \times B$ drift, and resulting gradients in TEC has been demonstrated in recent studies. Separately, areas of large TEC gradients were shown to be associated with formation of irregularities. We propose to investigate the entire causal chain of ionospheric variability and irregularity morphology in response to enhanced planetary wave levels.

advance. If Link C can be established, this level of predictability eventually can lead to the development of a multi-hour and multi-day forecast of ionospheric behavior. The proposed effort addresses unresolved science problem (cause of day-to-day ionospheric variability), societal objective (reliability of communication systems), and operational objective (progress towards development of multi-day ionospheric forecast).

As a first step of the program, we propose to focus on the morphology of ionospheric variability and ionospheric irregularities around SSW events. If the link C is established, in the next step studies will be extended to other types of planetary waves, which are prevalent during other seasons.

2. Background: current state of knowledge in the field

2.1. Wave coupling of atmospheric regions

The dynamics and phenomenology of the quiet time ionosphere is affected by vertical coupling in the atmosphere-ionosphere system through atmospheric waves (gravity waves, tides, and planetary waves) that propagate from troposphere and stratosphere. While the vertical coupling processes driven by tidal forcing and the consequent dynamo electric fields are responsible for major phenomenology of the equatorial and low-latitude ionosphere, interactive processes involving gravity waves and planetary waves are believed to play a significant role in day-to-day and short-term variability. The effects of tidal and gravity waves and their control of ionospheric densities are a matter of active research. Examples of this include studies of multi-peak longitudinal structures in ionospheric parameters (e.g. *Immel et al.*, 2006; *England et al.*, 2006; *Luhr et al.*, 2007; *Fejer et al.*, 2008) or studies related to gravity waves [*Vadas*, 2007, *Vadas and Liu*, 2009, *Fritts et al.*, 2009]. Tidal and gravity wave effects are relatively better understood than effects originating from vertical coupling due to planetary waves. Multiple studies highlight the role of planetary waves as progenitors of lower altitude forcing on the ionosphere [*Forbes and Leveroni*, 1992, *Chen*, 1992, *Abdu et al.*, 2006, *Fuller-Rowell et al.*, 2008, *Pedatella and Forbes*, 2009]. These studies focus on ionospheric variations with planetary wave periods, i.e. 2-, 5-, 10- and 16-days. There is growing evidence that interaction of planetary waves with tides leads to increased tidal variability in the mesosphere-lower thermosphere region [*Hagan and Roble*, 2001, *Liu and Roble*, 2002, *Lieberman et al.*, 2004, *Liu et al.*, 2007, *Palo et al.*, 2007, *Forbes et al.*, 2008]. However, the implications of increased tidal variability in the F-region are not yet understood.

2.2. Ionospheric variations associated with SSW events

Series of recent studies have focused on ionospheric response to stratospheric sudden warmings, which result from interaction of strongly enhanced quasi-stationary planetary waves with zonal mean flow. During SSW events, the Jicamarca ionospheric radar, located at the magnetic equator, reported anomalous variations in vertical drift [*Chau et al.*, 2009], as shown in Figure 2. The distinctive feature of this variation is a strong semidiurnal signature in the drift.

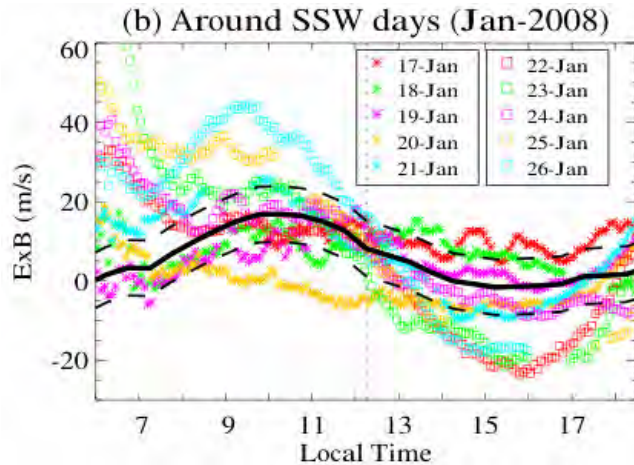


Figure 2. Daytime vertical ExB drifts measured over Jicamarca on January 17-26, 2008 (around SSW days), as function of local time. The expected average and standard deviation quiet-time values are indicated with black solid and dashed curves, respectively. A distinctive semidiurnal variation in vertical drift is observed during increase in stratospheric temperatures, with largest vertical drifts on Jan. 26, 2008. From *Chau et al.*, [2009].

The semidiurnal variation in equatorial vertical motion leads to a large-scale redistribution of electron density in the daytime ionosphere, as demonstrated by *Goncharenko et al.* [2010a] using total electron content (TEC) data measured by a global network of GPS receivers (Figure 3). An important aspect of the variation observed in TEC is its large magnitude, reaching 50-150% above and below the pre-SSW TEC behavior. Analysis of experimental data during other stratospheric warmings and from different instruments indicates that this phenomenon is a repeatable occurrence and is not a coincidence (*Anderson and Araujo-Pradere*, [2010], *Chau et al.*, [2009, 2010], *Goncharenko et al.*, [2010a,b], *Fejer et al.*, [2010], *Yue et al.*, [2010]). The ionospheric response to SSW events in general is characterized by a semidiurnal character, large amplitude, and persistence of perturbations for up to three weeks after the peak in high-latitude stratospheric temperatures.

Since the occurrence of SSW is related to the growth of planetary waves, observations of increased semidiurnal signature in the F-region vertical drift during SSW can be interpreted as direct evidence of modulation of the E-region dynamo by planetary wave. Numerical experiments with both the TIMEGCM model [*Liu et al.*, 2010] and the Whole Atmosphere Model [*Fuller-Rowell et al.*, 2010] demonstrate that when quasi-stationary planetary waves are introduced at the lower boundary at high latitudes, the low-latitude ionosphere experiences significant variations in both vertical ion drifts and total electron content. Although the quasi-stationary planetary waves do not propagate into the ionosphere or to low latitudes, nonlinear interactions of planetary waves with tides lead to large changes in tides [*Liu et al.*, 2010, *Fuller-Rowell et al.*, 2010]. Using GPS TEC data, *Pedatella and Forbes* [2010] presented experimental evidence that planetary wave-tide interaction is the primary mechanism coupling SSWs to ionospheric variability by demonstrating an increase in both migrating and non-migrating perturbations to the semidiurnal tide.

2.3. Electron density irregularities

Plasma irregularities are observed at all latitudes and altitudes in the ionosphere and exhibit a range of scale sizes, from centimeters to hundreds of kilometers. Irregularities act like a diffraction screen and are known to disrupt satellite communications and navigations systems, terrestrial communication links, ionospheric

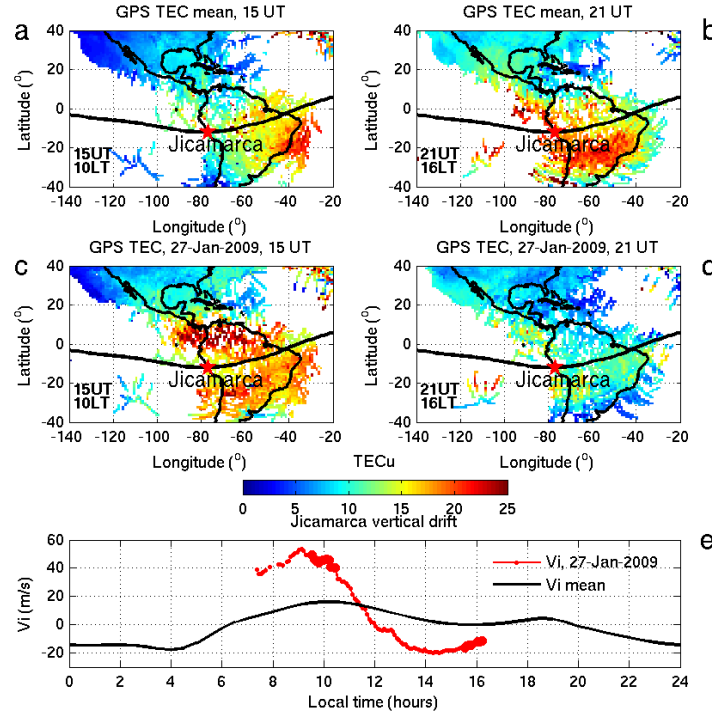


Figure 3. Observations of ionospheric behavior during stratospheric warming. (a) Typical distribution of total electron content (TEC) in the western hemisphere at 15 UT (morning sector, 10 LT at 75oW). Black line shows magnetic equator. (b) Same as Figure 1a, but for 21 UT (afternoon sector, 16 LT at 75oW). (c) TEC in the morning sector (15 UT) on Jan 27, 2009, during SSW. Increase in TEC is observed in the extended range of longitudes and latitudes. (d) TEC in the afternoon sector (21 UT) on Jan 27, 2009. In contrast to observations during morning hours, in the afternoon TEC is decreased. (e) Vertical drift observations by Jicamarca radar (12oS, 75oW). Red line – observations during SSW, black line – average drift for winter season. From *Goncharenko et al.*, [2010a].

sounders, and space-based remote sensing systems such as synthetic aperture radars. They are a major component of ionospheric space weather and must be understood if they are to be anticipated and mitigated.

Irregularities at the latitudes of the equatorial ionization anomaly (EIA) are often associated with regions of strong density gradients (*Basu et al.*, 2001, *Muella et al.*, 2008, 2010). Alternatively, collapse of EIA was linked to the suppression of equatorial plasma bubbles and scintillation (*Basu et al.*, [2009]). The vertical $E \times B$ drift is thought to be the primary factor controlling the peak plasma density and its gradients at low latitudes [e.g., *Anderson*, 1973]. A large vertical $E \times B$ drift is related to an enhancement of the EIA, while a smaller or downward drift leads to a suppression of the EIA. As the vertical drift prior to and shortly after sunset strongly influences the occurrence of F-region irregularities, understanding the causative mechanisms driving day-to-day variability in the drift has the potential for forming the basis of a forecast model for irregularities [*Basu et al.*, 1996, *Fejer et al.*, 1999].

We propose to investigate the link between quasi-stationary planetary waves, which have large amplitudes at stratospheric heights at high latitudes, and low-latitude ionospheric irregularities of different scales. Such a mechanism has not been indicated before as one of the possible causes affecting the onset and development of irregularities. This mechanism is expected to affect the day-to-day variability of the background ionosphere on the scale of hours and days, and thus to provide pre-conditioning important for wide variety of physical processes affecting the development of irregularities.

3. Suggested approach

We propose a cost-efficient observational strategy comprised of two stages. In the first stage, the research program will combine already existing predictive capabilities in the lower atmospheric community with a wide variety of instrumentation, techniques, and models currently available in the space physics community. In the stratosphere and up to the mesosphere-lower thermosphere, the key measured parameters will include temperatures (MLS, SABER on TIMED) and winds (data products from NCEP, ECMWF, or UKMO). In the MLT region, short-term variations in tides can be studied with available networks of ISRs, MF and meteor radars, and lidars. Observations of vertical drifts or electric field can be obtained by ISRs, magnetometers and C/NOFS. Ionospheric and thermospheric variability induced by planetary waves can be studied using such global datasets as GPS TEC, COSMIC, TIMED GUVI, CHAMP, C/NOFS, with inclusion of ground-based instruments. Variations in irregularities of different scales can be studied with C/NOFS PLP data, CERTO satellite radio beacon observations (150.012 MHz (VHF), 400.032 MHz (UHF) and 1066.752 MHz (L-band)), Jicamarca JULIA data (50MHz), and ionosondes (HF). Different aspects of the coupling mechanism can be investigated with capabilities of models like TIMEGCM, WAM, WACCM, SAMI, PBMOD. This list serves only as an example of existing capabilities, and many more potential contributors can be added. The first step has high degree of readiness including immediate availability of key needed resources. Collaborative work as a Focused Science Team can be used as a proof of concept, with total cost of \$5-7 mln for a 3-4 year program.

Due to the paucity of ground based instruments, accurate characterization of multiple tidal modes modified by planetary waves (e.g. migrating and non-migrating, 24-h, 12-h, 8-h) is impossible with current instrumentation. The final objective can be achieved through new satellite mission that includes measurements of multiple parameters (winds from the stratosphere to MLT to the upper thermosphere, electric fields, temperatures, electron density, irregularities of different scales). The cost of such mission will depend on inclusion of particular instruments and could vary between small (\$250-500 million) and medium (\$500-\$750 million).

References

- Abdu, M. A., P. P. Batista, I. S. Batista, C. G. M. Brum, A. J. Carrasco, and B. W. Reinisch (2006), Planetary wave oscillations in mesospheric winds, equatorial evening prereversal electric field and spread F, *Geophys. Res. Lett.*, 33, L07107, doi:10.1029/2005GL024837.
- Anderson, D. N. (1973), A theoretical study of the ionosphere F-region equatorial anomaly, II. Results in the American and Asian sectors, *Planet. Space Sci.*, 21, 421–442.
- Anderson, D., and E. A. Araujo-Pradere (2010), Sudden stratospheric warming event signatures in daytime ExB drift velocities in the Peruvian and Philippine longitude sectors for January 2003 and 2004, *J. Geophys. Res.*, 115, A00G05, doi:10.1029/2010JA015337.
- Basu, S., et al. (1996), Scintillation, plasma drifts, and neutral winds in the equatorial ionosphere after sunset, *J. Geophys. Res.*, 101, 26,795–26,809.
- Basu, S., et al. (2001), Ionospheric effects of major magnetic storms during the International Space Weather period of September and October 1999: GPS observations, VHF/UHF scintillations, and in situ density structures at middle and equatorial latitudes, *J. Geophys. Res.*, 106, 30,389–30,413, doi:10.1029/2001JA001116.
- Basu, Su., S. Basu, J. Huba, J. Krall, S. E. McDonald, J. J. Makela, E. S. Miller, S. Ray, and K. Groves (2009), Day - to - day variability of the equatorial ionization anomaly and scintillations at dusk observed by GUVI and modeling by SAMI3, *J. Geophys. Res.*, 114, A04302, doi:10.1029/2008JA013899.
- Chau, J. L., B. G. Fejer, and L. P. Goncharenko (2009), Quiet variability of equatorial ExB drifts during a sudden stratospheric warming event, *Geophys. Res. Lett.*, 36, L05101, doi:10.1029/2008GL036785.
- Chau, J. L., N. A. Aponte, E. Cabassa, M. P. Sulzer, L. P. Goncharenko, and S. A. Gonzalez (2010), Quiet time ionospheric variability over Arecibo during sudden stratospheric warming events, *J. Geophys. Res.*, 115, A00G06, doi:10.1029/2010JA015378.
- Chen, P. - R. (1992), Two - day oscillation of the equatorial ionization anomaly, *J. Geophys. Res.*, 97, 6343–6357, doi:10.1029/91JA02445.
- England, S. L., S. Maus, T. J. Immel, and S. B. Mende (2006), Longitudinal variation of the E - region electric fields caused by atmospheric tides, *Geophys. Res. Lett.*, 33, L21105, doi:10.1029/2006GL027465.
- Immel, T. J., E. Sagawa, S. L. England, S. B. Henderson, M. E. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton (2006), Control of equatorial ionospheric morphology by atmospheric tides, *Geophys. Res. Lett.*, 33, L15108, doi:10.1029/2006GL026161.
- Fejer, B. G., L. Scherliess, and E. R. de Paula (1999), Effects of the vertical plasma drift velocity on the generation and evolution of equatorial spread F, *J. Geophys. Res.*, 104, 19,859–19,869.

Fejer, B. G., J. W. Jensen, and S. - Y. Su (2008), Quiet time equatorial F region vertical plasma drift model derived from ROCSAT - 1 observations, *J. Geophys. Res.*, 113, A05304, doi:10.1029/2007JA012801.

Fejer, B. G., M. E. Olson, J. L. Chau, C. Stolle, H. Luhr, L. P. Goncharenko, K. Yumoto, and T. Nagatsuma (2010), Lunar-dependent equatorial ionospheric electrodynamic effects during sudden stratospheric warmings, *J. Geophys. Res.*, 115, A00G03, doi:10.1029/2010JA015273.

Forbes, J. M., and S. Leveroni (1992), Quasi - 16 day oscillation in the ionosphere, *Geophys. Res. Lett.*, 19, 981–984, doi:10.1029/92GL00399.

Forbes, J. M., X. Zhang, S. Palo, J. Russell, C. J. Mertens, and M. Mlynchzak (2008), Tidal variability in the ionospheric dynamo region, *J. Geophys. Res.*, 113, A02310, doi:10.1029/2007JA012737.

Fritts, D. C., Abdu, M. A., Batista, B. R., Batista, I. S., Batista, P. P., Buriti, R., Clemesha, B. R., Dautermann, T., de Paula, E. R., Fechine, B. J., Fejer, B. G., Gobbi, D., Haase, J., Kamalabadi, F., Kherani, E. A., Laughman, B., Lima, P. P., Liu, H.-L., Medeiros, A., Pautet, P.-D., Riggini, D. M., Rodrigues, F. S., Sao Sabbas, F., Sobral, J. H. A., Stamus, P., Takahashi, H., Taylor, M. J., Vadas, S. L., Vargas, F., and Wrasse, C. M.: Overview and summary of the Spread F Experiment (SpreadFEx), *Ann. Geophys.*, 27, 2141-2155, doi:10.5194/angeo-27-2141-2009, 2009.

Fuller - Rowell, T. J., et al. (2008), Impact of terrestrial weather on the upper atmosphere, *Geophys. Res. Lett.*, 35, L09808, doi:10.1029/2007GL032911.

Fuller - Rowell, T., F. Wu, R. Akmaev, T. - W. Fang, and E. Araujo - Pradere (2010), A whole atmosphere model simulation of the impact of a sudden stratospheric warming on thermosphere dynamics and electrodynamics, *J. Geophys. Res.*, 115, A00G08, doi:10.1029/2010JA015524.

Goncharenko, L. P., J. L. Chau, H.-L. Liu, and A. J. Coster (2010a), Unexpected connections between the stratosphere and ionosphere, *Geophys. Res. Lett.*, 37, L10101, doi:10.1029/2010GL043125.

Goncharenko, L. P., A. J. Coster, J. L. Chau, and C. E. Valladares (2010b), Impact of sudden stratospheric warmings on equatorial ionization anomaly, *J. Geophys. Res.*, 115, A00G07, doi:10.1029/2010JA015400.

Hagan, M., and R. Roble (2001), Modeling diurnal tidal variability with the National Center for Atmospheric Research thermosphere-ionosphere-mesosphere-electrodynamics general circulation model, *J. Geophys. Res.*, 106(A11), 24869-24882.

Lieberman, R.S., J. Oberheide, M. E. Hagan, E. E. Remsberg, and L. L. Gordley (2004) Variability of diurnal tides and planetary waves during November 1978-May 1979. *J. Atmos. Sol.-Terr. Phys.*, 66, 517-528.

Liu, H.-L., and R. G. Roble (2002) A study of a self-generated stratospheric sudden warming and its mesospheric-lower thermospheric impacts using the coupled TIME-GCM/CCM3, *J. Geophys. Res.*, 107(D23), 4695, doi:10.1029/2001JD001533.

- Liu, H.-L. et al. (2007) Comparative study of short-term diurnal tidal variability, *J. Geophys. Res.*, 112, D18108, doi:10.1029/2007JD008542.
- Liu, H.-L., W. Wang, A. D. Richmond, and R. G. Roble (2010) Ionospheric variability due to planetary waves and tides for solar minimum conditions, *J. Geophys. Res.*, doi:10.1029/2009JA015188.
- Luhr, H., K. Hausler, and C. Stolle (2007), Longitudinal variation of F region electron density and thermospheric zonal wind caused by atmospheric tides, *Geophys. Res. Lett.*, 34, L16102, doi:10.1029/2007GL030639.
- O'Neill, A. (2003), Stratospheric sudden warmings, in *Encyclopedia of Atmospheric Sciences*, edited by J. R. Holton, J. A. Pyle, and J. A. Curry, pp.1342-1353, Elsevier, New York.
- Palo, S. E., J. M. Forbes, X. Zhang, J. M. Russell III, and M. G. Mlynchak (2007), An eastward propagating two-day wave: Evidence for nonlinear planetary wave and tidal coupling in the mesosphere and lower thermosphere, *Geophys. Res. Lett.*, 34, L07807, doi:10.1029/2006GL027728.
- Pedatella, N. M. and J. M. Forbes (2009), Modulation of the equatorial F-region by the quasi-16-day planetary wave, *Geophys. Res. Lett.*, 36, L09105, doi:10.1029/2009GL037809.
- Pedatella, N. M. and J. M. Forbes (2010a), Evidence for stratosphere sudden warming - ionosphere coupling due to vertically propagating tides, *Geophys. Res. Lett.*, 37, L11104, doi:10.1029/2010GL043560.
- Vadas, S. L. (2007), Horizontal and vertical propagation and dissipation of gravity waves in the thermosphere from lower atmospheric and thermospheric sources, *J. Geophys. Res.*, 112, A06305, doi:10.1029/2006JA011845.
- Vadas, S. L., and H. Liu (2009), Generation of large-scale gravity waves and neutral winds in the thermosphere from the dissipation of convectively generated gravity waves, *J. Geophys. Res.*, 114, A10310, doi:10.1029/2009JA014108.
- Yue, X., W. S. Schreiner, J. Lei, C. Rocken, D. C. Hunt, Y. H. Kuo, and W. Wan (2010), Global Ionospheric Response Observed by COSMIC Satellites during the January 2009 Stratospheric Sudden Warming Event, *J. Geophys. Res.*, doi:10.1029/2010JA015466, in press.

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 $4\pi N \propto I/g$, 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 omnidirectional 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.

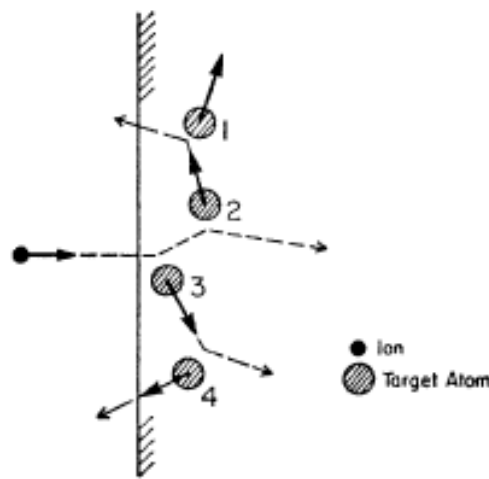


Figure 1: Ion sputtering schematic [11]

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 $4.4 \times 10^4 \text{ cm}^{-3}$. In this same paper, the sodium near-ground density is given as $1.7 - 3.8 \times 10^4 \text{ cm}^{-3}$, 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. O³⁺) the spectra are confined to the ultraviolet region from 195.86 to 3744.89 Å. Similarly, for O V (i.e. O⁴⁺) 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 (Å)	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	1.7-3.8 × 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_ν/I = 1.42×10^{-2} cm.
[†]I = 1.28×10^9 Ph cm⁻² s⁻¹ at 1 AU;
I_ν/I = 3.52×10^{-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_ν/I = 7.9×10^{-2} cm.
[§]I_ν = 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 1.8×10^6 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.

Magnetosphere Atmosphere Coupling Mission (MACM)

Introduction

Recent advances resulting from studies of the Earth's upper and lower atmosphere and magnetosphere have revealed the importance of the dynamic connection between these regions. The evidence is now overwhelming that any barriers that distinguish these regions are quite artificial. Energy, momentum and mass flow between the different regions is subject only to the varying importance of the physical processes at work. The influence of waves originating in the troposphere, on the upper atmosphere is now well established [Hagan *et al.*, 2007]. The influence of the potential of the ionospheric conducting layer on the electrification properties of the lower atmosphere is now receiving attention [Kniveton *et al.*, 2008]. The influence of variations in the ionospheric conductance on the electrodynamics of the magnetosphere is well established [Huba *et al.*, 2005] and the impact of mass flows between the ionosphere and the magnetosphere is actively being investigated. Overarching this coupled geospace system are the variable drivers originating directly from the Sun and propagating to the Earth through the solar wind.

Here we describe a space-based investigation to discover the spatial and temporal variability in the charged and neutral particle dynamics in the Earth's Atmosphere and its relationships to the variable inputs from the Sun that appear directly and are processed by the magnetosphere. This experiment called the Magnetosphere-Atmosphere Coupling Mission (MACM) must be able to accurately specify the imprint of magnetospheric and solar inputs on the atmosphere and accurately describe the response of the charged and neutral gas to those inputs.

Science Rationale

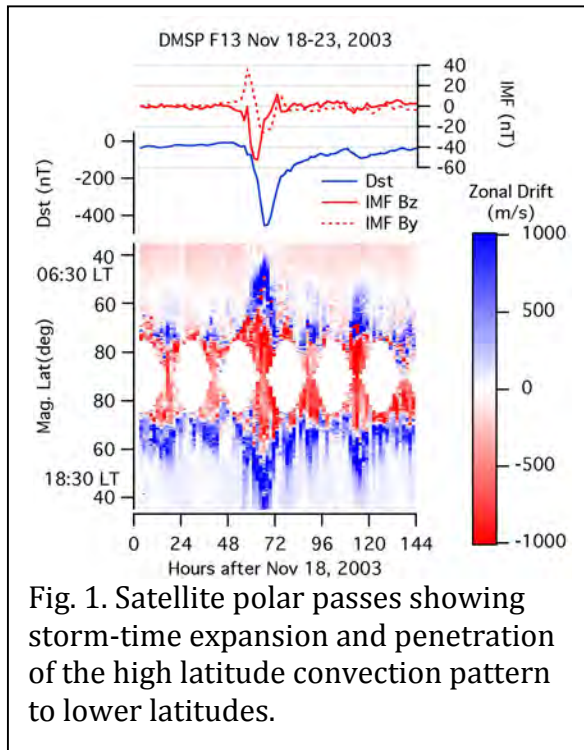


Fig. 1. Satellite polar passes showing storm-time expansion and penetration of the high latitude convection pattern to lower latitudes.

At high latitudes the ionosphere and thermosphere receive electromagnetic and particle energy inputs that can be characterized by the field-aligned Poynting flux and the flux of the precipitating charged particles. This energy is deposited over a range of spatial and temporal scales and is converted to heat and momentum through ion-neutral interactions. The ionospheric electric field across the polar cap and the field-aligned currents that deliver it, respond almost immediately to changes in the solar wind drivers. In the auroral zone the response time is controlled by other factors involving the motion of charged particles in the magnetosphere and the ionospheric conductivity and the neutral atmosphere dynamics [Wolf *et al.*, 2007]. But the ionospheric conductivity and neutral dynamics are themselves strongly influenced

by the dynamics resulting from the electric fields. Thus a complex coupling exists between the magnetosphere and the atmosphere over a range of spatial and temporal scales.

Figure 1 shows a sequence of profiles of the zonal ion drift from passes of the DMSP spacecraft across the northern polar region from 1830 to 0630 local time near 830 km altitude in the ionosphere. The transition between sunward (blue) and antisunward (red) flows clearly marks the convection reversal boundary. In the quiet period, from 0 to 48 hours, the equatorward edge of the strong westward auroral flows near dusk moves with the convection reversal boundary to slightly lower magnetic latitudes as the satellite passes through the convection pattern at slightly later local times. However, during a magnetic storm the convection reversal boundary may move to much lower latitudes. Near 70 hours for example, the convection reversal boundary appears near 60 degrees magnetic at dawn and near 55 degrees at dusk. This expansion of the convection reversal boundary is also associated with the equatorward movement of the auroral zone, which in this case extends to magnetic latitudes as low as 40 degrees with dramatic effects on the thermosphere.

While the storm-time expansion of the convection pattern is the most prominent feature in this figure it is important to notice that many other less dramatic expansion and penetration features exist. Some of these differences in consecutive latitude profiles of the ion drift are due to longitude variations, which repeat on a daily basis but are not well understood. Others are due to variations that appear in one pass and are absent in subsequent passes. ***These spatial and temporal variations are among the largest sources of variability in the ionosphere and thermosphere and cannot be examined with a single sample through a given volume.*** For example, Codrescu *et al.* [1995] have shown that temporal changes in the ion drift or electric field at high latitudes could change the estimated frictional heating rate by a factor of 2. But the heat deposited in the ionosphere-thermosphere and the momentum change to the neutral gas will depend on the temporal persistence of ion drift features with different spatial scale sizes.

At middle and high latitudes, electric fields created by the wind dynamo drive horizontal ion drifts that combine with those driven by external electric fields. The influence of neutral winds on the plasma temperature and the neutral and ion composition cannot be neglected, and their ability to modify the field-aligned currents must also be accounted for [Thayer, 2000]. While, it may be generally concluded that the effects of winds on the horizontal ion drift at high latitudes are small [Ridley *et al.*, 2003] it is not true that the neutral atmosphere is unaffected. Indeed, frictional heating can modify the neutral circulation patterns [Roble *et al.*, 1982] and may influence the Poynting flux delivered to the atmosphere from the magnetosphere [Gary *et al.*, 1994]. However, ***the relationships between the convection electric field, the ionospheric conductance, the neutral wind dynamo, and the solar wind drivers are poorly understood.***

Impulsive heating generates traveling atmospheric/ionospheric disturbances that propagate globally. The time scales over which these modifications take place can vary from minutes to hours and should be dependent on the magnitude and location of the energy inputs and the dynamic state of the system. ***The ability to specify the energy input and view the dynamic state of a volume at middle and high latitudes over time periods ranging from less than one minute to many tens of minutes is necessary for a quantitative assessment of the system response.***

These parameters will be functions of local time, longitude and season. Complete profiles of the latitude distribution of electric fields and currents at all local times and under a variety of conditions in the interplanetary medium are required to fill this knowledge gap. Even in the absence of storm-time perturbations in the high latitude convection pattern the offset between the geographic and geomagnetic poles produces very strong longitude and local time variations in the ion drift at auroral and sub-auroral latitudes. ***Only with measurements of winds and plasma drifts with sample periods less than an orbit period is it possible to describe the role of ion-neutral interactions in the overall dynamics.***

At middle and low latitudes the influence of winds and external, so-called penetration electric fields on the charged particle dynamics is also fundamental. Always present are electric fields dominated by tidal oscillations generated in the thermosphere by daily solar heating, and also at lower altitudes, from where they propagate into the E region. These influences are most evident in the zonal ion drifts at middle latitudes that are predominantly westward by day and eastward by night and in the zonal and vertical drifts at equatorial latitudes that also display a predominantly diurnal variation [Fejer et al., 1991]. In addition, the dynamo action of winds in the F region can be important at night. Finally, even at magnetically quiet times models show significant leakage of magnetospherically generated fields from high to low latitudes [Richmond et al., 2005]. ***Adequate specification of the behavior of the plasma in this region requires an understanding of the interplay between these dynamo fields and penetration fields.***

Figure 2 illustrates a series of latitude profiles of the zonal ion drift obtained from the DMSP F13 satellite during the course of a magnetic storm. These profiles, separated by approximately 100 minutes in UT and 25 degrees in longitude show dramatic changes at high, middle and low latitudes. Notice however, that the initial retreat of the auroral drifts seen in the first two profiles occurs instantaneously as does the expansion seen in the last two profiles. From this data set it is simply not possible to describe the evolution of the convective flow or to constrain the time scales over which it takes place. Yet such a description and constraints are required to understand the role of penetration fields and disturbance dynamo fields. It is possible to differentiate between these two physical mechanisms from knowledge of the temporal evolution of these features, since the polarization field features

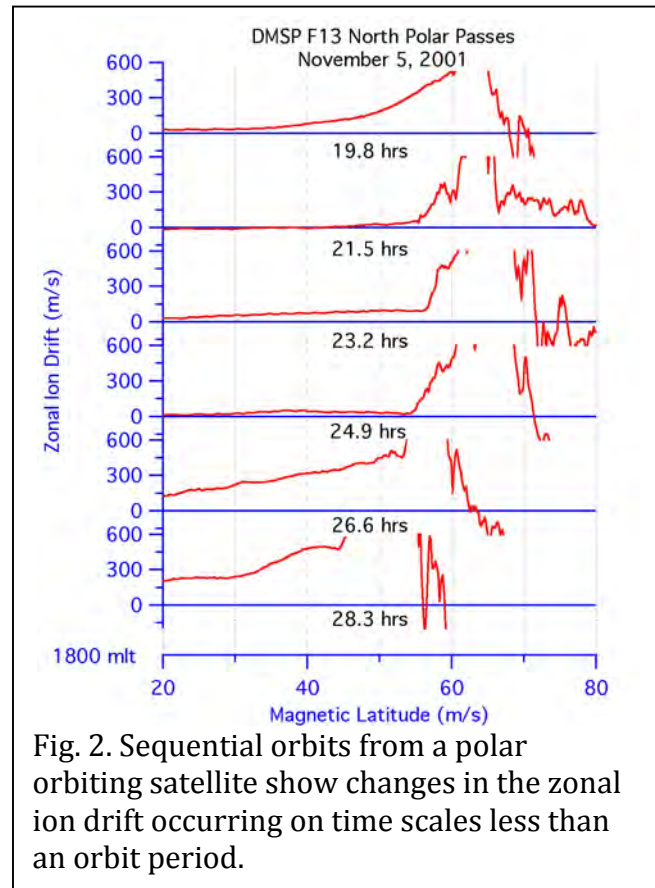


Fig. 2. Sequential orbits from a polar orbiting satellite show changes in the zonal ion drift occurring on time scales less than an orbit period.

should evolve over shorter time scales than those driven by the neutral wind. However *the evolutionary time scales are clearly less than an orbit period and thus at least two spacecraft following each other are required to make progress in understanding the role of these mechanisms.*

Mission Goals

- 1) To discover the spatial and temporal scales over which the ionosphere and thermosphere respond to magnetospheric energy inputs.**
- 2) To determine how magnetospheric sources and wind dynamos contribute to the electric field in the ionosphere and thermosphere.**

Mission Approach

One key to discovering the response of the atmosphere-ionosphere system is to magnetospheric inputs lies in discovering how the key parameters of temperature, composition and motion vary with latitude. This can only be accomplished with satellite borne measurements that traverse the region from the pole to the equator. In low Earth orbit a pole to equator transit will occupy about 25 minutes, during which time the longitude will change by about 6 degrees. A single satellite will revisit the same latitude range some 90 to 100 minutes, and in a longitude region that is shifted by about 25 degrees. This is much too long in time and space to discover the temporal persistence of key features or to quantify the change in the magnetospheric inputs. Two identically instrumented satellites in approximately the same orbit will however allow a temporal decomposition of the region with a resolution determined by their spatial separation.

In a given sample volume changes may result from internal process or from the motion of features through the volume. Satellites in approximately the same orbit can easily identify features that are moving meridionally. However, in order to quantify the zonal motions or to verify that they are small it is necessary to sample neighboring longitudes at the same time. This could be accomplished by a third satellite with the same orbit characteristics as the previous satellites but in an orbit plane that is longitudinally displaced from them.

Required Measurements

Both the state variables that specify the atmosphere and ionosphere and the imprint of magnetospheric inputs will need to be required.

Magnetospheric Inputs above the atmosphere

Energy and pitch angle distribution of the precipitating ions and electrons 10 eV to 50 KeV

Vector electric field.

Vector magnetic field

Ionospheric State Variables near and above the F-peak

Total Ion Concentration

Major Ion Composition

Ion and Electron Temperature

Ion Drift Vector

Neutral Atmosphere State Variables from 100 to 350 km

Neutral Density

Major Constituent Composition O/N₂

Neutral Temperature

Neutral Wind Vector

Instrumentation/Payload

Candidate instrumentation exists to make all the required measurements. Most demanding is the requirement for neutral density and wind measurements over a range of altitudes. While the altitude distribution of the ion density will be readily available from a global distribution of ground and space based GPS receivers, specification of the neutral atmosphere will be the highest mission priority.

A combination of remote and in-situ sensing is desirable. The challenge will be to accommodate the required measurements on a spacecraft for which the apogee and perigee will change during the mission to vary the spatial separation of the measurement platforms.

The spacecraft must accommodate seven or eight instrument packages.

Fields

1. Vector Electric Field (could be redundant with ion drift measurement depending on temporal cadence required)
2. Flux gate Magnetometer

Charged Particles

3. Electrostatic Analyzer for Medium Energy (10eV – 50 KeV) Pitch angle and energy
4. Ion Velocity Meter for thermal plasma density composition and drift.

Neutral Particles

5. Neutral Mass Spectrometer (may not be required depending on neutral wind instrument capability)
6. Neutral Wind Meter

7. FUV limb imager for neutral density profile
8. Doppler Michelson Interferometer or Fabry-Perot Interferometer for wind profile

Mission Implementation

In order to maintain a specified orbit period and a specified satellite separation around the orbit, each spacecraft will require propulsion. Observations should be made from a nominally circular orbit near 400 km. Propulsion is utilized to maintain the orbit near a constant pressure surface so that a mission lifetime of 2 to 4 years can be protected. Optimizing the orbit altitude in accord with the atmospheric pressure also preserves the ability to make in-situ measurements of the thermal plasma and the neutral wind and density.

In addition, the propulsion is utilized to change the ellipticity of a given orbit in order to adjust the spatial separation between the spacecraft. With very small changes in ellipticity, it is possible to vary the effective separation of the spacecraft from a few seconds (20 km) to a few minutes (1000 km) to many minutes (10000 km).

Orbits should be accomplished at high inclination near 75 degrees. This allows access to the auroral oval and the polar cap in the northern and southern hemispheres while providing a precession rate that allows all local times to be sampled in less than 4 months.

References

- Codrescu, M. V., T. J. Fuller-Rowell, and J. C. Foster (1995), On the importance of E-field variability for Joule heating in the high-latitude thermosphere, *Geophys. Res. Lett.*, **22**(17), 2393–2396.
- Fejer, B. G., E. R. de Paula, S. A. González, and R. F. Woodman (1991), Average Vertical and Zonal F Region Plasma Drifts Over Jicamarca, *J. Geophys. Res.*, **96**(A8), 13,901–13,906, doi:10.1029/91JA01171.
- Gary, J. B., R. A. Heelis, W. B. Hanson, and J. A. Slavin (1994), Field-aligned Poynting Flux Observations in the High-latitude Ionosphere,, *J. Geophys. Res.*, **99**, 11417.
- Hagan, M. E., A. Maute, R. G. Roble, A. D. Richmond, T. J. Immel, and S. L. England (2007), Connections between deep tropical clouds and the Earth's ionosphere, *Geophys. Res. Lett.*, **34**, L20109, doi:10.1029/2007GL030142.
- Huba, J. D., G. Joyce, S. Sazykin, R. Wolf, and R. Spiro (2005), Simulation study of penetration electric field effects on the low- to mid-latitude ionosphere, *Geophys. Res. Lett.*, **32**, L23101, doi:10.1029/2005GL024162.
- Kniveton, D.R., B.A. Tinsley, G.B. Burns, E.A. Bering, O.A. Troshichev (2008), Variations in global cloud cover and the fair-weather vertical electric field, *J. Atmos. Sol. Terr. Phys.*, **70**, 13, 1633-1642, DOI: 10.1016/j.jastp.2008.07.001.
- Ridley, A. J., A. D. Richmond, T. I. Gombosi, D. L. De Zeeuw, and C. R. Clauer (2003), Ionospheric control of the magnetospheric configuration: Thermospheric neutral winds, *J. Geophys. Res.*, **108**(A8), 1328, doi:10.1029/2002JA009464.
- Roble, R.G., E.C. Ridley, and R.E. Dickinson (1982), Global circulation and temperature structure of the thermosphere with high latitude convection, *J. Geophys. Res.*, **87**, 1599-1614.
- Thayer, J. P. (2000), High-latitude currents and their energy exchange with the ionosphere-thermosphere system, *J. Geophys. Res.*, **105**(A10), 23,015–23,024.
- Wolf, R.A., R.W. Spiro, S. Sazykin, F.R. Toffoletto (2007), How the Earth's inner magnetosphere works: An evolving picture, *J. Atmos. Solar Terr. Phys.*, **69**, 3, 288-302, DOI: 10.1016/j.jastp.2006.07.026.

Space-Atmosphere Boundary Layer Electrodynamics (SABLE)

Introduction

At the very edge of Earth's atmosphere lies a region that marks the transition from the neutral atmosphere to the plasma state of space, called geospace. In this threshold region both neutral and charged gases coexist, interact and come in contact with the magnetic and electric fields and energetic particles of space. This region is an electrically charged planetary boundary layer that processes almost all of the electromagnetic and particle energy resulting from our planet's interaction with the heliosphere.

Despite the proximity of this layer to the surface of the Earth and its crucial role in moderating the interaction of the atmosphere with the magnetosphere, it remains one of the least measured and most poorly understood regions of our geospace environment.

The SABLE mission will explore the role of this boundary layer in the coupling between the overlying space environment and our planet's atmosphere.

The physical processes that characterize the ionosphere-thermosphere boundary layer are dominated by interdependent interactions of the ionized and neutral gases and by distant, coupled, source regions. Significant energy is transferred into this region by the particles, fields, and currents of the high-altitude magnetosphere as well as by upward-propagating tides and waves from the lower atmosphere. The I-T boundary layer generates an upward feedback into the magnetosphere and modulates influences of the space environment on the lower regions. The major features of this boundary layer are illustrated in figure 1.

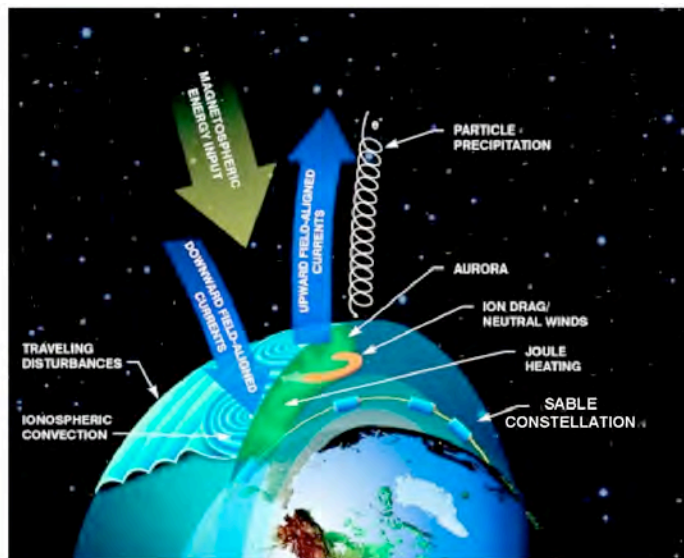


Figure 1. Schematic illustration of major inputs and response of the Space- Atmosphere Boundary Layer.

The major features of this boundary layer are illustrated in figure 1.

Mars and Venus, the most Earth-like planets in the solar system, have a contact region like the Earth's except for one very important difference: neither of these planets has a large intrinsic magnetic field. Resolving these processes in the Earth's I-T boundary layer is the primary goal of the SABLE mission.

Science Priorities

First the full effects of collisions between the ion and neutral gases must be understood. Then the relationships between different scales in the ion and neutral gases can be assessed, leading finally to an ability to uncover how the many different scales contribute to the exchange and redistribution of energy within and through the region.

1) Ion-neutral collisional interaction.

How do the collisionally coupled charged and neutral atmospheric constituents respond to electrodynamic input imposed from space?

Currently our knowledge of the magnitude and scale-size of the dynamic interaction below 200 km altitude is insufficient to describe the response of the system to the complex electric fields and particle precipitation imposed on our upper atmosphere. In this region there exist vertical and horizontal spatial gradients that can only be sensibly resolved with in-situ measurements. Thus, to close this knowledge gap, the SABLE mission must:

- Measure ion and neutral drifts, and neutral composition and density, over an altitude region where the difference between the ExB direction and the ion drift velocity can be determined.
- Measure the electric field and the energetic particle distribution to specify the input drivers and their spatial structure.
- Use measurements of ion and neutral particle dynamics, composition, and density, to determine the ion-neutral collision frequency.
- Measure the temperature and density of the ion and neutral species to locate variations of pressure that modify neutral circulation.
- Describe the changes in horizontal circulation at different altitudes in the lower thermosphere.
-

The electrodynamic connection of the Earth's atmosphere to space is determined by the collisional interaction of the charged gases of the lower ionosphere and neutral gases of the lower thermosphere. Within this Ionosphere-Thermosphere region (I-T region), heat and momentum are exchanged between the ionized gas and the neutral gases. Heating alters the pressure profile of the neutral gas, which modifies the neutral circulation and subsequently the momentum exchange. Momentum transfer modifies the ion-neutral velocity difference, which in turn modifies the electrical conductivity and the heating rates.

Among the least well-known, yet most critical, factors that control these non-linear interactions is the ion-neutral collision frequency. The absolute magnitude of heating and momentum transfer rates depend critically on this parameter, as does the electrical conductivity of the region. Without in-situ measurements of the various electrodynamic parameters in the altitude region where the collision frequency significantly affects the observations, it can only be inferred from estimates of the atmospheric composition and density, and laboratory measurements of collision cross sections, which are obtained under pressures and temperatures that differ significantly from I-T conditions. Hence, a critical element of the SABLE mission will be excursions of the orbit to altitudes where the appropriate measurements can be obtained.

2) Scales of electrodynamic drivers and responses:

How are the spatial variations and persistence of the electrodynamic drivers related to the neutral responses?

Spatial and temporal variations of the electrodynamic drivers incident on the I-T system can produce a wide range of responses that depend on the scales and persistence of those variations. Responses are strongly dependent on altitude due to the changing influence of the magnetic field

and the ion-drag force with changing altitude. Models indicate that localized interactions can lead to localized heating that will in turn generate upwellings of the neutral gas and atmospheric waves, which carry some portion of the incident energy away from the region where it is deposited. More distributed or more gradual interactions should lead to more distributed responses that are likely to produce large scale horizontal motions of the neutrals gas (winds) rather than upwelling. Our current ability to address these spatially and temporally varying inputs and responses is severely limited by an inadequate knowledge of the significant scales. Parameterizations in models are based upon an inadequate knowledge. To close this knowledge gap the SABLE mission must:

- Measure the electric field and energetic particle flux with a sample rate that allows resolving static structures with scale-size of 1 km and greater.
- Measure the ion and neutral drifts with a resolution of 1 km over an altitude region where the difference between the ion drift direction and the “ExB” direction can be determined.
- Determine the persistence and correlation of the inputs and responses by obtaining the measurements with temporal spacings ranging from 10 seconds to 30 minutes.

3) Evolution and feedback.

How, and under what conditions, do the responses evolve and feedback on the drivers?

Electrodynamic inputs to the system produce responses in the neutral and ionized gasses that differ in both spatial and temporal scales because of the greater influence of inertia and viscosity in the neutral gas. The neutral response may extend well beyond the period and spatial region over which the input is delivered. Inertia in the neutral atmosphere can maintain an imprint of the input that is conveyed back to the ionized gas in a flywheel effect. Thus, while currents may be driven in the region from outside influences, they are modified and at times maintained by the response of the system. But, currents must flow in closed loops such as those connect the external driver regions to the I-T system, leading to the ability of the atmospheric response to modify the driver. Currently, we have little information on these influences and how they evolve. Understanding them, however, is critical to understanding the interaction between the atmosphere and the geospace environment. To rise to this challenge the SABLE mission must:

- Measure electric and magnetic fields to distinguish between changes in energy flow produced by changes in the electric fields from changes in currents.
- Measure the neutral wind, ion drift, and electric fields, in-situ with temporal separations from 10 seconds to 30 minutes.
- Correlate the rate of change of the electric field and energetic particles with changes in the ion and neutral drifts to expose the influence of feedback from the atmosphere.

Required Measurements

Table 1 summarizes the key mission requirements to accomplish the SABLE science objectives. For each geophysical parameter, the range is dictated by its expected geophysical variability. This variability results from the need to sample over all local time regions and over an altitude region where the ion-neutral coupling varies considerably. The accuracy is dictated by the magnitude of changes associated with the magnetospheric drivers and the I-T system response. For example, to measure changes in the field aligned currents will require a sensitivity in the measurement of magnetic field perturbations of 10 nT. Changes in vertical neutral winds resulting from localized Joule heating will require a sensitivity of a 5 m/s.

Science Parameter	Science Goal	Measurement Requirements		
		Range	Accuracy	Temporal Res.
Electric Fields (vector)	Magnetospheric and Dynamo Influences	± 300 mV/m	0.25 mV/m	0.1 sec
Ion Velocity (vector)	Parallel and Perp Plasma Motions	± 6 km/s	5 m/s	1 sec
Neutral Velocity (vector)	Neutral Circulation	± 2 km/s	5 m/s	1 sec
Magnetic Field (vector)	Field-aligned Currents	± 65000 nT	10 nT	0.1 sec
20 eV – 300 keV (all pitch angles, 20° res.)	Electron Precipitation	$10^6 - 10^{10}$ eV/cm ² /s/sr/eV	15% fluxes, 15% energy	0.1 sec
Energetic Ion Fluxes 20 eV – 20 keV (all pitch angles, 20° res.) 20 keV – 1MeV (downward flux, 90° res.)	Ion Energization and Precipitation	$10^5 - 10^9$ eV/cm ² /s/sr/eV	15% fluxes, 15% energy	0.1 sec
	Ring Current and Radiation Belt Location	$10^3 - 10^7$ eV/cm ² /s/sr/eV	20% fluxes, 25% energy	1 sec
Plasma Density	Ion-Neutral Drag and Plasma Continuity	$10^2 - 10^7$ cm ⁻³	$\pm 10\%$	0.1 sec
Ion & Electron Temperature	Energy Balance	300 -10000 K	$\pm 10\%$	1 sec
Neutral Density	Ion-Neutral Drag and Neutral Gas Continuity	$10^9 - 10^{12}$ cm ⁻³	$\pm 10\%$	1 sec
Neutral Temperature	Energy Balance	300 – 3000 K	$\pm 10\%$	1 sec
Ion Composition (H⁺, O⁺, NO⁺, O₂⁺, Mg⁺, Fe⁺)	Production, Loss and Chemistry	relative concentration	$\pm 1\%$	1 sec
Neutral Composition (O, N₂, O₂)	Dynamics and Chemistry	relative concentration	$\pm 1\%$	1 sec

Table 1. Measurement Requirements

Mission Implementation

To accomplish the goals of the SABLE mission in-situ measurements of the neutral and plasma properties within the contact region are required. As such, a “deep-dipping” element of the mission is imperative to reach the regions where fundamental interactions connect the atmosphere and space. As the response and scale dependencies of the neutral gas and the plasma are vastly different, a multi-spacecraft mission is needed to quantify the persistence of the energy inputs and the response of the contact region. In this report, we outline a cost-effective reconfigurable constellation of small satellites that will resolve these key issues.

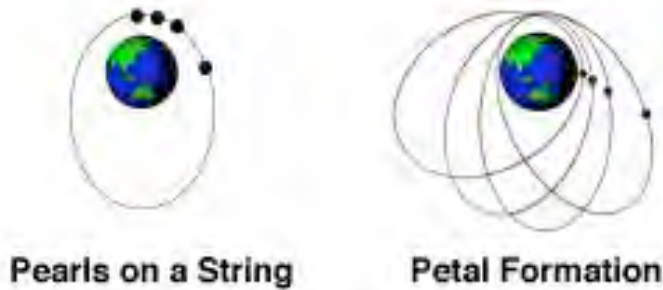


Figure 2. Orbital configurations for SABLE that allow space-time ambiguities to be resolved and altitude profiles to be identified.

Orbit Requirements

Two arrangements of the satellite orbits are required to conduct the SABLE mission: “deep dipping” in a string of pearls configuration and a “petal orbit”. They are illustrated in figure 2. In brief the pearls on a string configuration allows SABLE to sample a particular region in the sequence of one spacecraft followed by another. Separations between 1 and 30 minutes will allow temporal variations of these periods to be specified for features that have corresponding longitude coherence. This configuration is a powerful and necessary tool to discover the impact of features with different levels of persistence. Vertical coupling is most efficiently investigated by employing a petal configuration that provide for the satellites to simultaneously sample the same latitudinal region at different altitudes. Studies of the locations of apogee and perigee, and the science return that they provide, will be required to maximize the productivity of the SABLE mission.

Access to altitudes below 200 km.

The possible approaches to dipping campaigns are many because of the flexibility inherent in the locations, durations, and amounts of perigee excursion. The trade space is large in terms of science return and resource expenditure, so guidelines for developing a dipping strategy will be necessary. The high-priority phenomena manifest themselves most profoundly below 150 km; so spending a significant amount of time below 150 km is a high priority. The phenomena are not uniformly distributed in local time, so measurements over a range of local times are necessary. Additionally, the requirement to make measurements through the high latitude auroral zone and many local times makes an orbital inclination of 75 degrees optimal.

INTRODUCTION:

We propose a mission to study coupling of the thermosphere and ionosphere with emphasis on times of magnetic activity. During these periods, energy in the form of Poynting flux enters the high-latitude ionosphere. This energy is converted and partitioned into Joule heat (thermal energy) and neutral winds (kinetic energy). The winds transport Joule heat to lower latitudes. The effects are manifest in the thermosphere as an increase in neutral density at any specific altitude. This has been demonstrated in many studies since the CHAMP and GRACE spacecraft were launched.

While the scenario described above has been indirectly verified, the physics has received little attention due to a lack of observations in this region. In particular, neutral winds are almost unknown in the global ionosphere-thermosphere (IT) system, the only useful data *in situ* coming from DE-2, a short-lived mission launched in 1981.

The specific goal of our proposed mission is an improvement in forecasting of satellite drag. We describe the details below. However there are multiple science questions which are addressed, all relevant to an understanding of ion-neutral coupling in the IT system.

Question #1: *What are the primary observations necessary to understand energy deposition, partitioning and transport in the IT system?*

Question #2: *How accurately can the neutral winds be measured remotely, and what is their global distribution with respect to solar flux and geomagnetic activity?*

Question #3: *What is the altitude profile of neutral wind?*

Question # 4: *What is the thermospheric response to energy input from geomagnetic and solar disturbances in terms of density, composition, and winds perturbations? What are the time scales of the responses and how do they manifest themselves locally and globally?*

Question #5: *What is the amount, spatial and temporal distribution of energy deposited in the IT system during geomagnetic disturbances?*

Question #6: *Exactly how and where is the input energy distributed in the thermosphere?*

Question #7: *Quantitatively how do variations of the satellite drag coefficient affect the measurement of neutral density and how much error can they introduce to this measurement?*

THERMOSPHERE-IONOSPHERE COUPLING AND ORBIT PREDICTION SYSTEM:

A schematic of the physics addressed in the mission is given below. Input to the IT system is from the sun above and the troposphere below, and output in the form of drag is shown at right. The boxed area at center left indicates the ion-neutral coupling which takes place in the IT system, in particular energy input at polar and auroral latitudes, partitioning into thermal and kinetic energy, and mass transport.

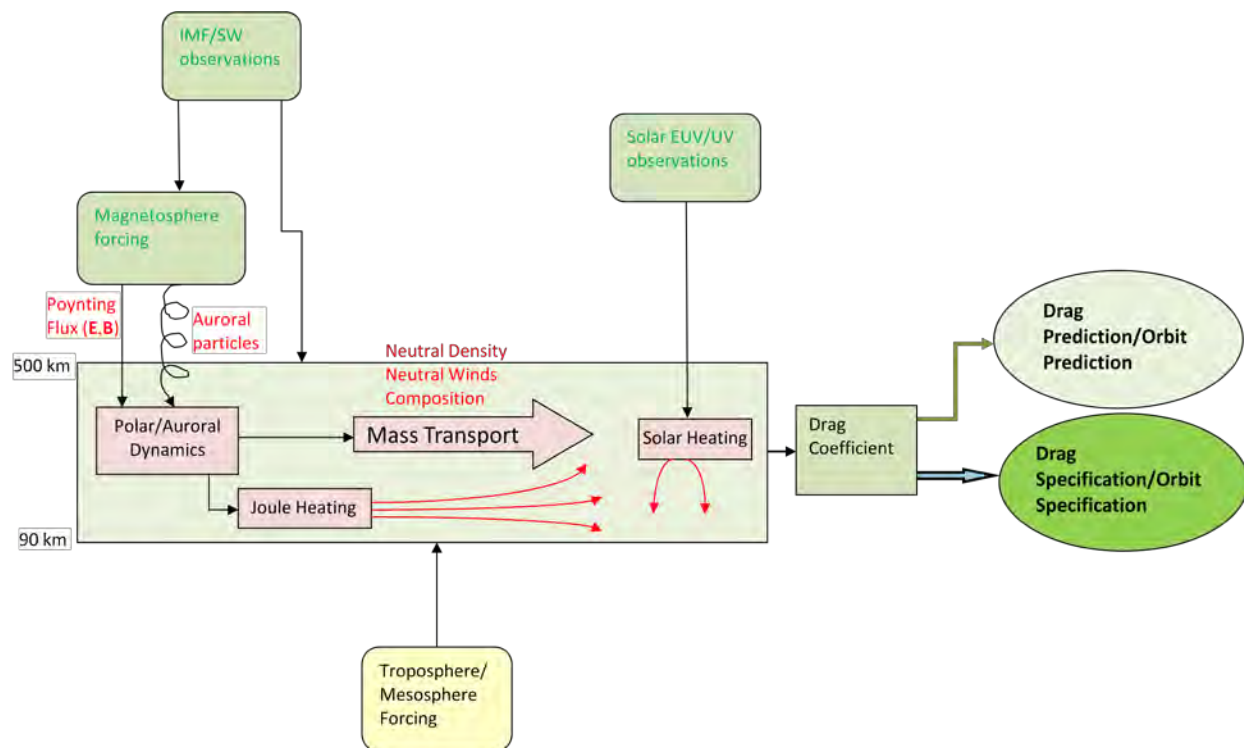


Figure 1: Schematic of the IT system

The mission, titled “Thermosphere-Ionosphere Coupling and Orbit Prediction System” will measure all the quantities in the satellite drag equation, thus enabling improved specification and nowcast of drag. It will measure the required observables in the IT system –neutral density, winds and composition, ion temperature and velocity, particle precipitation and magnetic field. These will enable the fundamental physics of coupling to be addressed and also provide data to assimilation models of the IT system. Results from this mission will improve forecasting of satellite drag and precise orbit determination.

The mission leverages off a Multidisciplinary University Research Initiative awarded to the University of Colorado to study the processes leading to improved input for physical modeling of the thermosphere relevant to satellite drag.

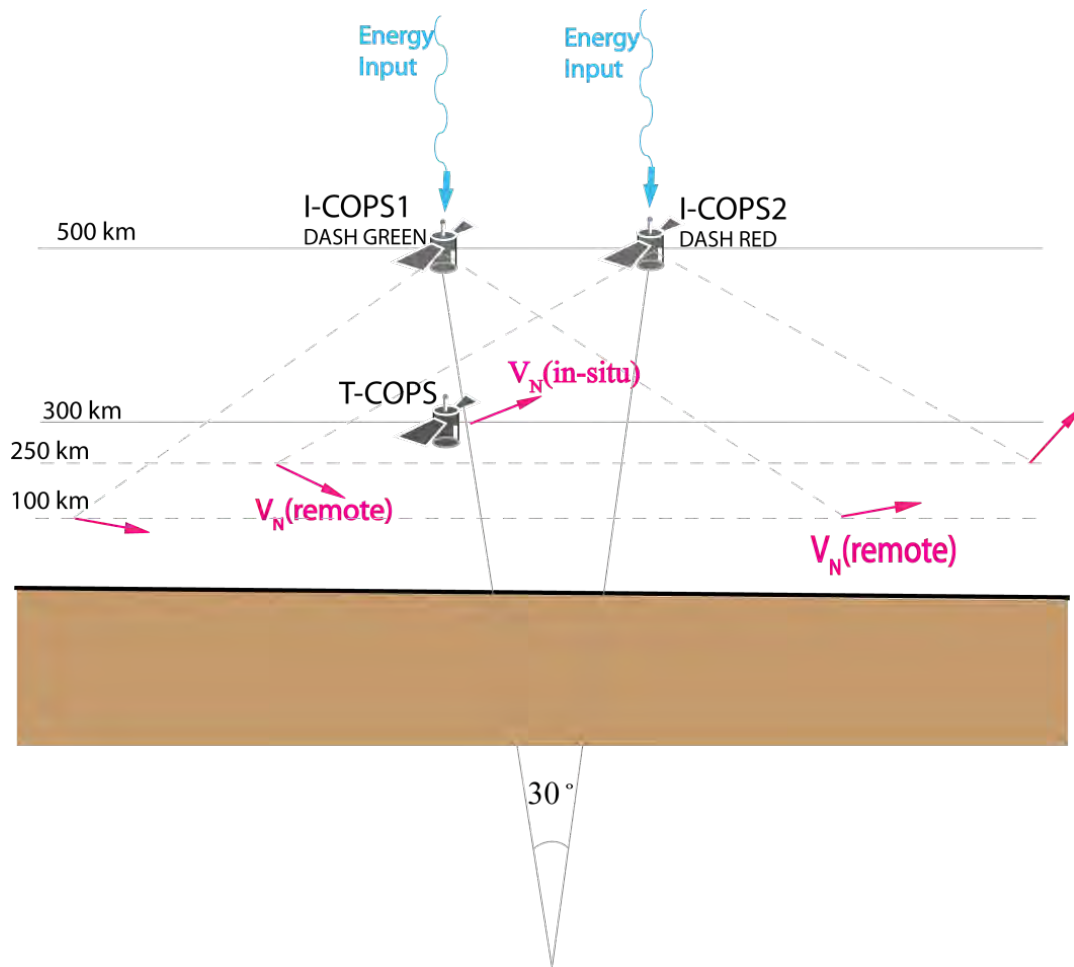


Figure 2: Schematic of proposed mission

TICOPS is a constellation of satellites to measure directly ALL the quantities in the satellite drag equation, enabling (1) immediate radical improvement in nowcasts and forecasts of atmospheric density, winds and composition; (2) input to physics-based data assimilative models which will form the next generation products used in drag prediction.

The mission comprises 3 satellites, 2 I-COPS at high altitude (500km), and T-COPS at low altitude (300km), each flying in high-inclination orbit. The two I-COPS will each carry a remote neutral wind sensor, the Doppler Asymmetric Spatial Heterodyne spectrometer (DASH). DASH is a remote sensing limb imager. Directed toward the horizon, a vertical profile of neutral winds can be obtained. DASH can also be tuned to specific frequencies. One I-COPS DASH will be tuned to the red line, the second to the green line, enabling detailed wind measurements in both E and F layers. This captures energy deposition and mass transport across the ionosphere. Angular spacing of the two I-COPS by 30° allows for full remote sensing across the globe with overlap in coverage from the two spacecraft.

In addition to DASH, I-COPS will carry instruments to measure ion drifts and densities, magnetometers and precipitating particle detectors. These specify the energy input to the IT system in the form of Poynting flux obtained from the magnetometers and ion drift meters, and particle energy. Particle precipitation also allows Pedersen conductivity below the satellite to be estimated.

T-COPS will carry instruments to measure the ion velocity, neutral density, composition and *in situ* winds. Ion measurements are necessary to analyze the conversion of electromagnetic energy into thermal and kinetic energy. Neutral density and composition measurements specify the thermospheric response to energy deposition and dissipation. *In situ* winds are included as a validation experiment for DASH and to expand the database of neutral winds.

T-COPS will also carry a deployable structure which allows the coefficient of drag to be estimated. By changing the acceleration due to drag, we obtain a direct measure of the drag coefficient. Experiments to determine drag coefficients have not been performed for over 40 years. Current ballistic coefficients (drag coefficient*area/mass) have error bars comparable with the values themselves. This leads to errors in the estimates of neutral mass density and hence, satellite drag. By measuring the coefficient of drag, the neutral density, winds and composition we will have direct quantitative measurements of the drag acceleration. This will be the first time these quantities have been all been measured simultaneously and will be valuable in future research into drag prediction.

SUMMARY:

- TICOPS is a constellation of satellites to study ion-neutral coupling in the IT system. It will measure directly ALL the quantities in the satellite drag equation, enabling (1) immediate radical improvement in nowcasts and forecasts of atmospheric density, winds and composition; (2) input to physics-based data

assimilative models which will form the next generation products used in drag prediction.

- Mission comprises 3 polar satellites, two at 500 km altitude and one at ~300 km altitudes. Conjunctions of the 3 satellites provides simultaneous measurements of the full electromagnetic and kinetic energy input to the thermosphere as well as the full thermospheric response

A Comprehensive, First-principles Model of Equatorial Ionospheric Irregularities and Turbulence

J.D. Huba, G. Joyce, and J. Krall

Introduction

Post-sunset ionospheric irregularities in the equatorial F region were first observed by Booker and Wells (1938) using ionosondes. This phenomenon has become known as equatorial spread F (ESF). During ESF the equatorial ionosphere becomes unstable because of a Rayleigh-Taylor-like instability: large scale (10s km) electron density ‘bubbles’ can develop and rise to high altitudes (1000 km or greater at times) [Haerendel, 1974; Ossakow, 1981; Hysell, 2000]. Attendant with these large-scale bubbles is a spectrum of electron density irregularities that can extend to wavelengths as short as 10 cm [Huba *et al.*, 1978]. 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. In fact, for this reason, it is the focus of the Air Force Communications/Navigation Outage Forecast Satellite (C/NOFS) mission [de La Beaujardiere, 2004].

The impact ESF can have on operational systems is shown in Fig. 1 [Makela, private communication, 2005; Kelley *et al.*, 2002]. The top panel is a composite of optical images at $7774 \text{ \AA}$ from Mount Haleakala; the darker areas are regions of low electron density (i.e., ESF plasma bubbles). Note that these bubbles can have complex structures (i.e., tilts, bifurcations). Overlaid on this figure are the orbits of three GPS satellites; the actual satellite locations at this time are denoted by the symbols. Of note is the ‘blue’ satellite. The path of the GPS signal from this satellite to the ground receiver passes through the plasma bubble. The signal is strongly scintillated as shown in the lower left panel where $S_4 \sim 1$ (the S_4 index is a normalized measure of the standard deviation of the GPS satellites’ received signal power as seen on the ground). In this case it is very

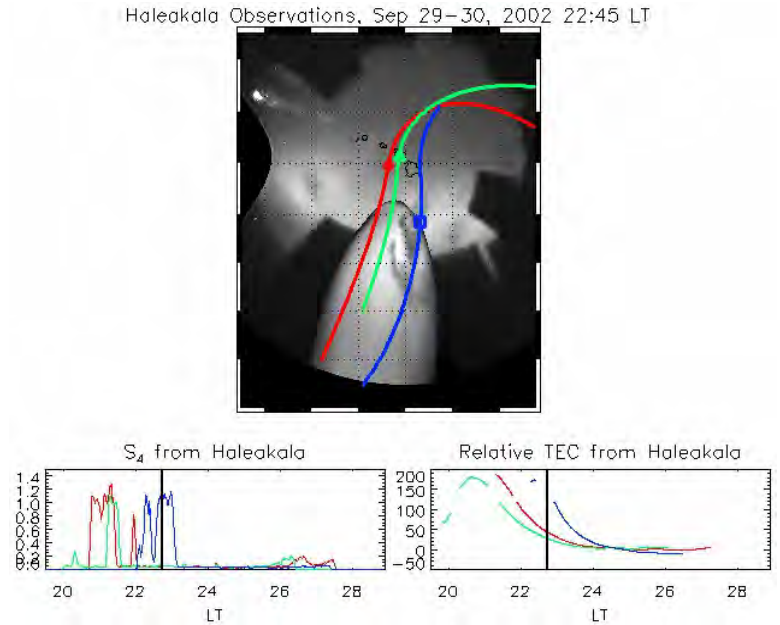


Figure 1: ESF plasma bubbles (dark regions) and GPS signal loss (bottom right) [courtesy of J. Makela].

significant because it causes a complete loss of signal as shown by the break in the blue line at 22:45 LT in the relative TEC measurement in the bottom right panel.

Theoretical [Sultan, 1996] and computational [Zalesak *et al.*, 1982; Huba and Joyce, 2007; Huba *et al.*, 2008; Retterer, 2010] models have been developed to describe the large-scale bubble development shown in Fig. 1; an example is shown in Fig. 2. However, these models are decoupled from the global electrodynamics of the earth’s ionosphere. Only recently has a model been developed that can capture both large-scale global ionospheric dynamics (~ 1000 s km) and small-scale bubble dynamics (~ 10 s km) [Huba and Joyce, 2010]. Although this model provides a new capability to understand the day-to-day variability of ESF, it falls far short of modeling the spectrum of irregularities associated with ESF and the small-scale turbulence (~ 100 s m) that is responsible for scintillating radio wave signals. The purpose of this white paper is to propose a comprehensive theoretical and modeling program to address this deficiency.

Scientific and Computational Challenges

The major computational challenge is to seamlessly couple different, spatially overlapping, physics models. These include (1) thermospheric models that describe neutral dynamics associated with tidal motions, planetary waves, and gravity waves [Fritts *et al.*, 2009; Fritts and Lund, 2010], (2) global ionospheric models that describe the neutral wind dynamo and bubble development [Huba and Joyce, 2010], (3) sub-grid electrostatic fluid models that capture electron density turbulence on scales 10s m to 100s m [Hysell, 2000], and finally (4) hybrid and particle-in-cell codes that model small-scale turbulence below the ion Larmor radius (10s cm to 10s m). This is a daunting challenge since it covers a spatial range spanning 8 orders of magnitude and a temporal range spanning 6 orders of magnitude.

As a first step in this challenging project, existing thermospheric, ionospheric, and electrostatic fluid turbulence models can be coupled to improve current modeling capabilities and to identify computational difficulties that need to be overcome in developing a comprehensive, integrated model. However, there is also a need to improve and extend existing models. One example is the electrodynamics of the ionosphere/thermosphere (IT) system. All global models of the IT system

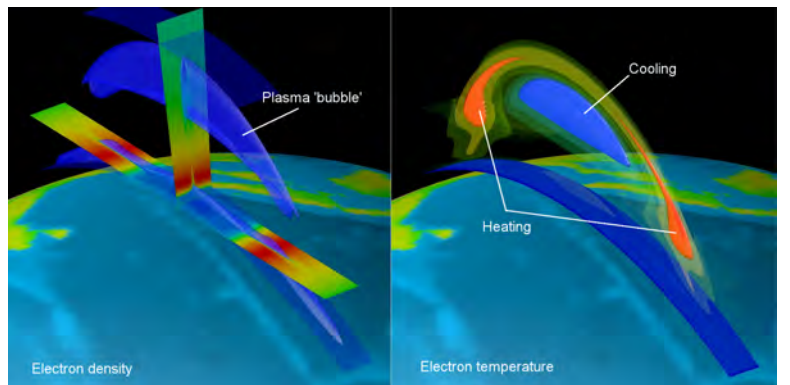


Figure 2: Electron density [Huba *et al.*, 2008] and temperature [Huba *et al.*, 2009] structure of an equatorial plasma bubble.

assume that the geomagnetic field lines are equipotentials; this assumption reduces the potential equation to two dimensions. This assumption is nominally considered valid for global scale sizes $L \gtrsim 1$ km because of $\sigma_{\parallel} \gg \sigma_{\perp}$ where σ is the conductivity of the plasma [Farley, 1963]. However, recent simulation studies by Aveiro and Hysell (2010) suggest this simple scaling argument may not be valid and that 3D electrodynamic model are required. Thus, to accurately describe the self-consistent coupling of large scale to small scale electron density irregularities it is necessary to have a fully three-dimensional electrodynamic model. This will require the solution to a 3D electrostatic potential equation. It will be necessary to have a 3D solver that is robust, efficient, and uses parallel processing on a nonuniform grid. A second example is the need to include finite Larmor radius effects in electrostatic fluid turbulence models [Hysell, 2000]. The oxygen ion Larmor radius is ~ 3 m and the spectrum of ionospheric turbulence is altered on this scale [Huba and Ossakow, 1979].

Proposed Program

A comprehensive modeling program is proposed to attack this difficult and important problem. Resources are needed (1) to develop new numerical algorithms to improve both the physics and efficiency of existing models, (2) to develop potentially new models of IT system and sub-system, and (3) to couple existing and new models seamlessly into a comprehensive framework. This effort requires long term, stable funding involving several research groups. A nominal program would be a 5 year effort at \$2M/year for a total of \$10M. This is consistent with the type of programs suggested by the Advanced Computational Capabilities for Exploration in Heliophysical Science (ACCEHS). In addition to model development, the program also needs to include an data/model comparison component to ensure that simulation results are consistent with experimental observations.

References

- Aveiro, H. and D.L. Hysell, Three-dimensional numerical simulation of equatorial F -region plasma irregularities with bottomside shear flow, to be published in *J. Geophys. Res.*, 2010.
- Booker, H.G. and H.G. Wells, Scattering of radio waves by the F -region of the ionosphere, *Terr. Mag. Atmos. Elec.* **43**, 249, 1938.
- de La Beaujardiere, O., C/NOFS: a mission to forecast scintillations, *JASTP*, **66**, 1573, 2004.
- Farley, D. T., A plasma instability resulting in field-aligned irregularities in the ionosphere, *J. Geophys. Res.* **68**, 6083, 1963.
- Fritts, D. C. et al., Overview and summary of the Spread F Experiment (SpreadFEx), *Ann. Geophys.* **26**, SpreadFEx special issue, 2141, 2009.
- Fritts, D. C. and T. Lund, Gravity wave influences in the thermosphere and ionosphere: Observations and recent modeling, *Aeronomy of the Earths Atmosphere and Ionosphere*, Springer, in press. (see <http://www.co-ra.com/dave/Springer/Springer.pdf>), 2010.

- Haerendel, G., Theory of equatorial spread F , preprint, Max-Planck Inst. für Extraterr. Phys., Munich, Germany, 1974.
- Huba, J.D., P.K. Chaturvedi, S.L. Ossakow and D.M. Towle, High frequency drift waves and wavelengths below the ion gyroradius in equatorial spread F , *Geophys. Res. Lett.* 5, 695, 1978.
- Huba, J.D. and S.L. Ossakow, On the generation of 3 meter irregularities during equatorial spread F by low frequency drift waves, *J. Geophys. Res.* 84, 6697, 1979.
- Huba, J.D. and G. Joyce, Equatorial spread F modeling: Multiple bifurcated structures, large density ‘bite-outs,’ secondary instabilities, and supersonic flows, *Geophys. Res. Lett.* 34, L07105, doi:10.1029/2006GL028519, 2007.
- Huba, J.D., G. Joyce, and J. Krall, Three-dimensional equatorial spread F modeling, *Geophys. Res. Lett.* 35, L10102, doi:10.1029/2008GL033509, 2008.
- Huba, J.D., G. Joyce, J. Krall, and J. Fedder, Ion and electron temperature evolution during equatorial spread F , *Geophys. Res. Lett.* 36, L15102, doi:10.1029/2009GL038872, 2009.
- Huba, J.D. and G. Joyce, Global modeling of equatorial spread F , *Geophys. Res. Lett.* 37, L17104, doi:10.1029/2010GL044281, 2010.
- Hysell, D.L., An overview and synthesis of plasma irregularities in equatorial spread F , *J. Atmos. Sol. Terr. Phys.* 62, 1037, 2000.
- Kelley, M.C., J.J. Makela, B.M. Ledvina, and P.M. Kintner, Observations of equatorial spread- F from Haleakala, Hawaii, *Geophys. Res. Lett.* 29, doi:10.1029/2002GL015509, 2002.
- Retterer, J.M., Forecasting lowlatitude radio scintillation with 3D ionospheric plume models: 1. Plume model, *J. Geophys. Res.* 115, A03306, doi:10.1029/2008JA013839, 2010.
- Sultan, P. J., Linear theory and modeling of the Rayleigh-Taylor instability leading to the occurrence of equatorial spread F , *J. Geophys. Res.*, 101, 2687526891, 1996

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 recommended 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:

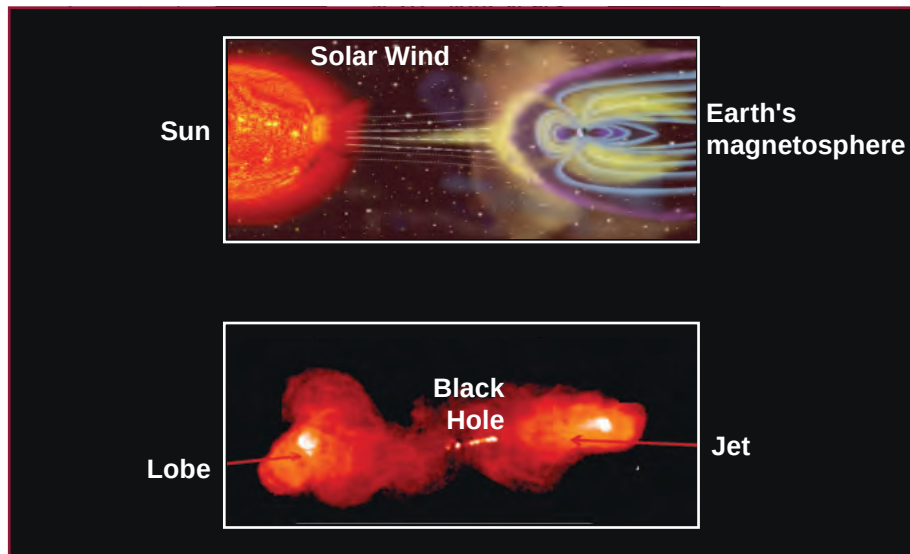
<http://www.pppl.gov/conferences/2010/WOPA>

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 $\sim 10^{-4}$ light years) to the large scale of extra-galactic jets, $\sim 10^6$ 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.

Detecting and Tracking Solar Ejecta with Next-Generation Heliospheric Imaging Systems

Janet C. Johnston

AFRL/RVBXS, Space Weather Center for Excellence, Air Force Research Laboratory, Hanscom AFB, MA 01731-3010

David F. Webb

ISR, Boston College, Chestnut Hill, MA 02467 USA

1. Background and Need

Most significant space weather is initiated at the Sun. Coronal mass ejections (CMEs), huge blobs of plasma with embedded magnetic structure, are the primary cause of severe space weather at Earth because they drive shocks and trigger geomagnetic storms that can damage spacecraft and ground-based systems. Magnetic storms can damage both military and civilian, space and ground assets. These hazardous storms are difficult to forecast and there are many false alarms.

What makes solar and heliospheric disturbances geoeffective, in terms of causing storms, is primarily southward interplanetary magnetic field (IMF) and compression. Southward IMF is important because it allows merging of the IMF and Earth's magnetic field and transfer of solar wind energy and mass into the magnetosphere. Compression is important because it strengthens existing southward IMF and, to a lesser extent, increases density. CMEs usually contain long-duration flows of southward IMF and fast CMEs compress any southward field ahead. Sometimes the geoeffective compressed, southward field can be entirely within the shock "sheath". In addition, CMEs themselves can carry high-density structures, such as solar filaments. High-speed streams are geoeffective when they compress any southward IMF in corotating interactive regions (CIRs), and also because they are associated with the acceleration of electrons in the magnetosphere to hazardous levels.

While coronagraphs observe the launch of a CME close to the Sun, the wide variety of CME speeds and morphologies make it necessary to image and track them through interplanetary space to predict their arrival, duration, strength and effectiveness at Earth. Space imaging is required because of the difficulty to detect the dim CME light against brighter background sources such as the zodiacal light and starlight.

Tracking and specification of significant solar ejecta in the next decade will rely on a data assimilative approach, incorporating imagers, interplanetary scintillation, wide field radar arrays, and other assets not yet in operation. As DoD systems (communications, navigation, surveillance) become more complex, space weather effects will play a more significant role in the fidelity of Earth and space situational awareness. Anomaly assessments are needed both reliably and within a useful timeframe. Forecasting can also be useful both to postpone critical space operations and take mitigating actions to protect sensors or to take advantage of adverse conditions. "All clear" forecasts are also valuable to the operational community. Heliospheric imagery can lead to improvements to solar wind forecast models. The accuracy of these models is crucial should a failure of surveillance assets occur.

2. Prior Remote Sensing Missions

Since the Helios mission it has been known that electrons from CMEs can be viewed in white light from space through Thomson-scattering (Richter et al., 1982). Beginning early 2003, a new era in heliospheric imaging started with the operation of the Air Force Space Test Program demonstration experiment, the Solar Mass Ejection Imager (SMEI – Eyles et al., 2003). SMEI, launched into a circular, Sun-synchronous orbit, captures a full sky image (minus a 20° exclusion zone around the Sun and a small patch at the anti-solar direction) every orbit (102 minutes). A photometric limit of at least 0.01% of the background sky must be attained to detect and track CMEs (Howard et al., 2006; Webb et al., 2006; 2009).

Since late 2006 SMEI's data have been augmented by imagery from the SECCHI suite of staring instruments (Howard et al., 2009) on the twin STEREO spacecraft in ~1 AU orbits around the Sun. Each STEREO has two heliospheric imagers (HI-1 and HI-2; Harrison et al., 2008; Eyles et al., 2009) which together have 70° x 90° views of the heliosphere centered east and west of the ecliptic. STEREO-A is ahead of the Earth and B lags behind, and thus provides an even longer lead time forecast for some phenomena. The STEREO craft diverge from each other at the rate of 22° each per year and will soon be 180° apart. The SMEI and STEREO HI data are complementary. HI-1 fills the gap between a coronagraph and the SMEI inner camera and HI-2 overlaps with SMEI over tens of degrees and has better spatial resolution. SMEI suffers from area obscurations due to particles in Earth orbit, but despite these gaps, SMEI observes nearly the entire sky. Both HI fields have about the same temporal resolution, about 2 hours. The improved spatial resolution of the HIs allows detection of plasma, or density waves that may be associated with stream interaction regions. Although their fields of view are more limited than SMEI's, HI data are free of the contamination suffered by SMEI from particle hits from traversing the South Atlantic Anomaly, the horns of the auroral zones and the noise introduced by high altitude auroras.

Radio techniques have also been used to remotely detect and track and coarsely image heliospheric disturbances related to CMEs. These use km-wavelength radio observations from space and interplanetary scintillation (IPS) observations from the ground. The km-wavelength observations can track the type II emission typically from strong shocks traveling ahead of fast CMEs. Such instruments have been flown on the ISEE-3, Wind, Ulysses and STEREO spacecraft. The IPS technique relies on measurements of the fluctuating intensity level of signals from strong, point-like distant radio sources from one or more ground arrays operating in the MHz range. IPS arrays detect changes in density inhomogeneities in the (local) IP medium moving across the line of sight to the source. Disturbances are detected by either an enhancement of the scintillation level and/or an increase in velocity.

3. Space Weather Requirements and Tradeoffs

Extensive studies have shown that a SMEI-like instrument can detect and track most earthbound CMEs capable of generating moderate to severe geomagnetic storms (e.g., Webb et al., 2009). However, “halo” CMEs, those CMEs which have an earthward component when viewed from a detector along the Sun-Earth line, require aggressive processing for determining the distance to Earth of the incoming material. With such detectors, only the angular distance (elongation

angle) from the Sun is measured and for halo events, there can be an ambiguity between a wide CME near the Sun and a narrow CME near the Earth. And the geometric shape of the leading edge is in question. White light imagers detect light produced by Thomson scattering, so the brightest part of an image is a convolution of the amount of material and its angle from the illumination source.

We are currently testing several methods and models that use HI data for improving forecasts. One method is to assimilate SMEI data into a forecast model. Currently the HAFv2 model is AFWA's operational solar wind model. SMEI data can enable a mid-course correction to the HAF-only forecast up to about a day before CME arrival. AFRL developed and transitioned a simple point-&-click tool/operator interface that permits a forecaster to fit the leading edge of a CME using SMEI data to estimate the arrival time at 1 AU. This tool also yields estimates of the CME speed and its direction angle with respect to the Sun-Earth line.

The Tappin-Howard (2009) model is a 3-D interplanetary CME reconstruction model based on leading edge measurements from heliospheric images and deriving the CME geometry using Thomson-scattering physics. As input, the model currently uses SMEI and/or STEREO HI data. It outputs (1) the basic structure, speed, direction, size, height- and speed-time profile, and (2) an estimation of time of arrival and likelihood of impact with the Earth. The systems impact for the AF/DoD is that it has the potential to provide a fast and accurate prediction of the arrival time of any observed CME at the Earth, or any other point in space.

In the near future we should be able to incorporate not only enhanced plasma density data but also velocity and IMF information in assimilative 3-D interplanetary reconstruction models. Such assimilative processes are increasingly being used operationally by AFWA and other government agencies in geospace weather models such as the Global Assimilation of Ionospheric Measurements (GAIM) model (e.g., Schunk et al., 2004). The velocity data are available now from the ground-based IPS arrays, but these are scattered in time and space coverage. International studies are underway to collect, intercalibrate and make accessible all the IPS data (Oberoi and Benkevitch, 2010).

An important parameter of CMEs is the orientation of the embedded magnetic field. A structure having a southward-pointing magnetic field is more likely to cause significant geomagnetic storming at Earth because it reconnects with the predominantly north-pointing geomagnetic field allowing solar wind plasma and energy to flow into geospace. During the next decade radio facilities will be developed to image and track the magnetic structures within CMEs. The MWA sited in western Australia is currently being constructed and will be capable of obtaining both greatly enhanced IPS data and unique IMF imaging data via the Faraday rotation technique (Jensen et al., 2010).

Relevant requirements for operational space weather support include: time of arrival (ToA) accuracy 1 -3 hours; lead or warning times > 12 hours; a 15 min. cadence, high precision measurements of CMEs to establish initial speed, geoeffectiveness degree; 30 arc-second sampling resolution; the ability to detect and track corotating interaction regions (CIRs); and an optimum CME viewing perspective. This latter involves the orbits into which an HI system could be placed. We discuss three possibilities (Table 1 and Figure 1):

- Earth orbit: Any spacecraft/instrument in Earth orbit suffers from radiation damage and penetrating trapped particles. Also low earth orbit (LEO) and others that pass through the South Atlantic Anomaly (SAA) or auroral zones suffer from particle effects and/or bright auroral emission. In a LEO, near-equatorial orbit, the particle and auroral contamination is reduced, and other Earth orbits such as medium Earth orbit (MEO) are possible.

- Along the Sun-Earth line: For deep space, an L1 halo orbit is probably the easiest and cheapest to get to and to maintain an orbit. However, coronagraphs and HIs cannot directly observe CMEs heading towards Earth, which appear as halos around the Sun, because of the necessity to block the bright sunlight around the Sun.. Thus, that part of the Earth-directed CME that is most geoeffective is not seen. Therefore, CMEs and CIRs need to be observed from a different vantage point that provides a full view of the CMEs, a profile en route, and the CIRs well before they arrive at Earth.

- L5: Coronagraphs and an HI at the fifth Sun-Earth Lagrange point (L5) orbit position (60° trailing the Earth at 1 AU) will be able to detect CME speed and direction for a range of solar origin locations, especially those expected to be most geoeffective. In addition, L5 has the advantage that it views beyond the eastern limb of the Sun as viewed from Earth. Thus, active regions and coronal holes can be viewed before they arrive on the Earth-facing disk, and their location, size and activity assessed. Finally, the geoeffective space weather resulting from high speed solar wind streams, such as radiation belt relativistic electron enhancements, can be forecast days in advance. Observations at or near this location using STEREO-B as a pathfinder for such a mission are discussed by Webb et al. (2010).

4. Summary

Interplanetary space is a dynamic environment, driven to first order by the Sun. The DoD has invested in the capability to specify conditions and phenomena in the heliosphere. An L5 imager, and/or an L1 or LEO Sun-Earth line imager can be accomplished with low risk, banking on knowledge gained from SMEI and STEREO. The ideal system would incorporate both and lead to a system producing stereo viewing of a CME, unambiguous profiling of its leading edge, advanced viewing of solar activity, and an inexpensive Sun/Earth line instrument giving full sky images. Interplanetary scintillation and wide-field radio arrays promise to add details of embedded magnetic fields to complete the CME specification for use in research and space weather forecasting.

References

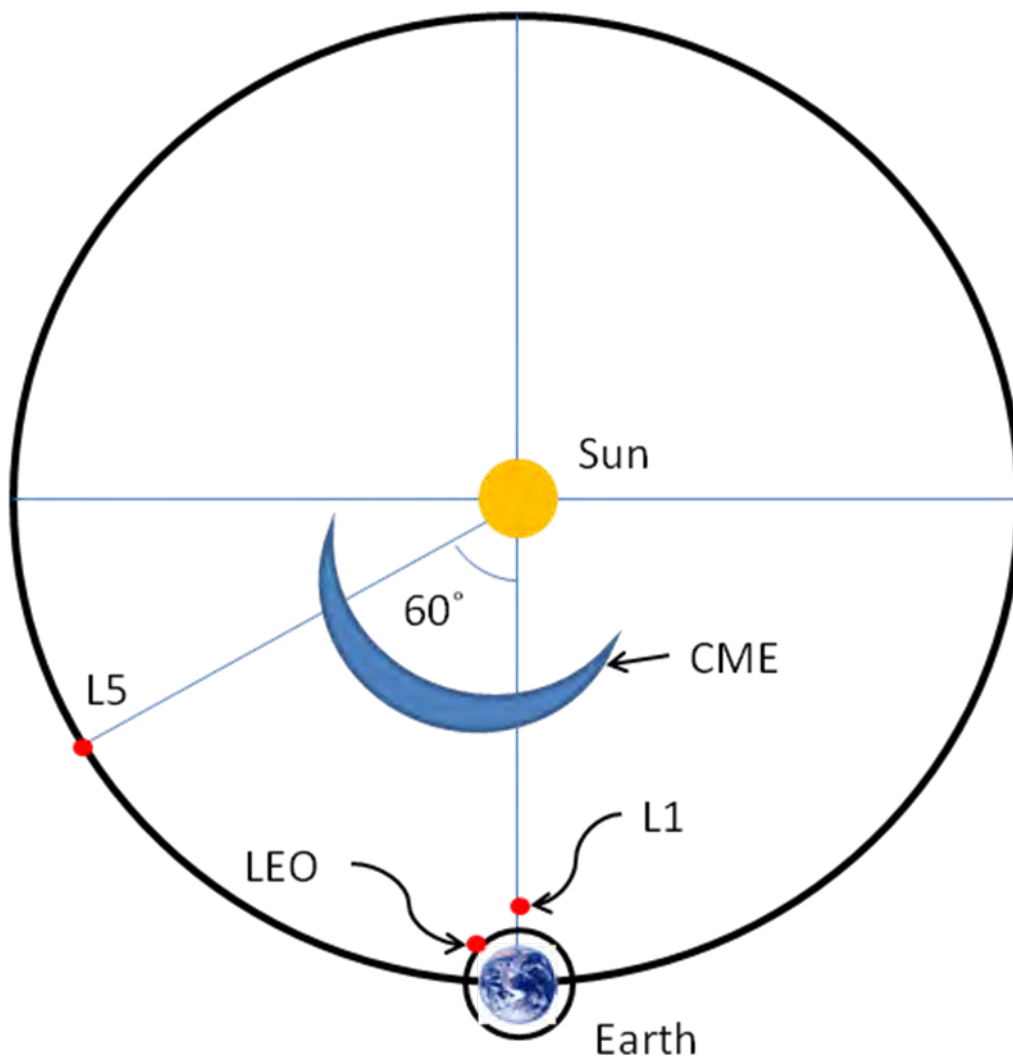
- Eyles, C.J., G.M. Simnett, M.P. Cooke, B.V. Jackson, A. Buffington, P.P. Hick, N.R. Waltham, J.M. King, P.A. Anderson, and P.E. Holladay, *Solar Phys.*, 217, p. 319-347 (2003)
Eyles, C.J. et al., *Solar Phys.*, 254: 387– 445 (2009)
Harrison, R.A. et al., *Solar Phys.*, 247, 171 (2008)
Howard, R.A. et al., *Space Sci. Rev.*, 136, 67 (2008)
Howard, T.A., D.F. Webb, S.J. Tappin, D.R. Mizuno and J.C. Johnston, *J. Geophys. Res.*, 111 (2006)
Jensen, E.A., Hick, P.P., Bisi, M.M., Jackson, B.V., Clover, J., Mulligan, T.: *Solar Phys.*, 265:

- 31–48 (2010).
- Oberoi, D. and L. Benkevitch, *Solar Phys.*, 265: 293–307 (2010)
- Richter, I., C. Leinert, and B. Planck, *Astron. Astrophys.*, 110, 115 (1982)
- Schunk, R. W., et al., *Radio Sci.*, 39, RS1S02 (2004)
- Tappin, S.J. and T.A. Howard, “Interplanetary Coronal Mass Ejections Observed in the Heliosphere: 2. Model and Data Comparison”, *Space Sci. Rev.*, 147, 55 (2009)
- Webb, D.F. et al., *J. Geophys. Res.*, 111 (2006)
- Webb, D.F., T. A. Howard, C. D. Fry, T. A. Kuchar, D. R. Mizuno, J. C. Johnston and B. V. Jackson, *Space Weather*, 7, S05002 (2009)
- Webb, D.F., D. A. Biesecker, N. Gopalswamy, O. C. St. Cyr, J. M. Davila, C. J. Eyles, B. J. Thompson, K. D. C. Simunac and J. C. Johnston, *Space Research Today*, 178, 10 (2010)

Table 1. Options for HI Orbits/Locations

<u>Location</u>	<u>Advantages</u>	<u>Disadvantages</u>
1. Earth orbit Polar LEO	Lower cost; more launch options.	Must be >1000 km. height; Bright auroral emission; Solar & auroral zone particles; SAA particles & radiation dose.
Near-equatorial LEO	Lower cost; more launch options; Few solar particles; Reduced auroral emission & radiation damage.	Must be >1000 km. height; Still some auroral emission and solar & auroral zone particles.
MEO, etc.	Minimize auroral emission.	Still some auroral emission and solar & auroral zone particles.
2. Sun-Earth Line L1 orbit	Lowest cost & ease of maintaining orbit; avoids Earth orbit problems; Detectors (CCDs) easily cooled	More expensive than LEO; HIs cannot observe nose/core of CME; Hard- & software: rad hard & deep space conditions
Other		L2 has Earth in Sun view; no other desirable orbits.
3. L5 orbit	With a coronagraph can view geo- eff. CMEs from side & from Sun to Earth; Early views of solar features east of Earth view; Early detection of geoeff. high speed streams.	Higher cost with weight & power restrictions; Hard- & software: rad hard & deep space conditions.
L5-L1 Constellation	Combines two views to reconstruct 3D kinematics & geometry of CMEs & CIRs.	Higher cost to launch & support two missions; Hard- & software: rad hard & deep space conditions.

Figure1: Proposed orbits for a heliospheric imager with a CME (not to scale).



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 per 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 New 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.

The technological case for TinySats (CubeSats and Nanosatellites) in support of Heliophysics research and the National Space Weather Program

David M. Klumpar, Montana State University, Space Science and Engineering Laboratory,
Bozeman, MT 59717-3840 klump@physics.montana.edu

Brian A. Larsen, Los Alamos National Laboratory, Los Alamos, NM, 87545, balarsen@lanl.gov

Robyn M. Millan, Dept. of Physics and Astronomy, Dartmouth College,
Robyn.Millan@dartmouth.edu

Executive Summary

The goal of this white paper for The Decadal Strategy for Solar and Space Physics (Heliophysics) is to establish the technical readiness of satellite subsystems needed to develop very small satellites for Heliophysics research. The document argues that rapid developments in electronics miniaturization, new manufacturing techniques, and new materials are continuing to reduce size and power requirements, and increase packaging density permitting increased capability for very small satellite platforms for Space Physics research. It is further demonstrated that an industry has developed capable of providing complete satellite subsystems for very small satellites including power systems, communications, attitude determination and control, and propulsion. The time is ripe to include very small satellites as a tool for Heliophysics research.

Very small satellites are being advocated as potential tools to further Heliophysics Research and to support the National Space Weather Program. The class of satellites under consideration ranges from very small, basic “1U” CubeSats with volumes as little as 1000 cm^3 , and masses as small as 1 kg, to CubeSat-like derivatives with volumes exceeding 6000 cm^3 and massing up to 14 kg (“6U” configurations), and upwards to very small microsatellites in the 10 – 50 kg class. “TinySats” such as these are being utilized by the National Science Foundation’s “CubeSat-based Science Missions for Space Weather and Atmospheric Research” program to conduct fundamental Upper Atmospheric, Ionosphere and Magnetosphere research and support goals of the National Space Weather Program. The Air Force is entertaining the use of CubeSats for operational Space Weather missions through its SENSE program. The National Reconnaissance Office has proffered industry contracts to purchase as many as 50 CubeSats under its Colony One and Colony Two programs. More than 40 CubeSats, developed world-wide by a vast variety of institutions (mostly universities) for various purposes, have been launched into space over the past decade. Many of these Cubesats continue to operate after several years on orbit. The most recent launch of CubeSats occurred on Sunday July 11th at 0352 GMT (8:52 am PST), when the Polar Satellite Launch Vehicle (PSLV) lifted off from the Satish Dhawan Space Center on India’s east coast carrying two CubeSats. Additional CubeSats are scheduled for launch at an increasingly rapid rate in the months and years ahead; indeed the launch of the STP-S26 mission from Kodiak Launch Complex, Alaska carrying several CubeSats, a 30-kg-class nanosatellite pair, and other satellites (7 in total) is scheduled for launch one week from today (November 19, 2010.) Several of these satellites are carrying Heliophysics science payloads.

But what are the technical capabilities of these very small orbiting platforms? Are these miniature satellites merely hobbyist toys, or do they have the potential to make a meaningful contribution if employed for Heliophysics and Space Weather research? What unique capabilities might these potential new space mission platforms bring to the table that are not currently available to more traditional space platforms? Do technical subsystems exist to support science missions and are these spacecraft support subsystems sufficiently mature to support science missions? What is the TRL of existing solutions, and are they readily available to developers of Heliophysics and Space Weather (HaSWx) missions. This white paper examines the current capabilities of TinySat subsystems available for developers of HaSWx missions. Other white papers (e.g. Swenson, et al.) being submitted in response to this call describe the scientific case for CubeSats. In this document, the case is advocated that sufficient technological solutions are readily available to support many HaSWx missions, and that ongoing and future technical advancements are being made at a such a rapid pace to enable even more sophisticated HaSWx missions during the coming decade. It is the right time for the HaSWx community to endorse TinySats, including CubeSats as new unique capability in our Heliophysics toolbox.

To establish the technological capabilities of TinySats, we briefly describe the capabilities of readily available technical subsystems including: attitude determination and control systems that provide sub-degree pointing control and attitude determination, miniature high efficiency power conversion units that optimize solar array power collection and that provide highly regulated output voltages, communications systems capable of Mb/sec telemetry downlinks that dynamically adopt bit rates to the available link margin, propulsion systems, command and data handling subsystems with sufficient processor speed and memory storage capability, and new fabrication techniques capable of producing an entire spacecraft structure in mere hours. Many of these subsystems are available for purchase as commercial off-the-shelf subsystems, leaving the HaSWx developer free to focus on his/her scientific mission and his/her HaSWx payload. In this document we purposefully will not discuss scientific payloads, or the science enabled by specific payloads. The purpose of the document is to enlighten the reader on current capability, availability of, and the technology readiness of spacecraft subsystems needed to support TinySat HaSWx missions today. Many of the examples are specific solutions known to exist by the authors. This document does not purport to represent anything resembling a thorough expose of the available subsystems.

Access to space for small satellites has been the Achilles Heel of previous attempts to provide the research community with small low-cost orbital missions (e.g. the UNEX, STEDI mission concepts that failed to take hold in the mid-nineties and early years of this century.) Even NASA's highly successful Small Explorer program has experienced diminishing launch frequency owing to cost growth. SMEX investigations, now consuming 3-4 years in development and most recently capped at \$105M, have clearly eclipsed categorization as small, low-cost, quick turn-around missions for focused scientific research. SMEX-class investigations no longer represent small science, are no longer within reach of the vast majority of research institutions, and no longer represent opportunities for university students to learn to develop space flight systems through direct hands-on involvement. Only a small hand-full of large space institutions have the institutional resources to conduct SMEX-class missions. HaSWx scientific advancements are being hindered by the high barrier to entry that excludes all but the largest institutions. HaSWx needs frequent, and truly low-cost access to space as an element of its full

systems approach, and to return vibrancy to the program. This paper represents that CubeSats, formerly described as mere hobbyist toys, and their current and developing derivatives are a viable tool, even today, to advance the state of HaSWx research.

The examples given here demonstrate that reasonably mature technological capability exists to warrant the aggressive addition of this new tool based on low-cost small satellite platforms to our experimental program of HaSWx research.

Unique Attributes of the CubeSat System: Standardization and Containerization

The existence of an internationally adopted CubeSat standard, combined with a simple, rugged launch dispenser that is readily accommodated on a wide range of launch vehicles is the key to low-cost, frequent access to space for satellites that conform to the CubeSat Design Specification.

The CubeSat Standard

A CubeSat Design Specification (CDS) is in place to allow the greatest flexibility to developers of CubeSats, while simultaneously mandating only the essential constraints to permit launch by the PPOD. The standard is a simple specification that controls only the essential design parameters required for successful integration into the launch deployer and protects other stakeholder's hardware during launch and injection into orbit. The CDS is maintained by California Polytechnic University in San Luis Obispo. [available at <http://cubesat.atl.calpoly.edu/index.php/documents/developers>]. It specifies the exact external dimensions of the satellite, the mass, required deployment switches and springs, and areas that may be used for access ports for ground testing and battery charging after the satellite has been placed in the PPOD.

Containerized Dispenser Standard

The PPOD provides standardized launch accommodation for up to three 1U CubeSats (or combinations involving a single "3U", a pair of "1.5U"s, or a "2U" and a "1U"). The simple design of the PPOD has allowed it to be qualified for launch on a very wide range of both small and large launch vehicles in the US and abroad. The PPOD interface control documents and other PPOD technical information is available at: <http://cubesat.atl.calpoly.edu/index.php/documents/launch-providers>.

Maturity and availability of Subsystems

Attitude Determination and Control Subsystems

Attitude or pointing control and attitude determination and adjustment solutions for CubeSats providing pointing accuracies to better than one degree and accommodating both spinners and three-axis stabilization are commercially available. Two such sources include Maryland Aerospace, Inc.[<http://www.miniadacs.com/>] and Sinclair Interplanetary [<http://www.sinclairinterplanetary.com/>].

Maryland Aerospace specializes in development of attitude determination and control systems for small and miniature spacecraft. They manufacture the IMI-100/200 product line which are

the first self contained ADACS systems specifically intended for CubeSats and other microsats. Their mini ADACS products include Sun sensors and Earth sensors compatible with TinySats.

Products available from Sinclair Interplanetary include Digital Sun Sensors, Reaction Wheels, Magnetic Torque Rods and Star Trackers designed and manufactured specifically for small satellites

Structural Subsystems

Commercial vendors have developed multiple-line subsystems for CubeSats. In the United States Pumpkin, Inc [<http://www.cubesatkit.com/>] supplies a wide range of structural component kits that permit CubeSat developers to design and build sensors that fit into CDS-conforming structures in sizes ranging from 1U to 3U configurations. Pumpkin's low-cost kit approach and their unique structural system that eliminates much of the machining involved in more typical approaches permits rapid CubeSat development at quite low cost. Pumpkin is mentioned elsewhere as a source of a variety of Command and Data Handling Systems, and as a vendor for various other subsystems described in this white paper.

In Europe Innovative Solutions In Space (ISIS) self-described [<http://www.isispace.nl/>] as "a young and dynamic company in the space industry. The company specializes in the miniaturization of satellite systems with a particular emphasis on the design and development of subsystems for micro- and nanosatellites. Located in the Netherlands in Delft, ISIS supports small satellite projects and missions with its services and products." ISIS, like Pumpkin, specializes in multiple-lines of subsystems for CubeSats. Their structural approach is based on a modular frame, allowing integration of a stack of PCBs into the secondary structure with the subsequent addition of side frames and shear panels for stiffness and solar panel mounting. Their modular structural system readily allows 2U and 3U units to be fabricated.

Innovative new spacecraft structural approaches are being developed compliant to the CubeSat standard by a consortium of developers for the RAMPART CUBESAT (RAPidprototypedMEMSPropulsionAndRadiationTest CUBEflowSATellite) mission. While not in commercial production, this technology utilizes rapid prototyping methods of building one-piece satellite structures, propellant tanks, printed circuit board cages, erectable solar panels, antenna deployment mechanisms, etc. at a fraction of the cost of current methods. New laser-sintered materials with improved mechanical and thermal properties are being adapted for space use from the Formula 1 Racing field. Polymer sealants and metal platings have been utilized on surfaces inside and outside the satellite to eliminate outgassing and to aid in thermal management. The mission and the technology is described in further detail in a manuscript from the proceeding of the 24th Annual AIAA/USU Conference on Small Satellites [ref. paper number SS10-III-8, "3D Printing and MEMS Propulsion for the RAMPART 2U CUBESAT", Moore, et al., August 9-12, 2010]. RAMPART is nearing its flight build in preparation for launch by the Space Test Program in 2011. Innovative concepts such as the "print-and-fly" technique being used on RAMPART point to an entirely new paradigm in rapid low-cost satellite manufacturing.

Communications Systems

CubeSat communications systems are being produced by a number of companies for telemetry and command uplink on frequency bands including UHF, VHF, and S-band. At UHF and VHF telemetry rates are available up to 19.2 kbaud.

Higher telemetry rates including dynamic bit rate systems that self adjust to up to 10 Mbps are available. Adaptive Radio Technologies, LLC (A.R.T.) offers elegantly designed, ground-breaking communications systems for space, weight, and power-constrained applications. A.R.T.'s comprehensive systems dramatically increase the value of missions by increasing communications bandwidth a factor of 10x or more without a corresponding increase in power demand. A.R.T. is based in Los Alamos, New Mexico
[<http://adaptiveradiotech.com/aboutus.shtml>].

AstroDev is a U.S. producer of CubeSat telecommunications transceivers. ...
ISIS, mentioned previously provides a full duplex CubeSat transceiver system operating in the UHF and VHF bands, and has an S-band system under development.

While the system specifications seem meager for TinySats it is important to remember to scale the argument. One example is the downlink systems; a modest system can achieve a 9600baud downlink which combined with 2 ground contacts per day provides a spacecraft with about 160 bits/s orbit average data rate. Already readily available are systems that provide 19.2kbaud providing 320 bits/s there are only antenna size (physics) and power (engineering and physics) limitations on pushing this limit higher. Standard particle measurements use between 3 and 10 kbits/s (e.g. RBSP-ECT), which is accommodated in large satellites by more ground contact time and faster downlink speeds. A TinySat carrying one such instrument would require downlink of 192 kbaud for the same 2 ground contacts, the Colony Two program's desired performance specifications call for data rates from the high 10's to low 100's of kbaud. Assuming that the requirements are met by the developer this will be demonstrated soon.

Propulsion

The RAMPART mission, mentioned above, will raise its own apogee altitude from 450 km to 1200 km, just below the equatorial inner Van Allen Radiation Belt, utilizing a self-pressurizing warm gas propulsion subsystem and microelectromechanicalsystem (MEMS) technologies in its resistojet propulsion system. The propellant tanks are 3D-printed as an integral element of the mechanical structure manufactured of WINDFORM materials, which is currently used extensively in the manufacture of Formula One racing cars and wind tunnel models.

Solar Cell Technologies and electrical power systems

TinySats are readily able to take advantage of the same advances in solar cell technology as any other mission. Large solar arrays are commonly multiple strings of small cells wired in parallel and series. These cells are fully appropriate for TinySats with no modification. New technologies, including flexible, deployable, roll-up cells are being developed.

Several companies provide full electrical power systems for CubeSats and other small satellites. ISIS, already mentioned, has a full line of EPS systems and components, including provision of

fully assembled solar panels. ClydeSpace, [<http://www.clyde-space.com/>] also produces a full line of power systems and components for CubeSats.

Future Developments of the CubeSat Standard

Future prospects of technology

There are many HaSWx missions that do not lend themselves to TinySats based on a physics constraint (e.g. collection area, geometric factor), there however are many classes of measurements that do. The many different sizes and shapes of TinySats that are in development from 1U cubesats to the 50kg small satellite allow for the spacecraft to really be tailored for the mission and the budget constraints. Further constraints will only serve to strengthen the technological drive for TinySat systems. The community is passing beyond the “chicken and egg” problem of not being able to do missions because there is not technology and no technology because there are no missions which will enable systems no yet thought of, not unlike the clunky car phone of the 1990’s to the iPhone of today, what comes next? Only fostering of launches and opportunities for development are required for vast improvements in bus and system technologies.

Going beyond the current capabilities of the 3U launch dispenser, current efforts have led to a 6U class of cubesats. Seeded by initiatives at NASA Ames, and at AFRL-Albuquerque, new canisterized launch dispensers have been built that now allow development and launch of 6U satellite configurations (measuring roughly 10 cm x 20 cm x 30 cm). These spacecraft have the capability to carry significantly more payload mass with far fewer constraints on payload footprint and volume. Spacecraft technical resources include many tens of watts of orbit average power, enhanced pointing control, higher delta-v propulsion capability to enable formation flying. Yet the basic containerized system still provides protection to the launch vehicle and the primary payload when manifested as secondary payloads flying under excess launch capacity.

Detecting Tsunami Generated Ionospheric Perturbations Using GPS Measurements

Attila Komjathy, David Galvan and Anthony Mannucci
Jet Propulsion Laboratory, California Institute of Technology

November 12, 2010

The December 26, 2004 earthquake-generated Sumatra tsunami caused enormous loss of life and property, even in locations relatively far away from the epicentral area (see Figure 1). The losses would likely not have been so massive had an effective worldwide tsunami warning system been in place. A typical tsunami travels ~ 200 m/s taking approximately 7-8 hours to cross the Indian Ocean, for example. A warning system could detect a tsunami and provide an alert to coastal areas in its path. Among the strengths of a tsunami early warning system would be the capability to provide an estimate of the magnitude and location of an earthquake and to confirm the amplitude of any associated tsunami before it reaches populated areas. In the aftermath of the Sumatra tsunami, an important effort is underway to interconnect seismic networks and to provide early alarms quantifying the level of tsunami risk after an earthquake hits.



Figure 1. The Indonesian province of Aceh was devastated by a tsunami on December 26, 2004 (left image, before tsunami – right image, after tsunami) [Mustain, 2010].

However, the seismic estimation process cannot quantify the exact amplitude of a tsunami, and so the second step, that of tsunami confirmation, is still a challenge. The earthquake fault mechanism at the epicenter cannot exactly explain the initiation of a tsunami as it is only approximated by the estimated seismic source. The fault slip is not transmitted linearly at the ocean bottom due to various factors including the effect of the bathymetry, the fault depth, and the local lithospheric properties as well as possible submarine landslides associated with the earthquake.

In the open ocean, detecting, characterizing and imaging tsunami waves is possible only to a limited degree. The offshore vertical tsunami displacement (on the order of a few

centimeters up to half a meter in the case of the Sumatra tsunami) is hidden in the natural ocean wave fluctuations, which can be several meters or more. In addition, the number of offshore instruments capable of tsunami measurements, such as tide gauges and buoys, is very limited. For example, there are only about 70 buoys in the whole world. As a tsunami propagates with a typical speed of ~ 700 kilometers per hour, a 15-minute confirmation system would require a worldwide buoy network with a 150 kilometer spacing. Such a system would be prohibitively expensive, considering such buoys cost $\sim \$250,000$ each to build, and more to maintain.

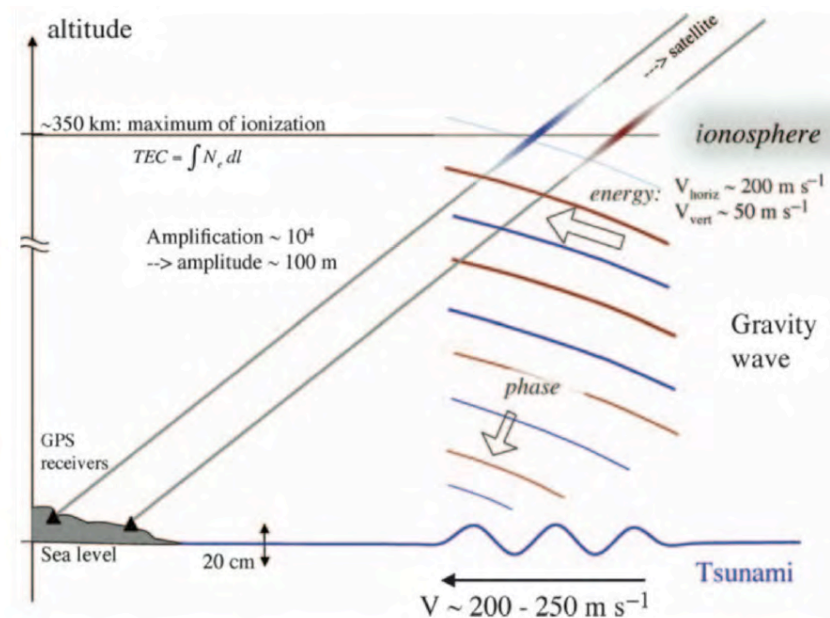


Figure 2. Illustration for tsunami-generated gravity waves in the ionosphere [Artru *et al.*, 2005].

Satellite altimetry has recently proved capable of measuring the sea surface variation in the case of large tsunamis, including the December 2004 Sumatra event [Occhipinti *et al.*, 2006]. However, satellites only supply a few snapshots along the sub-satellite tracks. Optical imaging from the shore has successfully measured the wave arrival at the coastline but is very inefficient in the open sea. At present, only ocean-bottom sensors and GPS buoy receivers supply measures of ocean vertical displacement. In many cases, the tsunami can only be identified several hours after the seismic event, due to the poor distribution of sensors. This delay is necessary for the tsunami to reach the buoys and for the signal to be recorded for a minimum of one wave period (a typical tsunami wave period is between 10 and 40 minutes) to be adequately filtered by removing the noise due to normal wave action.

In the case of the December 2004 Sumatra event, the first tsunami measurements by any instrumentation were only made available about 3 hours after the earthquake. It was supplied by the real-time tide gauge at the Cocos Islands, an Australian territory in the south-east Indian Ocean. Up until that time, the tsunami could not be fully confirmed and

coastal areas remained vulnerable to tsunami damage. This delay in confirmation is a *fundamental weakness* of the existing tsunami warning systems.

Recently, observational and modeling results have confirmed the existence and detectability of a *tsunami-driven signature* in the ionosphere. Physically, the displacement induced by tsunamis at the sea surface is transmitted into the atmosphere where it produces internal *gravity waves* (IGWs) propagating upward. When a fluid or gas parcel is displaced at an interface to a region with a different density, gravity restores the parcel toward equilibrium resulting in an oscillation about the equilibrium state; hence the term gravity wave. The normal ocean surface variability has a typical high frequency of several to tens of seconds (compared to tsunami wave periods of 5 to 60 minutes) and does not transfer detectable energy into the atmosphere. In other words, the Earth's atmosphere behaves as an "analog low-pass filter". Only a tsunami produces gravity waves in the atmosphere that can successfully propagate to ionospheric altitudes. During the upward propagation, these waves are strongly amplified by the double effects of the conservation of kinetic energy and the decrease of atmospheric density resulting in a local displacement of several tens of meters at 300 kilometers altitude in the atmosphere. This displacement can reach a few hundred meters for the largest events.

At an altitude of about 300 kilometers, the neutral atmosphere is strongly coupled with the ionospheric plasma, and the gravity waves produce perturbations in the ionospheric electron density. These perturbations are visible in GPS and satellite altimeter data since those signals have to transit the ionosphere. Figure 2 shows a schematic representation of the propagation of tsunami-driven IGWs up into the ionosphere, and the effect that resulting ionospheric perturbations can have on total electron content (TEC) measurements from GPS radio signals. In Figure 3, it is shown that the dual-frequency signal emitted by GPS satellites can be processed to obtain the integral of electron density along the path between the satellites and the receiver, the total electron content (TEC). The upper plot displays the absolute slant TEC (STEC) values observed in the line-of-sight between the ANIP receiver on the island of Hawaii and the SVN40 GPS satellite, with a 10th order polynomial curve fit to the measurements. The middle panel shows the residual differences between the observations and a 10th order polynomial fit to the observations. The bottom plot shows the variations in the absolute TEC after running the data through a band-pass filter, which removes variations with periods outside the typical range of tsunami wave periods. Note that the variations in TEC are clearly visible in all three plots between ~23:50 on Sep 29 and 00:50 on Sep 30 UT. This ionospheric variation occurs just as the ocean tsunami itself is moving through the same region. [for detailed analysis see e.g., *Galvan et al.*, 2010a; 2010b].

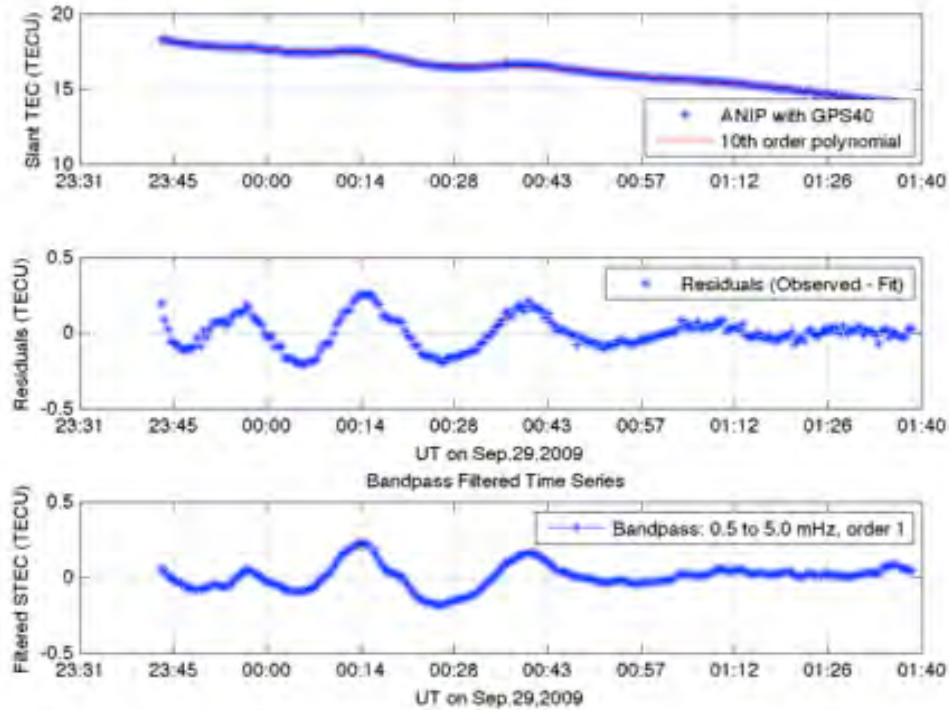


Figure 3. Time series observations of STEC (top), residuals with a fit to a high-order polynomial (middle), and band-pass filtered STEC (bottom) for the Anipeahi (ANIP) GPS receiver on Hawaii in communication with GPS satellite SVN40.

Within 1.5 to 2 hours, the waves generated at the sea surface reach ionospheric altitudes creating measurable fluctuations in the ionospheric plasma and consequently in the TEC. This indirect method of tsunami detection should be helpful in ocean monitoring, allowing us to follow an oceanic wave from its generation to its propagation in the open ocean.

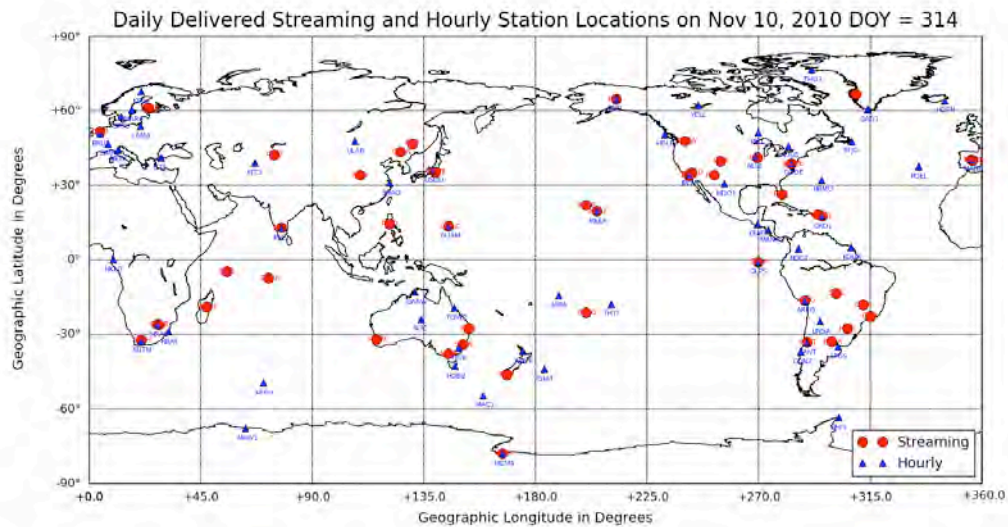


Figure 4. Real-time GPS network that can potentially be used to monitor tsunami-generated TEC perturbations for tsunami-early warning (<http://www.gdgps.net/>).

We argue for the concept of ionospheric sounding that would provide a robust method of tsunami confirmation using NASA's global network of real-time GPS receivers such as shown in Figure 4. It is expected that in the future this technique can be incorporated into a tsunami early warning system and complement the more traditional methods of detection including tide gauges and ocean buoys. Research has shown that ground-based GPS TEC measurements combined with a numerical model of the tsunami-ionosphere coupling could be used to detect tsunamis robustly. Such a detection scheme depends on how the ionospheric signature is related to the amplitude of the sea surface displacement resulting from a tsunami. In the near future, the ionospheric monitoring of TEC perturbations should become an integral part of a tsunami warning system that could potentially augment present systems by significantly increasing area of coverage.

REFERENCES

- Artru, J., V. Ducic, H. Kanamori, P. Lognonné, and M. Murakami (2005), Ionospheric detection of gravity waves induced by tsunamis, *Geophysical Journal International*, **160**, 840–848, 10.1111/j.1365-246X.2005.02552.x.
- Galvan, D., A. Komjathy, M. Hickey, J. Foster and A.J. Mannucci (2010a), “Observing Traveling Ionospheric Disturbances Caused by Tsunamis Using GPS TEC Measurements.” On the CD-ROM of Proceedings of Institute of Navigation (ION) GNSS 2010 Meeting, Portland, OR, Sept 20-24. (Won Best Presentation Award).
- Galvan, D.A. A. Komjathy, M.P. Hickey and A.J. Mannucci (2010b), “Understanding Two Recent Tsunami Events Observed in the Ionosphere Using GPS Total Electron Content.” Submitted to *Journal of Geophysical Research*, doi:2010JA016204
- Mustain, Andrea (2010), “Tsunamis Send Mysterious Waves into Atmosphere” Our Amazing Planet website, <http://www.ouramazingplanet.com/tsunamis-send-mysterious-waves-into-earths-atmosphere-too-0632/> Accessed October 28, 2010.
- Occhipinti, G., P. Lognonné, E. A. Kherani, and H. Hébert (2006), Three-dimensional waveform modeling of ionospheric signature induced by the 2004 Sumatra tsunami, *Geophysical Research Letters*, **33**, L20,104, 10.1029/2006GL026865.

White Paper for the Space Physics Decadal Survey

Title: Physics-Based Modeling of the Plasmasphere

Authors: J. Krall and J. D. Huba

Summary

We argue case that a key priority of U.S. space physics enterprise should be to develop a physics-based numerical model of the coupled magnetosphere-ionosphere system that will describe the plasmasphere and its interactions with the ionosphere, magnetosphere and ring current during both quiet and stormtime conditions.

Motivation/Background

The plasmasphere is a toroidal region of cold dense plasma ($10^2 - 10^4 \text{ cm}^{-3}$) surrounding the earth and extending out to several earth radii [Lemaire and Gringauz, 1998 and references therein]. Results from the IMAGE satellite [Burch, 2000]

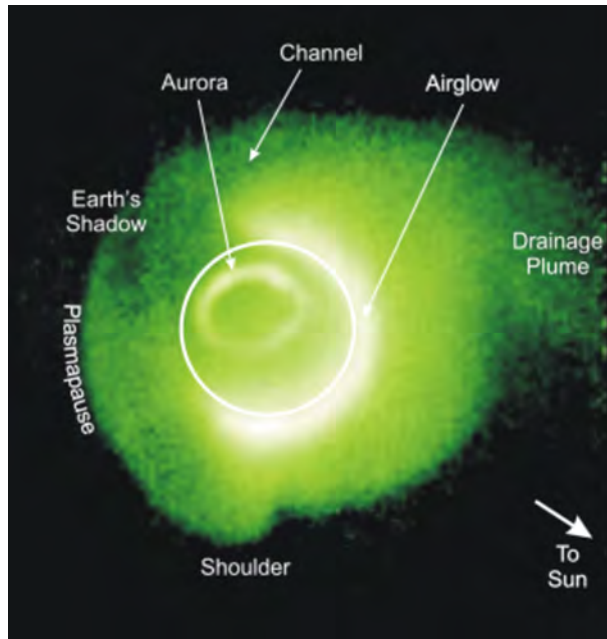


Figure 1: IMAGE EUV image showing the salient parts of the plasmasphere [http://euv.lpl.arizona.edu/euv/]. Note that EUV shows He^+ ions instead of more abundant H^+ .

sparked a renewed interest in plasmasphere dynamics, which continues to this day. EUV (Extreme Ultraviolet Imager) [Sandel et al., 2000] and HENA (High Energy Neutral Atom) [Mitchell et al., 2000] images from the IMAGE satellite showed the global plasmasphere to be a dynamic and complex system (see Fig. 1) that is tied to both the ionosphere (as a source of plasma) and the magnetosphere (whose electric and magnetic fields shape the plasmasphere). During magnetically active periods the outer layers of the plasmasphere undergo erosion. The location and shape of the plasmapause [Nishida, 1966] can change dramatically in several hours [Burch et al., 2001; Sandel et al., 2001; Goldstein et al. 2003a,b]. Subsequently, during magnetically quiet periods, the plasmasphere undergoes refilling by out-flowing ionospheric plasma with timescales of hours to many days [Sojka et al., 1985; Song et al., 1988; Krall et al., 2008].

The plasmasphere plays a significant role in space weather [Goldstein and Sandel, 2007; Goldstein *et al.*, 2007]. First, erosion of the plasmasphere during magnetic storms causes the Van Allen radiation belts to move earthward, potentially exposing spacecraft to high levels of radiation. Second, during the erosion process, large ionospheric density enhancements can be convected from mid-latitude into the auroral region. These enhancements can disrupt communication and navigation systems, and are associated with plasmaspheric plumes [Foster *et al.*, 2007; Yizengaw *et al.*, 2008a]. A global, physics-based model is needed to describe this complex physical system and to interpret and understand the observations.

Key Issues

Plasmaspheric effects on the magnetosphere: It is often assumed that the plasmasphere largely responds to ionospheric and magnetospheric forcings, and that the plasmasphere density or morphology has little effect on the forces that shape it. In fact it is known that the plasmasphere above 1000 km can contribute as much as 60% to global TEC [Yizengaw *et al.*, 2008b; Thompson *et al.*, 2009]. It is reasonable to expect that this electron content impacts the electrodynamics of the inner magnetosphere. Because very little first-principles modeling of this system has been performed, however, this expectation remains untested. Further, the presence of the plasmasphere has influences beyond the macroscopic level. For example, plasma wave and particle processes can depend on plasmasphere density, composition, and density gradients.

Plasmaspheric plumes seen on Earth: Stormtime enhanced density (SED) structures are observed at

Millstone Hill during the early-phase magnetic storms in the pre-midnight sub-auroral zone [Foster and Rich, 1998]. These are high-TEC regions that are convected sunward by poleward electric fields. Sub-auroral polarization streams (SAPS) [Foster and Vo, 2002; Foster and Burke, 2002] form just equatorward of the nightside auroral zone during disturbed conditions and involve rapid westward plasma flows (> 1 km/s). Foster *et al.* [2002] presented observational evidence that strongly suggests that the “storm enhanced density results from the erosion of the outer plasmasphere by sub-auroral polarization stream electric field and the SED/TEC plumes seen in the ionosphere map directly into the drainage plasmaspheric tails” (shown in Fig. 1).

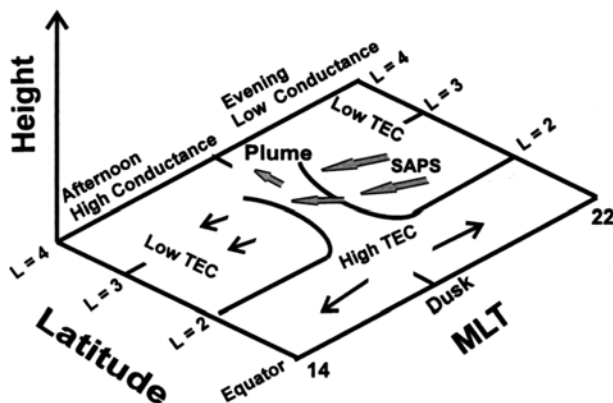


Figure 2: Schematic of the relationship between SAPS, SEDs, and drainage plumes [from Lin *et al.*, 2007]

Recently, Lin et al. [2007] studied a plasmasphere drainage plume using satellite (DMSP, ROCSAT) and IMAGE data for the Bastille Day storm. Their conclusions are similar to those of Foster et al. [2002]. Figure 2 is a schematic of the plasma flow vectors and relative TEC values based on their analysis.

There appears to be an intimate relationship between ionospheric processes (SED/SAPS) and plasmaspheric dynamics (drainage plumes). The degree to which the SAPS phenomenon is associated with, feeds, or even drives SEDs and plasmaspheric structures such as plumes, is not fully understood. Without first-principles, physics-based simulations, this understanding will be difficult to obtain. For example, a plasmasphere-ionosphere model that includes several independently-computed ion species would be needed to address questions of plume composition [Berube et al., 2005; Kim et al., 2007] for various conditions. The limited simulation work that has been done generally supports the theoretical construct of Fig. 2, but much more remains to be done. For example, simulations using the Comprehensive Ring Current Model (CRCM) [Fok et al., 2001] have demonstrated the low trough conductances and field-aligned currents associated with SAPS [Zheng et al., 2008].

Past and Present Efforts to Model the Plasmasphere

Considerable progress toward understanding the plasmasphere has been made through observations and through the construction of empirical models based on the data. For example a plasmaspheric model used in several studies is the Dynamic Global Core Plasma Model (DGCPM) of Ober et al. [1997]. This code solves a continuity equation for the flux tube content of cold ‘thermal’ ions; it uses semi-empirical models for losses and refilling on each field line. DGCPM uses second-order accurate numerical schemes to convect the ions from one grid cell to another in the equatorial plane based on $\mathbf{E} \times \mathbf{B}$ drifts. The code can use various empirical magnetic and electric field specifications. For example, the DGCPM was linked with the Ring current Atmosphere interaction Model (RAM) [Jordanova et al., 1996; Liemohn et al., 2001] and the ionospheric electric potential solver of Ridley and Liemohn [2002] in a study by Liemohn et al. [2004]. The aim was to consider various empirical electric field models to see which best reproduced IMAGE-determined plasmapause positions during the magnetic storm of 17 April 2002. No single electric field model was found to capture all key aspects of the observed plasmapause and plume structure. This result illustrates the limitations of empirical models: they can only reproduce the most common elements of the observations. In order to move beyond the current state of the art, a more comprehensive, physics-based approach is needed.

A number of physics-based models do exist that model elements of this complex system. For example, the Comprehensive Ring Current Model (CRCM) [Fok et al., 2001] self-consistently solves the bounce-averaged Boltzmann convection-diffusion equation of ring current particles and the closure of the currents between the inner magnetosphere and ionosphere. This model has been successful in computing ring current dynamics and in obtaining self-consistency in the electric field between the ring current particles and the

connected ionosphere. Another example is the Lyon/Fedder/Mobarry (LFM) model [Lyon *et al.*, 2004]. The LFM model is a 3D, MHD simulation of the response of the earth's

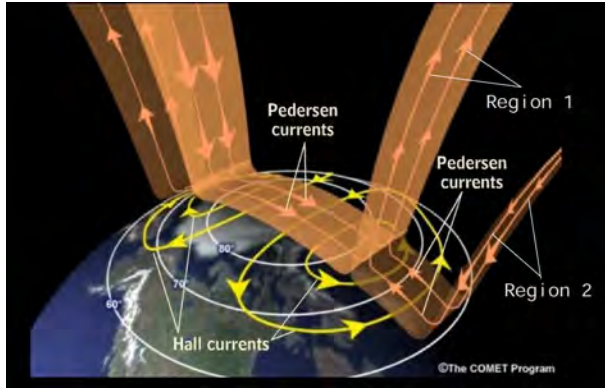


Figure 3: Schematic of the region 1 (outer magnetosphere) and Region 2 (inner magnetosphere) current systems [from Wiltberger, private communication]

magnetosphere to the solar wind and the interplanetary magnetic field (IMF). The LFM code has been used to study magnetospheric issues for model problems [Fedder *et al.*, 1995a], for actual events [Fedder *et al.*, 1995b; Fedder *et al.*, 1997; Slinker *et al.*, 1998, 1999a], and for event simulations with model-data comparisons [Slinker *et al.*, 1999b]. A third example is the SAMI3 ionosphere code [Huba *et al.*, 2000, 2010], which includes seven ion species on an equal footing (i.e., the full continuity, velocity equations are solved for these ion species). SAMI3 self-consistently computes the wind-driven dynamo electric fields, which dominate near Earth and at low latitudes. SAMI3 includes both the vertical and zonal plasma drifts.

These three simulation codes are examples of the state-of-the-art in near-earth space physics simulations. The problem is that they compute electric fields driven by three different current systems: Region I, Region II, and the wind-driven dynamo (see Fig. 3). Each computes the dominant electric potential in the region where that potential is expected to dominate. The difficulty is that these regions overlap—there is no valid prescription for combining these solutions. **What is needed is a global solution of the electric field that encompasses all three of these current systems.** This is a tall order but is doable and is needed to make progress for the future.

A successful physics based model is one that will be able to address these major science questions:

- How do ionosphere, ring current, and magnetosphere conditions affect the density, composition, temperature, and refilling rate of the plasmasphere?
- What is the influence of the plasmasphere on the magnetospheric ring current and on the TEC of the ionosphere?
- What are the necessary conditions for the occurrence of a plasmaspheric plume and a sub-auroral polarization stream, and how are they related to each other and to stormtime enhanced density (SED)?
- What is the effect of the plasmasphere on the global ionosphere-magnetosphere electric fields?

References:

- Berube D., M. B. Moldwin, S. F. Fung, and J. L. Green, A plasmaspheric mass density model and constraints on its heavy ion concentration, *J. Geophys. Res.*, *110*, A04212, doi:10.1029/2004JA010684, 2005.
- Burch, J. L., IMAGE mission overview, *Space Sci. Rev.*, *91*, 1, 2000
- Burch, J. L., Mitchell, D. G., Sandel, B. R., Brandt, P. C., and Wuest, M., Global dynamics of the plasmasphere and ring current during magnetic storms, *Geophys. Res. Lett.*, *28*, 1159, 2001
- Fedder, J.A., J.G. Lyon, S.P. Slinker, and C.M. Mobarry, Topological structure of the magnetotail as a function of interplanetary magnetic field direction, *J. Geophys. Res.*, *100*, 3613, 1995a.
- Fedder, J.A., S.P. Slinker, J.G. Lyon, and R.D. Elphinstone, Global numerical simulation of the growth phase and expansion onset of a substorm observed by VIKING, *J. Geophys. Res.*, *100*, 19083, 1995b.
- Fedder, J.A., S.P. Slinker, J.G. Lyon, C.T. Russell, F.R. Fenrich, and J.G. Luhmann, A first comparison of POLAR magnetic field measurements and magnetohydrodynamic simulation results for field-aligned currents, *Geophys. Res. Lett.*, *24*, 2491, 1997.
- Fok, M.-C., R. A. Wolf, R. W. Spiro, and T. E. Moore, Comprehensive computational model of the Earth's ring current, *J. Geophys. Res.*, *106*, 8417, 2001.
- Foster, J.C. and R.J. Rich, Prompt mid-latitude electric field effects during severe geomagnetic storms, *J. Geophys. Res.*, *103*, 26,367, 1998.
- Foster, J.C. and H.B. Vo, Average characteristics and activity dependence of the subauroral polarization stream, *J. Geophys. Res.*, *107*, 1475, DOI 10.1029/2002JA009409, 2002.
- Foster, J.C., P.J. Erickson, A.J. Coster, J. Goldstein, and F.J. Rich, Ionospheric signatures of plasmaspheric tails, *Geophys. Res. Lett.*, *29*, 10.1029/2002GL015067, 2002.
- Foster, J. C. and Burke, W. J., SAPS: A New Categorization for Sub-auroral Electric Fields, *EOS Trans. AGU*, *83*, 393, 2002.
- Foster, J.C., W. Rideout, B. Sandel, W.T. Forrester, F.J. Rich, On the relationship of SAPS to storm-enhanced density, *J. Atmos. Sol.-Terr. Phys.*, *69*, 303, 2007.
- Goldstein, J., Spasojevic, M., Reiff, P. H., Sandel, B. R., Forrester, W. T., Gallagher, D. L. and Reinisch, B. W., Identifying the plasmopause in IMAGE EUV data using IMAGE RPI in situ steep density gradients, *J. Geophys. Res.*, *108*, 1147, doi:10.1029/2002JA009475, 2003a.
- Goldstein, J., Sandel, B. R., Forrester, W. T., and Reiff, P. H., IMF-Driven plasmasphere erosion of 10 July 2000, *Geophys. Res. Lett.*, *30*, 1146, doi:10.1029/2002GL016478, 2003b.
- Goldstein, J., Borovsky, J., Foster, J., and Carpenter, D., Formation, dynamics, and impact of plasmaspheric plumes, *EOS Trans. AGU*, *88*, 247, 2007.
- Goldstein, J. and B. Sandel, Space Weather Effects of the Earth's Plasmasphere, *AMS Meeting, Fourth Symposium on Space Weather*, San Antonio, 2007.
- Huba, J.D., G. Joyce, and J.A. Fedder, SAMI2 (Sami2 is Another Model of the Ionosphere): A New Low-Latitude Ionosphere Model *J. Geophys. Res.*, *105*, 23,035, 2000.

- Huba, J. D., G. Joyce, J. Krall, C. S. Siefring and P. A. Bernhardt, Self-consistent modeling of dawn density depletions with SAMI3, *Geophys. Res. Lett.*, 2010
- Jordanova, V.K., L.M. Kistler, J.U. Kozyra, G.V. Khazanov, A.F. Nagy, Collisional losses of ring current ions. *J. Geophys. Res.*, 101, 111, 1996.
- Kim K.-H., J. Goldstein, and D. Berube, Plasmaspheric drainage plume observed by the Polar satellite in the prenoon sector and the IMAGE satellite during the magnetic storm of 11 April 2001, *J. Geophys. Res.*, 112, A06237, doi:10.1029/2006JA012030, 2007.
- Krall, J., J. D. Huba, and J. A. Fedder, Simulation of field-aligned H^+ and He^+ dynamics during late-stage plasmasphere refilling, *Ann. Geophys.*, 26, 1507–1516, 2008.
- Lemaire, J. and K.I. Gringauz, *The Earth's Plasmasphere*, [Cambridge University Press, New York, NY], 1998.
- Liemohn, M. W., J. U. Kozyra, M. F. Thomsen, J. L. Roeder, G. Lu, J. E. Borovsky, T. E. Cayton, Dominant role of the asymmetric ring current in producing the stormtime Dst*, *J. Geophys. Res.*, 106, 10883, 2001.
- Liemohn, M. W., A. J. Ridley, D. L. Gallagher, D. M. Ober, J. U. Kozyra, Dependence of plasmaspheric morphology on the electric field description during the recovery phase of the 17 April 2002 magnetic storm, *J. Geophys. Res.*, 109, A03209, 2004.
- Lin, C.S., H.-C. Yeh, B.R. Sandel, J. Goldstein, F.J. Rich, W.J. Burke, and J.C. Foster, Magnetospheric convection near a drainage plume, *J. Geophys. Res.*, 112, A05216, doi:10.1029/2006JA011819, 2007.
- Lyon, J. G., Fedder, J. A., and Mobarry, C M., The Lyon-Fedder-Mobarry (LFM) global MHD magnetospheric simulation code, *J. Atmos. Terr. Phys.*, 66, 1333, 2004.
- Mitchell, D.G., et al., High energy neutral atom (HENA) imager for the IMAGE mission, *Space Sci Rev.*, 91, 67, 2000.
- Nishida, A., Formulation of plasmapause, or magnetospheric plasma knee, by the combined action of magnetospheric convection plasma escape from the tail, *J. Geophys. Res.*, 71, 5669, 1966.
- Ober, D., J. Horwitz, and D. Gallagher, Formation of density troughs embedded in the outer plasmasphere by subauroral ion drift events, *J. Geophys. Res.*, 102, 14,595, doi:10.1029/97JA01046, 1997.
- Ridley, A.J. and M.W. Liemohn, A model-derived storm time asymmetric ring current driven electric field description, *J. Geophys. Res.*, 107, 1151, 2002.
- Sandel, B.R., et al., The extreme ultraviolet imager investigation for the IMAGE mission, *Space Sci Rev.*, 91, 197, 2000.
- Sandel, B. R., King, R. A., Forrester, W. T., Gallagher, D. L., Broadfoot, A. L., and Curtis, C. C., Initial results from the IMAGE Extreme Ultraviolet Imager, *Geophys. Res. Lett.*, 28, 1439, 2001.
- B. R. Sandel and M. H. Denton, Global view of refilling of the plasmasphere *Geophys. Res. Lett.*, 34, L17102, doi:10.1029/2007GL030669, 2007.
- Slinker, S.P., J.A. Fedder, J. Chen, and J.G. Lyon, Global MHD simulation of the magnetosphere and ionosphere for 1930-2330 UT on November 3, 1993, *J. Geophys. Res.*, 103, 26,243, 1998.

- Slinker, S.P., J.A. Fedder, W.J. Hughes, and J.G. Lyon, Response of the ionosphere to a density pulse in the solar wind: simulation of traveling convection vortices, *Geophys. Res. Lett.*, *26*, 3549, 1999a.
- Slinker, S. P., J. A. Fedder, B. A. Emery, K. B. Baker, D. Lummerzheim, J. G. Lyon, and F. J. Rich, Comparison of global MHD simulations with AMIE simulations for the events of May 19-20, 1996, *J. Geophys. Res.*, *104*, 28379, 1999b.
- Sojka, J. J., and Wrenn, G. L., Refilling of geosynchronous flux tubes as observed at the equator by GEOS 2, *J. Geophys. Res.*, *90*, 6379, 1985
- Song, X., Gendrin, R., and Caudal, G., Refilling process in the. plasmasphere and its relation to magnetic activity, *J. Atmos. Terr. Phys.*, *50*, 185, 1988
- Thompson, D. C., L. Scherliess, J. J. Sojka, and R. W. Schunk, Plasmasphere and upper ionosphere contributions and corrections during the assimilation of GPS slant TEC, *Radio Sci.*, *44*, RS0A02, doi:10.1029/2008RS004016, 2009.
- Yizengaw, E., J. Dewar, J. MacNeil, M. B. Moldwin, D. Galvan, J. Sanny, D. Berube, and B. Sandel, The occurrence of ionospheric signatures of plasmaspheric plumes over different longitudinal sectors, *J. Geophys. Res.*, *113*, A08318, doi:10.1029/2007JA012925, 2008a.
- Yizengaw, E., M. B. Moldwin, D. Galvan, B.A. Iijima, A. Komjathy, A.J. Mannucci, Global plasmaspheric TEC and its relative contribution to GPS TEC, *J. Atmos. Solar-Terr. Phys.*, *70*, 1541, 2008b.
- Zheng, Y., P. C:son-Brandt, A. T. Y. Lui, and M.-C. Fok, On ionospheric trough conductance and subauroral ion drift: Simulation results, *J. Geophys. Res.*, *113*, A04209, doi:10.1029/2007JA012532, 2008.

Diffusion and transport near the turbopause

G. Lehmacher and M. F. Larsen

Department of Physics and Astronomy, Clemson University, Clemson, South Carolina,
USA

Summary

The purpose of this white paper is twofold. First, it highlights the fact that our understanding of the underlying dynamics responsible for turbulent diffusion near the so-called turbopause is poor. The limited available data show evidence for enhanced diffusion, i.e., greater than molecular diffusion, above the nominal turbopause, as well as evidence for two-dimensional turbulence processes, but the processes responsible for diffusion in this critical region are still poorly understood. There is also very limited information available about the seasonal, geographical, and day-to-day variations in the turbulence levels and diffusion coefficients. Second, it highlights the need for a more extensive set of in-situ observations, of the type that can be provided by simple sounding rocket payloads, in order to make progress on these problems.

Introduction

Sounding rocket measurements have played a critical role in studies of the ionosphere since the late 1950's. In the 1960's, a regular and extensive program of sounding rocket observations of the thermosphere and ionosphere was undertaken, including measurements of the plasma densities, electric fields, neutral winds, and composition over a broad range of geographical locations, seasons, and local times. The peak NASA launch rate in the mid 1960's exceeded 90 launches per year. An example of the information obtained by the rocket observing program is the evidence that was found for the existence of a region of large winds and strong shears in the altitude range between 95 and 115 km, based on more than 400 rocket wind measurements (*Larsen, 2002*). The existence of the large winds and near-critical shears has been confirmed recently by ground-based Doppler lidar measurements that extend to altitudes near 100 km, but the characteristics of the wind maximum would have been difficult to resolve without the rocket measurements.

In the 1980's a transition occurred in the sounding rocket, both in the U.S. and in other countries. As more sophisticated ground-based techniques became available, the need for high-resolution in-situ measurements that only sounding rockets can provide was still recognized, but the number of launches became much more limited and the experiments became more complex. There is no question that there is a need for the type of high-resolution, multiple-instrument, in-situ measurements that comprise the majority of the current sounding rocket program, but we argue here that there are equally compelling reasons to carry out more extensive measurements with simple payloads and smaller, less expensive vehicles. Although ground-based measurements are now providing much more extensive measurements than in the 1960's, for example, they still resolve poorly or

do not cover at all the critical altitude range in the lower thermosphere and D and E region that rockets cover well.

In the next section, we describe a specific critical science question that serves as an example of the need for more extensive high-resolution in-situ measurements.

Turbulent diffusion above the turbopause?

In the region from approximately 90 to 110-km altitude, horizontal wind speeds and shears increase rapidly with height. Wind speeds of 150 m/s and shears exceeding 40 m/s per km, i.e., a critical value for shear instability, are common. The temperature structure in the region is characterized by a succession of very stable and nearly unstable layers. Large-amplitude waves abound and disappear on hour time scales. A common feature of all trimethyl aluminum (TMA) trails released in the mesosphere and lower thermosphere region is that a sharp transition from three-dimensional fluctuations within the trail to a laminar flow in which the trail becomes smooth occurs over an altitude range of a few hundred meters or less and is usually found at altitudes between 100 and 105 km. The TMA is a tracer of the neutral motions, and the change in the turbulent fluctuations is therefore also indicative of a change in the dynamics of the flow.

In the atmosphere, molecular diffusion increases with altitude as the atmospheric density decreases. The transition height where three-dimensional structure disappears and the flow becomes laminar is taken to be the turbopause height, where molecular diffusion becomes large enough to damp out the small-scale fluctuations in the flow (see, e.g., the early paper by *Justus (1967)*). Other definitions of the turbopause are certainly possible. For example, wave dissipation or wave parameters can be used (see, e.g., *Ebel (1984)*), but the transition evident in the trails is associated with the passive scalar transition marked by the change in the behavior of the tracer.

The dynamical argument used to explain the turbopause transition evident in the trail images is straightforward, but the results in two cases where the image data have been analyzed quantitatively in detail have yielded a surprising result. Based on the simple argument, the expansion of the trail is expected to be due solely to the action of molecular diffusion above the transition height. Below the transition, eddy or turbulent diffusion enhances the expansion rate where turbulence is present. The strength of the turbulence will vary with altitude, depending on shears, thermodynamic stability of the background atmosphere, and wave processes, but the diffusion rates below the transition are generally greater than those expected from molecular diffusion alone at all heights.

The relationship between the trail expansion rates and the diffusion coefficients and turbulent energy dissipation rates was used to exploit the trail releases as a means for measuring the variation of the turbulence parameters with altitude. The study by *Justus (1967)* was the first attempt to make use of the technique, but the paper by *Rees et al. (1972)* gave the first detailed results, due largely to the availability of high-sensitivity Baker-Nunn cameras that provided high spatial and temporal resolution images of releases made from Woomera, Australia. The results in the latter study showed

interesting variations in the turbulent diffusion below the nominal turbopause altitude but also showed an enhanced or anomalous expansion in an altitude range that extended 10 to 15 km above the nominal turbopause. In that altitude range, the trails were found to expand significantly more rapidly than could be explained by molecular diffusion alone. A similar analysis of trail expansion properties was carried out for the releases in the TOMEX experiment launched from White Sands, New Mexico, in October 2000 (*Bishop et al., 2004*). The conclusions were consistent with those in the earlier *Rees et al.* paper, namely that the trail expansion rates were significantly greater than expected for molecular diffusion in an altitude range that extended at least 10 km above the nominal turbopause.

A sequence of images of trails from four launches at Poker Flat, Alaska, over a period of approximately two hours in February 2009 is shown in Figure 1 below. The nominal turbopause transition is evident. The sequence shows both a degree of continuity from profile to profile and also small-scale variations as the turbulence evolves with time.

In the 10 or 15-km altitude range above the nominal turbopause, the vertical structure, i.e., the small-scale structure parallel to the trail, is damped, but the horizontal diffusion rates are greater than those associated with molecular diffusion alone. The implication is therefore that the vertical structure is damped but the horizontal structure remains. In other words, the evidence suggests a transition from a three-dimensional turbulence process to a quasi-two-dimensional turbulence process.

In summary, the analyses from the earlier experiments show enhanced or anomalous diffusion rates in a 10 to 15 km altitude range above the nominal turbopause. The dynamics responsible for the effect are not understood, but it seems clear that there is a change in the turbulence process that occurs at the nominal turbopause height. The enhanced diffusion is interesting because it is a poorly understood physical phenomenon, but it also has significant consequences because of the altitude range where it occurs. Specifically, altitudes between 100 and 115 km represent an altitude range where the ionospheric conductivity increases rapidly and electrodynamical and plasma physics effects become important. There is also a rapid transition in the concentration of various minor chemical species and metallic layers associated with meteor ablation in that altitude range. The same altitude range is also a region where small-scale wave breaking effects are important and where the vertical scales of large-scale tides and planetary waves change rapidly with height. Larger than expected diffusion coefficients and the associated anomalous transport is expected to have a significant effect on all those processes.

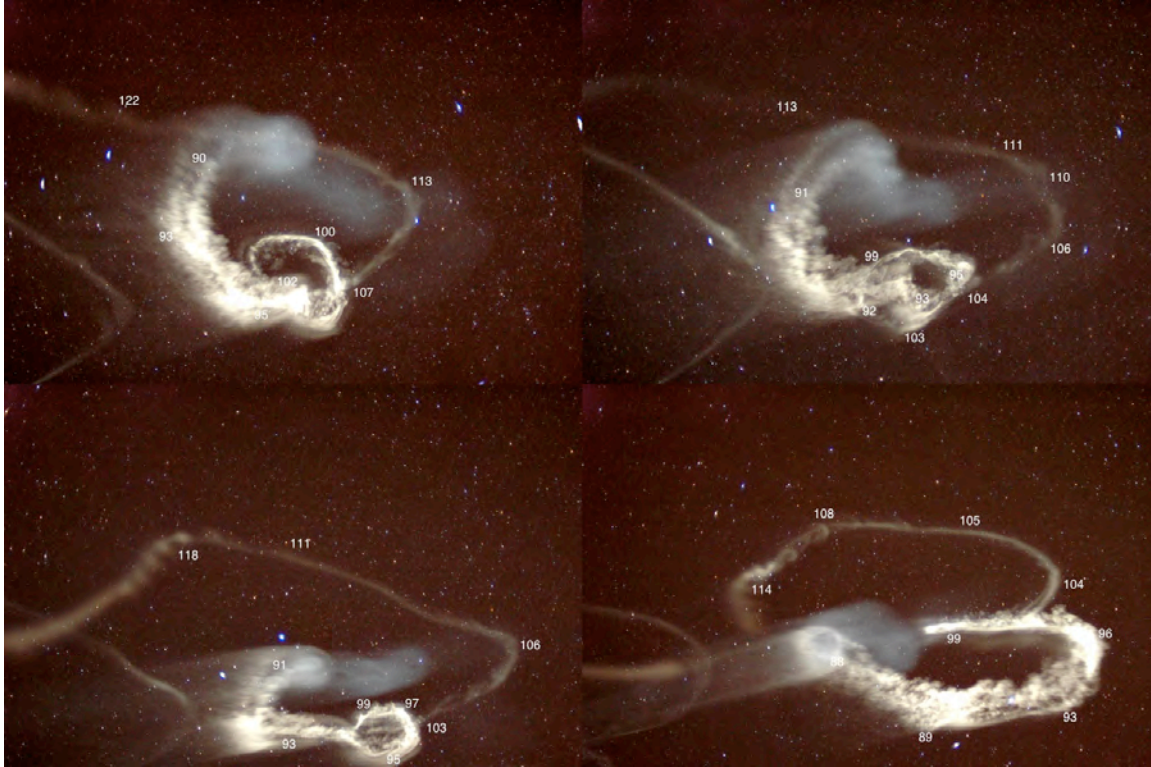


Fig. 1. Sequence of 4 chemiluminescent trails demonstrating wave and turbulent processes and the transition from small scale turbulence to molecular diffusion.

Large-scale transport

The results described in the papers by *Larsen (2002)* and *Larsen and Fesen (2009)* show that large wind speeds and local wind maxima are expected in the lower thermosphere, but it is not clear from the statistical distributions if the large winds vary rapidly with time and over small spatial scales or if they represent larger-scale coherent structures capable of producing enhanced transport. However, evidence of large-scale coherence is provided by the results of several studies of the global scale transport of Shuttle and other large rocket exhaust released in the altitude range between 100 and 110 km by *Siskind et al. (2003)* and *Stevens et al. (2002, 2003, 2005)*. Their studies show that there is a consistent rapid transport of the exhaust products from the southeastern coast of North America, where the Shuttle and the rockets are launched, to the Arctic or Antarctic regions over a period a few days. This type of transport implies a large-scale circulation with large mean winds in the same altitude range where the large rocket winds are observed. The time-scale and magnitude of the circulation, as well as the global-scale transport from low to high latitudes, is consistent with the transport expected for large-scale Rossby waves, for example, which have a similar behavior in the lower atmosphere.

Remote observations by limb-sounding satellites have provided long-term global coverage of lower thermospheric winds (e.g. HRDI, TIDI) and temperatures (e.g., SABER) measurements. However, neither of these techniques has the altitude resolution to resolve the wind shears and temperature gradients that are observed with the rocket

techniques. Sophisticated temperature and wind lidars, as well as meteor radars, have produced long-term studies of MLT dynamics at specific sites, but the measurements are generally limited to altitudes below the critical transition described in the previous section.

Atmospheric dynamics in the turbopause region leads to instabilities, nonlinear wave-wave interactions, and energy and enstrophy cascades, described as turbulence or eddy diffusion. While radars can sometimes observe indirect effects of neutral turbulence or wave processes as plasma waves or plasma turbulence that produce coherent echoes, only sounding rockets are able to observe the neutral small-scale dynamics directly. In this region, scales of tens to hundreds of meters are relevant. The time evolution of chemical releases from rockets, in combination with high-resolution density and temperature sampling, provides detailed sampling of such small-scale dynamics.

The principles of momentum transport in the mesosphere and the zonal wind reversal are reasonably well understood thanks to systematic rocket observations in the 1960s and 1970s. Later ground-based and satellite observations and global circulation models complemented our picture of these processes in the mesosphere. The neutral dynamical processes of the lower thermosphere, especially at small scales, are still very poorly understood. Tides are the most prominent wave modulation in this region, but they do not appear to explain the large-scale transport and enhanced turbulent diffusion observed in this altitude range.

In summary, there is still very little known about the seasonal or geographic variations in the turbulence processes or turbulent diffusion in the critical region between 100 and 120 km altitude and the neutral instabilities that drive such processes. Since the neutral structures are poorly understood and generally not measured, the coupling between the neutral and ionized component of the atmosphere in the region is also poorly understood.

Science questions

A list of some of the critical science questions is as follows:

- What are the instabilities responsible for neutral wave and turbulence structure in the lower thermosphere?
- How long do these instabilities persist and where do they occur?
- What is the effect of horizontal and vertical diffusion in that part of the atmosphere?
- What is the impact on vertical mixing, esp. chemically active species, O, OH*, NO?
- What is the impact of dynamical processes in that altitude range on large-scale transport?
- What is the consequence for momentum transport and mean flow acceleration?

- How do models accomplish momentum and tracer transport in the turbopause region?
- What is the relative importance of quasi-two-dimensional and three-dimensional turbulence processes?

Conclusion

The high-resolution measurements and direct observations of neutral processes that sounding rockets can provide in the lower thermosphere are critical in obtaining answers to the science questions posed above, either by providing critical information directly or by providing information that is critical to the underlying processes. The infrequent experiments that have been carried out in the last decade have provided important new information on a case study basis, but a more extensive observing program is required to understand the geographical, seasonal, and day-to-day variation in the critical turbulence and neutral structure parameters. The structure of the current sounding rocket program is poorly suited to this type of program, in part because of budget limitations but also because of the mission-oriented perspective within the program. More extensive observations are clearly possible, since such observations have been carried out in the past. Many of the science issues outlined above can be addressed with relatively simple and low-cost payload packages flown on low-cost surplus vehicles with low apogees. A high launch rate, however, would require a more standardized vehicle and payload design and an acceptance of a somewhat higher failure rate than is currently considered to be acceptable so that the pre-launch testing requirements can be reduced.

References

- Bishop, R. L., M. F. Larsen, J. H. Hecht, A. Z. Liu, and C. S. Gardner (2004): TOMEX: Mesospheric and lower thermospheric diffusivities and instability layers, *J. Geophys. Res.*, 109, 10.1029/2002JD003079.
- Ebel, A. (1984): Contributions of gravity waves to the momentum, heat and turbulent energy budget of the upper mesosphere and lower thermosphere. *J. Atmos. Terr. Phys.*, 46, 727–737.
- Justus, C. G. (1967): The eddy diffusivities, energy balance parameters, and heating rate of upper atmospheric turbulence. *J. Geophys. Res.*, 72, 1035–1039.
- Larsen, M. F. (2002): Winds and shears in the mesosphere and lower thermosphere: Results from four decades of chemical release wind measurements, *J. Geophys. Res.*, 107, 10.1029/2001JA000218.
- Larsen, M. F., and C. G. Fesen (2009): Accuracy issues of the existing thermospheric wind models: Can we rely on them in seeking solutions to wind-driven problems?, *Ann. Geophys.*, 27, doi:www.ann-geophys.net/27/1/2009/.

Rees, D., R. G. Roper, K. H. Lloyd, and C. H. Low (1972): Determination of the structure of the atmosphere between 90 and 250 km by means of contaminant releases at Woomera, May 1968. *Proc. Roy. Soc. London, Ser. A*, 271, 631–663, 1972.

Siskind, D. E., M. H. Stevens, J. T. Emmert, D. P. Drob, A. J. Kochenash, J. M. Russell III, L. L. Gordley, and M. G. Mlynczak (2003): Signatures of shuttle and rocket exhaust plumes in TIMED/SABER radiance data, *Geophys. Res. Lett.*, 30, 1819, doi:10.1029/2003GL017627.

Stevens, M. H., C. R. Englert, and J. Gumbel (2002): OH observations of space shuttle exhaust, *Geophys. Res. Lett.*, 29, 1378, 10.1029/2002GL015079.

Stevens, M. H., J. Gumbel, C. R. Englert, K. U. Grossmann, M. Rapp, and P. Hartogh (2003): Polar mesospheric clouds formed from space shuttle exhaust, *Geophys. Res. Lett.*, 30, 1546, doi:10.1029/2003GL017249.

Stevens, M. H., R. R. Meier, X. Chu, M. T. DeLand, and J. M. Plane (2005): Antarctic mesospheric clouds formed from space shuttle exhaust, *Geophys. Res. Lett.*, 32, doi:10.1029/2005GL023054.

Small-scale neutral-ion coupling in the mesosphere

Gerald Lehmacher

Department of Physics and Astronomy, Clemson University, Clemson, SC 29631

Summary

This white paper addresses the fundamental science questions “What governs the coupling of neutral and ionized species in the mesosphere?” and “What is responsible for the variability of the ionization layers in the mesosphere?” We urge to include mesospheric investigations in the revitalization of the sounding rocket program with multiple benefits and broad impacts. Sounding rockets are needed for investigation of the small scale dynamics, ionization layers, and complex ion chemistry between 50 and 90 km. Simple and small sounding rockets measuring large and small-scale neutral temperature and ionization structure should be used in combination with large and sophisticated ground-based instrumentation, such as lidars and radars. The goal is to obtain coincident observations of the radar scattering layers in the middle mesosphere (60-80 km), which are crucial for our fundamental understanding of the mesosphere and D region.

Background

The mesosphere is the atmospheric layer from 50 to about 100 km, between the stratosphere and thermosphere. It is characterized by the upward propagating gravity waves, which deposit their momentum and lead to the closure and reversal of the mesospheric jets. The importance of dynamical coupling between the lower and upper atmosphere is strongly emphasized in the 2003 NRC Decadal Survey, A-I-M Panel Report, as well as in the 2009 Heliophysics Roadmap, and the mesosphere is the poster child for these processes, including small-scale dynamics and turbulent mixing. The latest draft NSF CEDAR Strategic Plan recommends the utilization of large ground based facilities, as well the development of innovative sounding rocket payloads.

The global meridional circulation with mean upwelling in the summer hemisphere leads to the extremely cold summer mesopause (<130 K), marvelous noctilucent cloud displays, and the accumulation of charged aerosol layers, while the mean downwelling and NO transport in the winter hemisphere is often disrupted by stratospheric warmings and reorganization of the polar vortex. At equatorial latitudes, semiannual variability is dominant, which is further modulated by the quasi-biannual oscillation, both manifestations of atmospheric wave coupling. There is also large variability on intermediate timescales, such as planetary waves, variable tidal forcing, and persistent temperature inversion layers.

The mesosphere is an important part in the A-I-M system, since it is both strongly under atmospheric and solar control. The mesosphere coincides approximately with the ionospheric D region, which very much dependent on solar ionizing

radiation, but also affected by solar storms via energetic particles, solar X rays and bremsstrahlung. The new solar cycle will give new opportunities to study the variability of D region ionization with increasing and fluctuating solar input. The variability of the neutral mesosphere and the penetration of the solar control is also an important goal in the Heliospheric Roadmap that has not been addressed. The ionization is contained in a wide spectrum of charged species, from free electrons, light ions, and heavy hydrated ions, positive and negative, up to nanoparticles of meteoric origin, which likely form transient layers in wind shear zones or convective regions.

Many of these properties have been discovered first by comprehensive and systematic rocket experiments beginning in the late 1950s through the 1970s and 1980s. Then satellites, in particular LIMS, SME, UARS, and TIMED, have provided a global picture, with the time and spatial resolution restricted due to the satellite sampling geometry. Lidars have difficulty in the middle mesosphere (50-80 km), below the metal layers; large powerful systems are needed for good time and altitude resolution.

Radars have been used in the D region for long time, most successful, MF radars for wind estimates, VHF radars for wind and turbulence measurements, and UHF radars for wind and electron density observations. However, the scattering process is still largely unknown and the in situ observation of scattering layers has not been accomplished, except for comprehensive studies of Polar Mesospheric Summer Echoes (PMSE), which are relatively easy and reliably observable by radar. Recently, the origin of low-latitude echoes and winter echoes has been questioned, and several theories have been brought forward.

Science questions

What neutral-ionosphere coupling processes govern the mesosphere?

What is the origin of mesospheric radar echoes at MF and VHF frequencies under normal conditions, i.e. the main scattering region 70-80 km, and non-PMSE?

What is the role of heavy ions in the scattering layers?

What is the role of ion-acoustic waves in the scattering process?

What is the nature of turbulence in the scattering layers?

What is the electron density, neutral wind and temperature associated with the scattering layers?

Implementation

Sounding rockets are the only way to observe small scale neutral and plasma structure the mesosphere and help solving these puzzling questions, some of which are decades old. The importance of the NASA Sounding Rocket Program as platform for scientific discovery and education has been pointed out in the last NRC and NASA surveys. The revitalization of the sounding rocket activities must include the mesosphere, the layer that can be reached with relatively simple or small rockets, while our understanding would benefit greatly. We do note that some mesospheric application may be appropriate for commercial flights to the mesosphere.

We emphasize the need for small mesospheric payloads that can be launched near large ground based radar and lidar observatories. This recommendation has also been addressed in the latest findings of the NASA Sounding Rocket Working Group (SRWG) . Several options have been discussed, from simple passive payloads, to standardized instrumented payloads. For the specific issue addressed here, a payload with a sensitive positive ion probe and a fast ionization gauge for mesospheric pressures would be appropriate. The development or use of instruments for measuring cluster ions, and heavier charged species is closely related.

Conclusion

Mesosphere should be part of the vocabulary of the new decadal survey, and not just as part of the acronym TIMED. In the future mesosphere should be no longer be equated with “ignorosphere”, as is still commonplace. New, systematic rocket investigations dedicated to the mesosphere should help achieving this goal.

White Paper for the NRC Decadal Survey for Solar and Space Physics

PREDICT NEUTRAL DENSITY

**Chin S. Lin, Space Vehicles Directorate, Air Force Research Laboratory
and
Frank A. Marcos, Boston College**

Orbital aerodynamic drag continues to be the largest uncertainty in determining trajectories of satellites operating in Earth's upper atmosphere below about 600 km. Neutral density of the upper atmosphere (thermosphere) is the major uncertainty in determining satellite drag. Winds are significant particularly during geomagnetic storms. Therefore understanding and forecasting the global density and wind fields permits accurate satellite drag forecasts. Many manned and unmanned space missions operating in Low Earth Orbit require forecasts of satellite drag. One major thermospheric variability societal impact is the space debris threat to these missions. Air Force Space Command (AFSPC) has responsibility for providing Collision Avoidance Warnings for all satellites. Their goal is 72 hour drag forecasts within 5%. AFSPC needs upgraded operational assimilative satellite drag space weather for precise knowledge of locations of all satellites and predicting where reentry will occur. Current AFSPC forecasts during unstable orbital drag environment conditions may be in error by as much as 50%.

Since the thermosphere is driven by solar EUV and solar wind inputs as well as upward propagating disturbances these are the parameters that must be forecast. The major space weather challenge is forecasting the spatial and temporal response to geomagnetic storms. Large geomagnetic storms produce large disturbances in the upper atmosphere with wind speeds as high as 1 km/s. In addition, Solar EUV activity is currently near its minimum but is expected to reach maximum values in 2013. During solar max, orbital predictions are far less accurate and the risk of collisions increases significantly; more objects tend to deviate from their expected paths and more reenter. While solar cycle effects dominate on longer time scales, a strong space storm can typically change the neutral density at 400 km by a factor of 3 in just a few hours.

Current state of art density models are semi-empirical representations of the atmosphere and are based in part upon proxy of solar and geomagnetic heating inputs. These empirical models can adequately describe the neutral density global mean structure within 15-20% during quiet times. However, errors of 100% are not unusual during disturbed periods. A significant advance in specification of satellite drag was achieved by The Air Force High Accuracy Satellite Drag Model (HASDM). This approach assimilates drag data from some 100 calibration satellites to correct the global density fields of an empirical model to reduce errors to about 4-8%. Due to coverage issues, accuracy at high latitudes is more limited, particularly during geomagnetic storms. Further, given its temporal and spatial resolution it will provide only limited accuracy during geomagnetic storms even with timely forecasts of appropriate indices.

Significant improvements beyond this state will come from the use of physics-based atmospheric density models in place of the current semi-empirical models. Considerable RDT&E is required from the space physics and astrodynamics communities for this advancement to occur. Additional capabilities to understand and forecast the satellite drag dependence on the space weather environment are coming from new solar and geomagnetic observations plus the large direct satellite-borne density data sampling of the atmosphere vs latitude, local time, day of year, solar cycle and geomagnetic activity from the CHAMP and GRACE satellites and orbital drag databases generated by AFSPC. The satellite drag prediction accuracy, especially over a 72 hour prediction period, is found to be mainly a function of the predicted

accuracy of the solar and geomagnetic indices coupled with the accuracy of the atmospheric model using these predicted indices. Physics-based models offer an advantage by providing realistic heating inputs and detailed, time-dependent structure not available in the semi-empirical models. Knowledge of the required fundamental heating processes resulting from the EUV and solar wind interaction with the coupled magnetosphere–ionosphere–thermosphere system are in an active research state.

The Multidisciplinary University Research Initiative (MURI) with the University of Colorado (Dr. Jeff Forbes, PI) was created by Air Force Office of Scientific Research (AFOSR) to explore the physical mechanisms that drive the Earth's thermospheric dynamics. Candidate physical modeling tools are available including the physics-based neutral density model called the Coupled Thermosphere-Ionosphere Model (CTIM). Considerable research has contributed to development of NCAR's Thermosphere-Ionosphere Electrodynamics General Circulation Model (TIEGCM). The MURI program attempts to address current model limitations through scientific analyses of solar phenomena and available satellite drag data. The scope of this MURI topic is not the development of a new physics-based model but "the elucidation of those physical concepts and predictable key indicators of energy inputs to the atmosphere that will update and calibrate current operational satellite drag models and lead to the accurate prediction of thermospheric densities and, hence, a precise prediction of the locations of satellites. The intent of this MURI is to develop fundamental concepts that will eventually be included in the large models." The MURI effort has been divided into seven focus areas including Forecasting Geomagnetic Activity, Forecasting Solar EUV/UV Radiation and Wave Forcing from the Lower Atmosphere. A variety of solar disturbance and solar wind propagation models are applied to this research. In addition to developing forecast tools the MURI effort is investigating the response of the thermosphere to heating inputs and, with the Whole Atmosphere Model (WM), to investigate upward propagating waves.

The fundamental next step after the MURI is to incorporate the benefits of a physical model within the assimilation process. Initializing the physical model with both direct in-situ and indirect observations in a statistically rigorous manner provides a practical approach for representing the time-dependent conditions of the thermosphere. Implementation of a Kalman filtering data assimilation technique by the space physics community has proven to be non-trivial. It requires detailed understanding of the physics based model as well as estimation theory and astrodynamics. Thus, this development requires cooperative efforts of the space physics and astrodynamics communities. Employing physics-based neutral density models coupled with an assimilation technique should be the basis of next-generation drag model improvements for satellite cataloging.

In summary, uncertainties in neutral density variations have been the major error source for LEO determination. The problem is being vigorously and fruitfully attacked by numerous space weather studies including data assimilation schemes, predictive solar indices, and in-situ measurement. While there is still a lot of research needed, the tools are finally becoming available. The culmination of these efforts will be steady progress in meeting the evolving, previously unattainable, stringent requirements for operations in the satellite drag environment. A key research goal of the AFOSR MURI is a physics-based understanding and prediction capability for the spatial and temporal distribution of thermospheric energy sources. This critical basic research will lead to a near-real-time, accurate operational capability to locate, track identify, and estimate future locations of satellites with high accuracy. With the additional underpinning of this effort, we anticipate significant future progress in satellite drag capability.

White Paper for the NRC Decadal Survey for Solar and Space Physics

CUBESAT ORBITAL DRAG EXPERIMENT

Chin S. Lin, Space Vehicles Directorate, Air Force Research Laboratory
and
Frank A. Marcos, Boston College

Abstract

A major thermospheric variability societal impact is the space debris threat to manned missions and to unmanned civilian commercial and scientific payloads. Exploiting the CubeSat opportunity provides the unprecedented capability of long-term, routine high-accuracy measurements of thermospheric variability by accelerometers.

A major thermospheric variability societal impact is the space debris threat to manned missions and to unmanned civilian commercial and scientific payloads. Understanding the global density and wind fields permits accurate debris orbit prediction and determination of debris object cross sections (by eliminating satellite drag variation errors). Orbital aerodynamic drag continues to be the largest uncertainty in determining trajectories of satellites operating in Earth's upper atmosphere below about 600 km. Neutral density of the upper atmosphere (thermosphere) is the major uncertainty in determining satellite drag. Air Force Space Command (AFSPC) has responsibility for providing Collision Avoidance Warnings for all satellites. Their goal is 72 hour drag forecasts within 5%. AFSPC needs upgraded operational assimilative satellite drag space weather particularly for: avoiding collisions with manned and unmanned missions, precise knowledge of locations of all satellites and predicting where reentry will occur.

The development of small, cheap reliable thermosphere monitoring tools provides the opportunity for future routine monitoring of space weather analogous to the current state in tropospheric weather. Synoptic global measurements can be made by developing inexpensive vehicles for space access similar to today's sounding rockets. The capability for repeated, long-term, inexpensive thermospheric monitoring leads to the needed capability to predict environmental changes due to both natural factors and human activities over the solar cycle.

Exploiting the CubeSat opportunity provides the unprecedented capability of long-term, routine high-accuracy measurements. Small, inexpensive sensors are available or coming on-line. These include the NASA winds and composition instruments, the miniature accelerometer under development by AFRL, and a micro-plasma Langmuir probe from AFRL. Current technology provides the needed Data Processing Unit, Data Storage, Electric Power System, Downlink Transceiver, Attitude Determination and Control System, Global Positioning System and Thermal Management. The objective is a very low cost, low weight and power space qualified miniature coordinated capability of high-accuracy and high resolution Space Weather monitoring.

The properties to be measured by the proposed small accelerometers on the CubeSat experiment are: total atmospheric density, neutral composition, satellite drag, neutral winds and relevant ionospheric processes. These are critical parameters for understanding the dynamic processes that affect the variability of the upper atmosphere and for driving assimilative operational space weather models. Since a primary goal is to measure geomagnetic storm phenomena when winds are of the order of 500-1000 m/s, we anticipate that our cross track and in-track winds will provide critical data with the sufficient accuracy needed. Lack of comprehensive winds measurements is a major limitation in developing realistic thermosphere and ionosphere models. These data will support the understanding of thermosphere dynamics that will

enhance the achievement of an accurate forecast capability. Distributed global monitoring is required to validate and upgrade these assimilative operational first principles models.

Persisting deficiencies in physical understanding of the thermosphere remains a limiting factor in Space Weather modeling. The thermosphere is driven by solar EUV and solar wind inputs as well as upward propagating disturbances. While solar cycle effects dominate on longer time scales, a strong geomagnetic storm can typically change the neutral density at 400 km by a factor of 3 in just a few hours. At 400 km, the neutral density changes by about a factor of 10 over the solar cycle. Significant daily variations are also common and changes of 20-50% can occur over the 27 day solar rotation period. Thermospheric density structure can be drastically modified by solar wind interactions at high latitudes. The solar wind, fast and dense ionized gas which interacts with the Earth's magnetic field, results in significant energy deposition in the upper atmosphere. This causes geomagnetic storms that heat the upper atmosphere unevenly and increases neutral mass density globally. A major space weather challenge is understanding the spatial and temporal response to geomagnetic storms. Major geomagnetic storms produce large disturbances in the upper atmosphere with wind speeds as high as 1 km/s. Data taken during the current solar minimum period, when solar and geomagnetic inputs were small have shown the importance of upward propagating tides as a persistent source of thermospheric variability.

Analyses of CHAMP and GRACE satellites have revealed several phenomena whose elucidation will permit detailed physical understanding of the thermospheric dynamics and variability. These include the dramatic effect of geomagnetic storms, with latitude dependent response and time delays and periodic oscillations, all related to solar wind conditions, density peaks in the polar cusp region, an equatorial thermosphere density anomaly analogous to the equatorial ionization anomaly and the finding that atmospheric tides, including a wave-4 structures, and observed longitudinal variability have a number of impacts on the thermosphere at both E and F region altitudes. These observations have been made with two satellites mainly at altitudes of 350 km and above, monitoring satellite drag. Despite these superb new data sources, empirical and physical models continue to have statistical errors of about 15-20%. During geomagnetic storms model errors can be a factor of two or greater.

Accurate modeling of the thermosphere requires comprehensive long-term solar cycle monitoring of all the neutral atmosphere state variables: neutral density, composition, winds and temperature as well as relevant ionospheric properties, coupled with drivers from solar, geomagnetic and upward propagating wave inputs. A constellation of relevant instruments can: (1) address missing physics in thermospheric models and (2) permit current empirical and forthcoming physics-based assimilative models to meet their potential. Physical models are inherently complex and their development for operational use depends upon a thorough understanding of the physical processes that limit our capability to meet space mission objectives. Complete understanding of satellite drag specification and forecast requires greatly improved understanding of magnetosphere-ionosphere-thermosphere coupling. This can only be done by collecting synoptic measurements of key atmospheric variables with very high accuracy sensors to specify and forecast the neutral atmosphere (and ionosphere) as a function of altitude, latitude, longitude, local time, season, solar flux and geomagnetic activity. The natural solution is a strategic constellation of probes that balances needs and resources combined with an integrated analysis and modeling effort.

Achieving a 5% 3-day forecast capability will require the following: (1) accurate specification of positions of all satellites, (2) accurate models of the upper atmosphere to project satellite positions forward in time and (3) accurate forecasts of upper atmosphere heat inputs that drive thermospheric variability. CubeSat constellations provide (1) a global database to initialize satellite position using assimilative models, (2) the requisite distributed data to understand thermosphere response to solar-geomagnetic inputs as a function of altitude, latitude, longitude local time and day of year and (3) a

capability to use physical modeling driver-response relations to back out what heat inputs are important and to utilize NASA, NOAA and DoD monitoring and forecasts of these inputs.

Conceptual Framework for Space Weather Dynamics: an Interplay of Large and Mesoscale Structure within the Nightside Magnetosphere-Ionosphere-Thermosphere System

Introduction

The highly varying structure and dynamics of the near-Earth space environment, including the energetic charged particles and fields within the magnetosphere, the neutral constituents of the thermosphere, and the ionosphere substantially affect man-made space systems and susceptible ground systems. Most space weather disturbances can be visually identified via production of bright and dynamic aurora that occur along plasma sheet magnetic field lines as a result of electrons and ions impacting the upper atmosphere. The plasma sheet is a vast space plasma reservoir extending 10's of Earth radii in the anti-solar direction, as shown in Figure 1. Understanding of space weather disturbances requires understanding the structure and dynamics along plasma sheet field lines, including the interplay of the plasma sheet with the thermosphere and ionosphere. However, while critically needed, in depth understanding has been limited because temporal changes are extreme and rapid, coupling to the ionosphere and thermosphere is complex, and there have been severe limitations on capabilities to measure and model the coupled system.

However, the opportunity to dramatically change this situation is now emerging due to NSF deployments of new and expanded radar systems and the first system for measuring the distributions of coupling electrical currents, the ground auroral imaging and multi-spacecraft of the NASA THEMIS program, new capabilities to measure the two-dimensional ionospheric density distribution, and modeling advances. First, because of ongoing deployment of new facilities, we now have vastly enhanced capability to evaluate the coupling between the plasma sheet and ionosphere. Due to electrical currents that flow along magnetic field lines (referred to as field-aligned currents, FACs), examining the flows and aurora in the ionosphere is like opening a window into the processes in the magnetosphere and plasma sheet. NSF has deployed new radars, which measure ionospheric plasma flows and related parameters, and AMPERE, which uses magnetometer data from the commercial constellation of Iridium satellites to give the first global view of the FAC systems that connect the ionosphere and magnetosphere. Also available now are the continental-scale, high-resolution all-sky auroral imager (ASI) array and multi-spacecraft of NASA's THEMIS program, designed for measuring the plasma sheet, and the dense array of Alaskan GPS receivers providing horizontal distributions of total electron content (TEC) and substantially enhancing capabilities for measuring ionospheric structures. Initial analyses of observations from these facilities have suggested a concept for viewing the processes leading to Space Weather dynamics based on an *interplay between a slowly-changing large-scale structure and much more dynamic meso-scale structures* associated with flow channels. Structures of both scales are expected to strongly affect

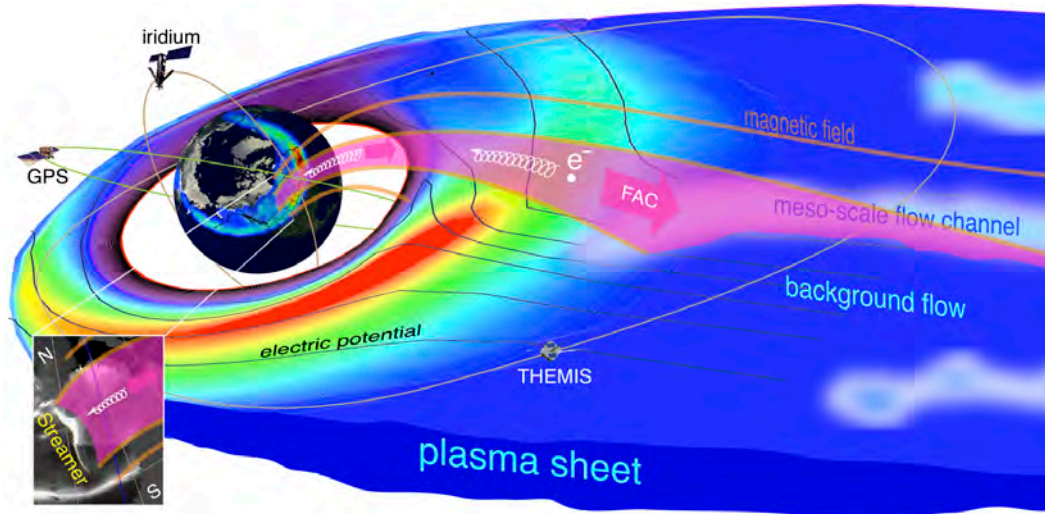


Fig 1. Coupling along magnetic field lines between plasma sheet particles and flows and the auroral ionosphere. Orbits of spacecraft to be used in the proposed study are indicated. Plasma pressures and electric potentials of the large-scale background from the RCM are shown in the equatorial plane, and the connection between meso-scale flow channels and auroral upward FAC features are illustrated.

ionospheric and thermospheric structure and dynamics in ways not previously understood. Additionally, recent analyses have revealed the unexpected possibility that flow structure formed within the high-latitude, polar-cap regions (which lie poleward of the plasma sheet and auroral oval along field lines connecting to the IMF) may drive plasma sheet meso-scale structures.

Model development has also added to our ability to evaluate and understand the above framework concept by reaching the point where new observations can be combined with the models to provide a far more comprehensive understanding of the integrated magnetosphere-ionosphere-thermosphere (MIT) system than has been previously feasible. This includes new capabilities for self-consistent modeling of MIT electrodynamical coupling, as well as for evaluation of how specific features of the coupling on both large- and meso-scales affect the structure and dynamics of the ionosphere and thermosphere, and their feed back on the coupling. With this unprecedented opportunity for system-wide observations and modeling, it is timely to focus on the following overall goals:

- 1. Determine the extent to which plasma sheet magnetosphere-ionosphere dynamics can be described as a slowly-varying large-scale state, with faster disturbances driven by embedded meso-scale flows.**
- 2. Determine and understand how the dynamics on both scales affect major features of ionospheric and thermospheric structure and dynamics, and their feedback on the full system dynamics.**
- 3. Determine the extent of, and effects from, the coupling between meso-scale structures within the polar cap regions and that within the plasma sheet and auroral oval.**

Attainment of these goals requires treating the magnetosphere, ionosphere, and thermosphere as a single, interdisciplinary field that includes plasma and neutral dynamics, and it requires coordination between analysis of data from the large variety of critical data sources and between the data and models. By taking advantage of the substantial new capabilities, a planned interdisciplinary approach has the potential for transforming our understanding of Space Weather dynamics by providing a coupled large-meso scale framework for both the more slowly evolving large-scale structure and the more rapidly varying, smaller-scale changes associated with the different types of disturbances.

Framework concept

Large- and meso-scale structure. Basic features of a slowly-evolving large-scale structure have been revealed by the Poker Flat Incoherent Scatter Radar (PFISR), SuperDARN radars, and low-altitude spacecraft [e.g., Lyons *et al.*, 2009a; Zou *et al.*, 2009a,b], and have been reproduced by a new version of the Rice Convection Model (RCM) [Gkioulidou *et al.*, 2009] with observationally based driving [Wang *et al.*, 2007]. The model shows development of the plasma sheet pressure distribution (illustrated in Figure 1), which leads to FACs flowing between the magnetosphere and the ionosphere. The FACs couple the ionosphere and magnetosphere and lead to an intimate linking between flows in the ionosphere and magnetosphere. This allows remote sensing of the flows in the magnetosphere, and more specifically within the plasma sheet, through radar measurements of ionospheric flows.

Figure 2 shows ionospheric FACs, height-integrated Pedersen conductivity, and equipotentials (solid curves in each panel) in the nightside ionosphere within 3 hr of midnight illustrating the slow time dependent evolution after the cross polar-cap potential drop ($\Delta\Phi_{PCP}$) was increased from 30 kV to 100 kV just after $T = 0$. Prior to the increase in $\Delta\Phi_{PCP}$, the model shows that the inner plasma sheet pressure (not shown) leads to a weak Region (R2) FAC system which maps to the ionosphere at $\Lambda \sim 66^\circ$ - 67° . The enhanced poleward directed electric fields on the pre-midnight side (close together equipotentials), which lie equatorward of the inner edge of the electron plasma sheet (at the model equatorward edge of enhanced Pederson conductivities), are subauroral polarization streams (SAPS) electric fields. These lie in the region of downward R2 currents driven by gradients of the strongest plasma sheet ion pressures.

The SAPS electric field can be seen to move equatorward within the ionosphere along with the electron plasma sheet after the increase in $\Delta\Phi_{PCP}$. There is also a gradual increase in plasma sheet pressures, pressure gradients, and R2 currents. Initially, the enhanced electric fields penetrate equatorward of the plasma sheet, but this gradually reduces. Some features of the above results have already been detected with PFISR, SuperDARN radar, and low-altitude spacecraft measurements [e.g., Lyons *et al.*, 2009a; Zou *et al.*, 2009a,b], demonstrating the potential realism of the RCM electrical coupling in determining the particle, current, and electric field structure along plasma sheet field lines.

Superposed on the large-scale structure are perturbations that can have a substantially larger magnitude than the background. Of particular importance for Space Weather disturbances are longitudinally localized bursts of equatorward directed flow that are often observed near the polar cap

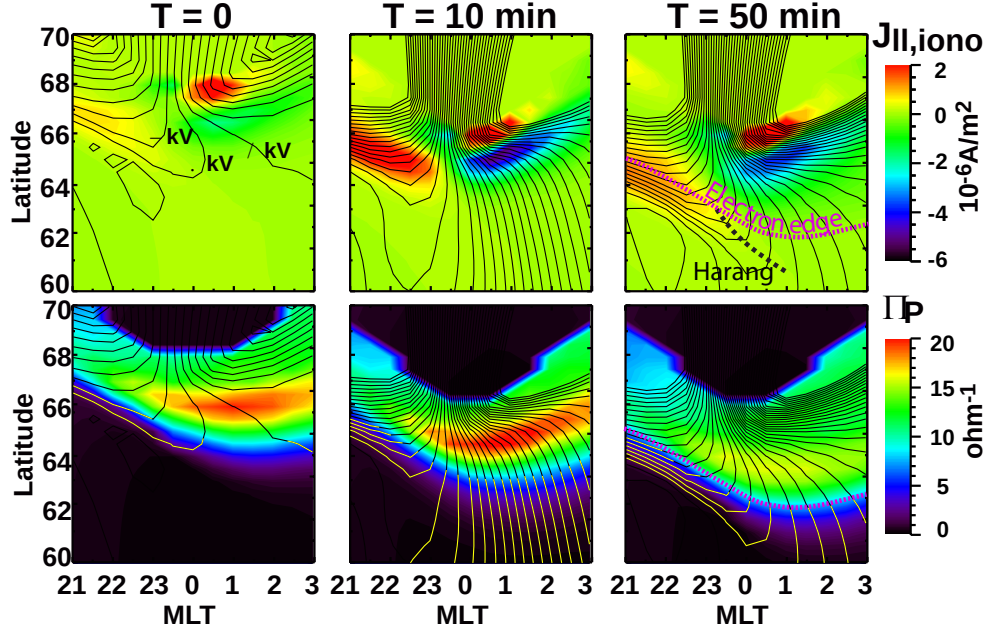


Fig 2. RCM results within the ionosphere using the Tsyganenko 96 magnetic field just prior to the convection enhancement ($T = 0$) and for later indicated times. FACs are positive downward. Height-integrated Pederson conductivity (bottom row) reflects the energy flux of plasma sheet electron precipitation. Equipotentials are overlaid in each panel. Dashed lines in the panels for $T = 50$ min identify the Harang reversal [Gkioulidou *et al.*, 2009].

boundary and have been associated with poleward boundary intensifications (PBI) along the auroral poleward boundary [e.g., *de la Beaujardière et al.*, 1994]. Furthermore, a number of studies show that these flow bursts can extend earthward within the plasma sheet, such extension being associated with auroral forms extending equatorward from PBIs that are referred to as north-south (N-S) auroras or auroral streamers. PBIs and streamers are two of the most common Space Weather disturbances and evidence suggests that they are associated with meso-scale flow enhancements.

These types of disturbances appear to originate quite close to the boundary between open and closed field lines and the radar observations [e.g., *de la Beaujardière et al.*, 1994] indicate that the flow enhancements leading to PBIs cross into the plasma sheet from the polar cap. Substorm onsets, on the other hand, have generally not been considered as directly related to flows from the polar cap boundary. However, recent ground based ASI and radar observations indicate the possibility that this long-held view may not be correct, suggesting a possible unified framework for understanding most Space Weather disturbances.

The very recent analysis [Nishimura *et al.*, 2010a] has led to the unexpected result that onset is preceded by PBI formation along the auroral poleward boundary, from which a streamer extends equatorward towards the equatorward boundary of the aurora oval. This can then lead to onset near where the streamer first reaches a diffuse-like arc located near the equatorward boundary, or after some azimuthal propagation. Auroral enhancements are expected to lie to the right of the enhanced flow direction, based on the enhanced upward FACs that form auroral enhancements being fed by enhanced ionospheric Pedersen currents within the flow channels. Evidence has been found for these flows in both radar [Lyons *et al.*, 2010] and THEMIS observations of the plasma sheet [Xing *et al.*, 2010].

The above results suggest that the plasma sheet may be viewable as having a background state, which can be described by the drift physics of the RCM, and with disturbances arising from meso-scale flow bursts. Note that the meso-scale flow bursts have been theoretically proposed to consist of depleted magnetic flux tubes (i.e., flux tubes with lower total entropy than the surroundings, leading to earthward interchange motion) [Pontius and Wolf, 1990; Chen and Wolf, 1993], and there are now some observations which support this depletion [e.g., Panov *et al.*, 2010].

We illustrate this concept for a new framework using the assemblage of THEMIS ASI images of the aurora in Figure 3 from an unusually clear, disturbed night over North America. The considerable activity

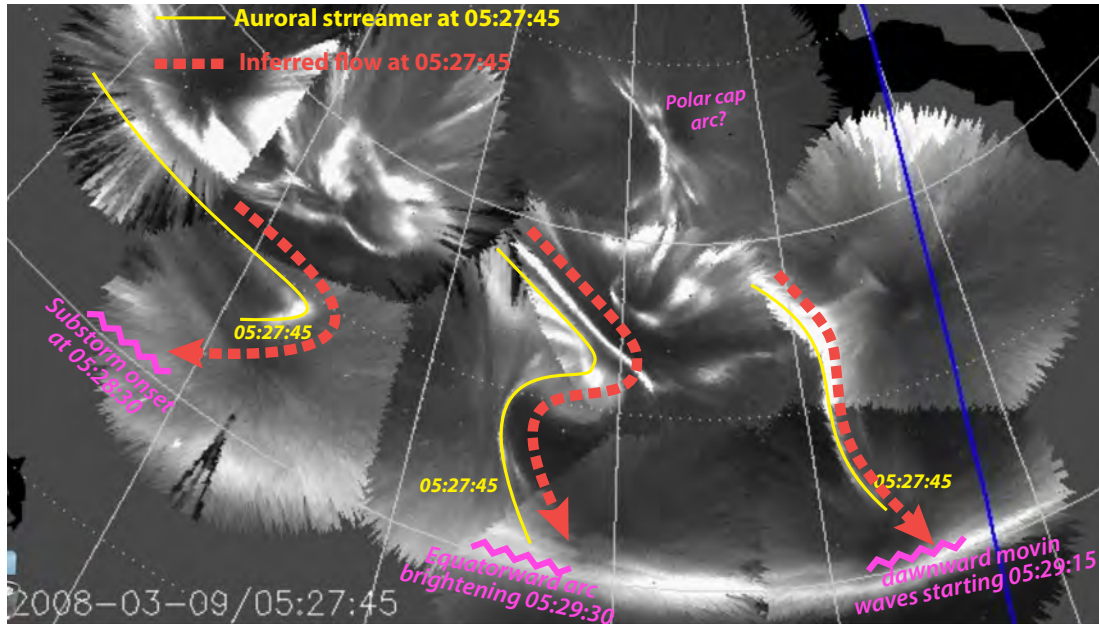


Figure 3. Auroral images from the THEMIS all-sky imager network, with north to the top. Magnetic latitude circles are drawn as solid lines at 60° and 70° , and longitude lines are drawn at 1 hr magnetic local time intervals. Three auroral streamers and associated flow channels are identified.

seen within the poleward portion of the auroral oval would be associated with reduced total entropy flow channels entering the plasma sheet. Several auroral streamers are seen simultaneously in the image. Based on the proposed framework, the streamers are like weather fronts, separating plasma masses of different properties. Here they separate higher entropy plasma from lower entropy plasma that is being brought closer to the Earth within the flow channels that extend well earthward from the outer plasma sheet boundary. Movies of the auroral images (not shown here) show these “fronts” slowly moving through the background plasma, as weather fronts do through the atmosphere. Three streamers (i.e., fronts) and inferred flow channels within the aurora oval are identified in Figure 3. Each of these flow channels led to a different type of auroral disturbance when it reached near the equatorward auroral boundary, as indicated. These new results suggest a unifying view that substorms, PBIs, streamers and other disturbances may be all related to meso-scale flow channels coming from the distant plasma sheet. If so, understanding the generation and propagation of these meso-scale flow channels and the cross-scale coupling with the large-scale background is critical for understanding Space Weather disturbances.

Coupling with thermospheric and ionospheric structure. The evolution of the above large and meso-scale structures have very recently been found to strongly affect the structure of the ionosphere. For example, by combining the large-scale distribution of TEC with meso-scale PFISR observations, the location and motion of the poleward boundary of a very low ionospheric n_e region, the mid-latitude trough, was found to be controlled by features of the evolving background pattern [Zou *et al.*, 2010]. PFISR measurements have shown regions of low n_e accompanying flow channels as might be expected for depleted flux tubes [Zou *et al.*, 2009c]. Strong n_e enhancements were observed to the right of the flow directions, indicating electron precipitation associated with auroral arcs. Meso-scale n_e enhancements have also been measured in the region of auroral arcs, which are associated with upward FACs, and a meso-scale high-latitude trough has been seen connecting to downward auroral FACs [Zou *et al.*, 2010].

While only a few cases have so far been identified and analyzed, the above observations indicate that specific features of both the large and meso-scale plasma sheet structures lead to substantial ionospheric density changes, as well as to temperature and flow changes. Such structure is expected to greatly influence the thermosphere via collisions, but specific description of this structure and models capable of evaluating structure effects have not been previously available. As an example, the atmosphere is generally considered to be in hydrostatic equilibrium, with gravity exactly balancing the pressure gradient. However, there are regions of extremely large vertical flows in the thermosphere, such as 20-100 m/s at 250 km altitude in the thermosphere [Spencer *et al.*, 1982; Crickmore *et al.*, 1991; Conde and

Dyson, 1995; Price *et al.*, 1995; Guo and McEwen, 2003], and we have little understanding of what causes these flows besides a general association with strong auroral heating and currents. Strong vertical winds have been investigated using the non-hydrostatic University of Michigan Global Ionosphere Thermosphere Model (GITM) of the coupled thermosphere-ionosphere system [Ridley *et al.*, 2006], and relatively strong vertical winds developed when there are rapid ($\lesssim 5$ min) changes in the aurora and its associated flows [Deng *et al.*, 2008]. These types of flows dramatically alter the upper atmosphere, changing the composition and the way that the auroral FACs can close through the ionosphere. Thus a self-consistent treatment of the coupled magnetosphere-ionosphere-thermosphere system appears to be necessary for proper modeling and understanding of the full system.

Possible relation to meso-scale, polar-cap structure. Potentially important for understanding the source of the meso-scale flow bursts is the observation that flow enhancements leading to PBIs can cross into the auroral oval from the high latitude, polar-cap region [de la Beaujardière *et al.*, 1994; Lyons *et al.*, 2010]. Very recently, Nishimura *et al.* [2010b] found narrow flow bursts moving from far within the nightside polar cap leading to PBIs that are followed by streamers, including those leading to near-Earth substorm onset. This association leads to the hypothesis that enhanced meso-scale flows formed within the polar cap may contribute to triggering the meso-scale earthward flows that lead to PBIs, streamers, and substorm onset.

Scientific Analysis Plan

Under the above framework concept, the background structure sets the conditions under which Space Weather disturbances arise. This large-scale structure should be evaluated using a model(s) that include energy dependent drift, such as the RCM, and the validity of the modeled structure and its temporal evolution in the magnetosphere and ionosphere should be tested for different strengths and durations of convection and IMF orientations. The testing should be done using the large array of facilities that is now becoming available, including PFISR, SuperDARN (with the mid-latitude the are now starting to become operational), low altitude satellites, and in situ THEMIS observations (see Wang *et al.* [2006], Xing *et al.* [2009]). Further testing can be accomplished using the global FAC pattern, soon to be available every 10 min from the AMPERE program. The goal here is to obtain a quite complete understanding of the large-scale background structure, including its time evolution. Locations of spacecraft are illustrated in Figure 1 and a portion of the extensive radar coverage relative to THEMIS ASIs is shown in Figure 4.

Both the radar and THEMIS observations can be used to determine the characteristics of the meso-scale flows relative to their auroral signatures in the ASI and the resulting disturbance type (e.g., PBI only, PBI followed by streamer, both followed by a substorm, and other disturbances such as the central one in Figure 2 and omega bands which have been related to streamers [Henderson *et al.*, 2002]).

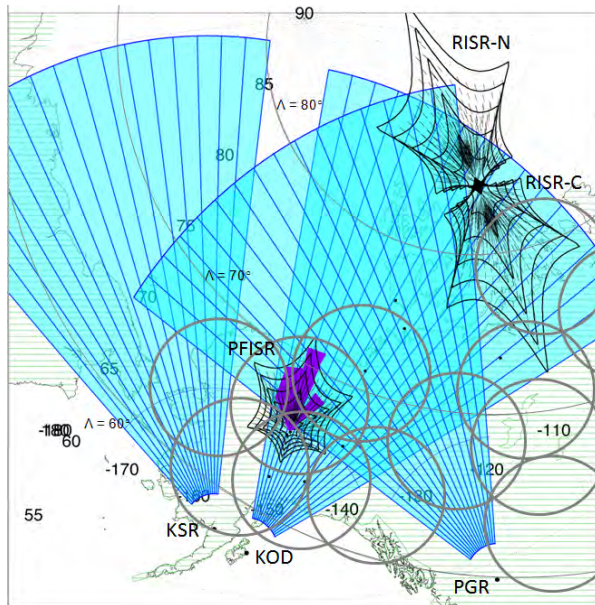


Figure 4. Field of views of AMISRs (PFISR, RISR-N and RISR-C), SuperDARN radars (blue), and the THEMIS ASIs (grey circles) over Alaska and Western Canada [based on W. Bristow, personal communication., 2010].

Evidence of depleted flux tubes can be tested via their signature in ISR observations of n_e resulting from reduced precipitation [Zou *et al.*, 2009c; Lyons *et al.*, 2010] and their signature in THEMIS spacecraft observations [Xing *et al.*, 2010]. Depleted flux tubes should also be added in models (as has been done with the RCM by Zhang *et al.*, [2008]) for comparison with the observations. An additional goal of such studies should be to learn why only some PBIs lead to streamers, why only some streamers lead to substorm onset, and why some streamers lead to other disturbances. Possibilities for consideration include features of the large-scale plasma and flow structure and of the plasma within the streamers, as well as the location of the flow channel relative to large scale flow pattern (e.g., can the flow channel be guided into the region of peak plasma pressures within the inner plasma sheet).

To determine the connections with ionospheric

structure, ionospheric n_e distributions relative to specific precipitation, electric field, and FAC features of the large and meso-scale structures should be determined. This can be accomplished using ISR observations, which provide altitude profiles of n_e in a limited region, and GPS TEC, which provides large-scale 2-D horizontal distribution of integrated n_e . Ionospheric tomography using signals from GPS and LEO beacon satellites received at receiver chains can be used to reconstruct the altitude profile of n_e [e.g. *Yizengaw and Moldwin*, 2005; *Kersley*, 2005], which can be done to quite small scales [*Moen et al.*, 1998]. Together with the 2-D horizontal GPS TEC distribution, this gives the potential to construct a 3-D n_e distribution. This will allow a combining of n_e observations, auroral images from the continental-wide THEMIS ASIs, convection flows from radars, and FACs from AMPERE, allowing for investigation of the mutual evolution of ionospheric densities and the large and meso-scale structures.

Evaluation of the effects on the thermosphere will require modeling capabilities such as those that have been incorporated in GITM, which has been run at 35 km latitudinal resolution over the entire globe, and can be run at substantially higher resolution for the auroral oval region. This is ≥ 10 times higher than the resolution of past ionosphere-thermosphere models, allowing analysis of the effects of the detailed observational features of the large and meso-scale disturbances. Coupling ionospheric-thermospheric models such as GITM with electrodynamic models such as the RCM, would allow evaluation of thermospheric responses, and also the feedback of conductivities for self-consistent modeling of the entire system. Coordination with the ionospheric density, radar, and FAC observations would also allow for model testing. Comparisons should also be made of the model results to thermospheric observations whenever possible.

Connections of meso-scale flows within the polar cap to those within the plasma sheet should also be evaluated, as well as the relations between polar cap flow structures and PBIs, streamers, and substorm onsets. This can be done with observations from the Sondrestrom ISR following de la Beaujardière et al. [1994] and *Lyons et al.* [2010], and the new SuperDARN radars that are located near the polar cap boundary (e.g., at Rankin Inlet) following *Nishimura et al.* [2010b]. The two Resolute Bay ISRs (RISR) within the polar cap will also be valuable for this. RISR-C is planned to begin operating in early 2012 and will be particularly important since it will be directed from well within the polar cap to the portion of the polar cap boundary covered by the ASIs at Rankin Inlet, Snap Lake and Taloyoak. Data that has been preliminarily examined from RISR-N for a few events shows the potential associations between polar cap flow structures and PBI, streamer, and substorm occurrence.

Summary and implementation. New facilities and model development give the exciting and unprecedented opportunity for comprehensive interdisciplinary study of the coupled MIT system. Coordination between a variety of types of observations, with modeling, and between the typically separated Magnetosphere and Ionosphere/Thermosphere communities is essential. The research approach described above has transformative potential through:

1. Showing the extent that plasma sheet magnetosphere-ionosphere dynamics can be described as a slowly varying background state, with more rapid disturbances driven by meso-scale flows, and determining the characteristics of the background and meso-scale features.
2. Using the cross-scale coupling features to characterize and understand how the dynamics on both scales affect major features of ionospheric and thermospheric structure, dynamics, and disturbances including the importance of non-hydrostatic effects and of self-consistent conductivities.
3. Determining the extent to which there is important coupling between meso-scale structure within the open polar cap region and that within the plasma sheet, and evaluating the effects of this coupling on major disturbances.

To accomplish these goals it is necessary that operations of the current and planned facilities described above continue with adequate funding for making data available to the community. This is already planned for the NSF facilities, but needs to be maintained also for the THEMIS ASIs and spacecraft. It is also necessary that groups and individuals be funded for the requisite coordinated research. On the order of 5-10 researchers should be able to make substantial progress provided coordination is maintained. This would also be an excellent opportunity for training student and post-docs so that they will be to take active roles in a future, coordinated magnetosphere-ionosphere-thermosphere research within the scientific community.

References

- Bristow, W. A., and P. Jensen (2007), A superposed epoch study of SuperDARN convection observations during substorms, *J. Geophys. Res.*, **112**, A06232, doi:10.1029/2006JA012049.
- Conde, M., and P. L. Dyson (1995), Thermospheric vertical winds above mawson, antarctica, *Journal of Atmospheric and Terrestrial Physics (ISSN 0021-9169)*, **57**, 589
- Crickmore, R. I., J. R. Dudeney, and A. S. Rodger (1991), Vertical thermospheric winds at the equatorward edge of the auroral oval, *Thermospheric and ionospheric dynamics; Symposium*, **53**, 485.
- de la Beaujardière, O., L. R. Lyons, J. M. Ruohonmiemi, E. Friis-Christensen, C. Danielsen, F. J. Rich, and P. T. Newell (1994), Quiet-time intensifications along the poleward auroral boundary near midnight, *J. Geophys. Res.*, **99**, 287.
- Deng, Y., A.D. Richmond, A.J. Ridley, and H.-L. Liu (2008), Assessment of the non-hydrostatic effect on the upper atmosphere using a general circulation model (GCM), *Geophys. Res. Lett.*, **35**, L01104, doi:10.1029/2007GL032182, 2008.
- Gkioulidou M., C.-P. Wang, L. R. Lyons, R. A. Wolf (2009), Formation of the Harang reversal and its dependence on plasma sheet conditions: Rice convection model simulations, *J. Geophys. Res.*, **114**, A07204, doi:10.1029/2008JA013955.
- Guo, W., and D. J. McEwen (2003), Vertical winds in the central polar cap, *Geophys. Res. Lett.*, **30**, doi:10.1029/2003GL017124.
- Henderson, M. G., L. Kepko, H. E. Spence, M. Connors, J. B. S. and L. A. Frank, H. J. Singer, and K. Yumoto (2002), The evolution of northsouth aligned auroral forms into auroral torch structures: The generation of omega bands and ps6 pulsations via flow bursts, in *Sixth International Conference on Substorms*, edited by R. M. Winglee, pp. 169–174, University of Washington.
- Kersley, L. (2005) Ionospheric tomography and its applications in radio science and geophysical investigations, *Ann. Geophysicae*, **48**, 535-548.
- Lyons, L. R., S. Zou, C. J. Heinselman, M. J. Nicolls, and P. C. Anderson (2009a), Poker Flat Radar Observations of the Magnetosphere-Ionosphere Coupling Electrodynamics of the Earthward Penetrating Plasma Sheet Following Convection Enhancements, *J. Atmos. Terr. Phys.*, **71**, 717-728, doi:10.1016/j.jastp.2008.09.025.
- Lyons, L. R., Y. Nishimura, Y. Shi, S. Zou, H.-J. Kim, V. Angelopoulos, C. Heinselman, M. J. Nicolls, and K.-H. Fornacon (2010), Substorm triggering by new plasma intrusion: Incoherent-Scatter Radar Observations, *J. Geophys. Res.*, doi:10.1029/2009JA015168, A7, A07223, 2010.
- Moen, J., S.T. Berry, L. Kersley and B. Lyvekk (1998) Probing discrete auroral arcs by ionospheric tomography, *Ann. Geophysicae*, **16**, 574-582.
- Nishimura, Y., L. Lyons, S. Zou, V. Angelopoulos, and S. Mende (2010a), Substorm triggering by new plasma intrusion: THEMIS all-sky imager observations, *J. Geophys. Res.*, doi:10.1029/2009JA015166, **115**, A7, A0722.
- Nishimura, Y., L. Lyons, S. Zou, X. Xing, V. Angelopoulos, S. Mende, J. Bonnell, D. Larson, U. Auster, T. Hori, N. Nishitani, K. Hosokawa, G. Sofko, M. Nicolls, and C. Heinselman (2010b), Pre-onset time sequence of auroral substorms: Coordinated observations by ASIs, satellites and radars, *J. Geophys. Res.*, in press.
- Panov, E. V., R. Nakamura, W. Baumjohann, V. Angelopoulos, A. A. Petrukovich, A. Retinò, M. Volwerk, T. Takada, K. - H. Glassmeier, J. P. McFadden, and D. Larson, Multiple overshoot and rebound of a bursty bulk flow, *Geophys. Res. Lett.*, **37**, L08103, 2010
- Pontius D. H, Jr., and R. A. Wolf, Transient flux tubes in the terrestrial magnetosphere, *Geophys. Res. Lett.*, **17**, 1, 49-52, 1990
- Price, G. D., R. W. Smith, and G. Hernandez (1995), Simultaneous measurements of large vertical winds in the upper and lower thermosphere, *J. Atmos. Terr. Phys.*, **57**, 631.
- Spencer, N. W., L. E. Wharton, G. R. Carignan, and J. C. Maurer (1982), Thermosphere zonal winds - vertical motions and temperature as measured from dynamics explorer, *Geophys. Res. Lett.*, **9**, 953, doi: 10.1029/GL009i009p00953.

- Wang, C.-P., L. R. Lyons, J. M. Weygand, T. Nagai, and R. W. McEntire (2006), Equatorial distributions of the plasma sheet ions, their electric and magnetic drifts, and magnetic fields under different interplanetary magnetic field B_z conditions, *J. Geophys. Res.*, **111**, A04215, doi:10.1029/2005JA011545.
- Wang, C., L. R. Lyons, T. Nagai, J. M. Weygand, and R. W. McEntire (2007), Sources, transport, and distributions of plasma sheet ions and electrons and dependences on interplanetary parameters under northward interplanetary magnetic field, *J. Geophys. Res.*, **112**, A10224, doi:10.1029/2007JA012522.
- Xing, X., L. R. Lyons, V. Angelopoulos, D. Larson, J. McFadden, C. Carlson, A. Runov, and U. Auster (2009), Azimuthal plasma pressure gradient in quiet time plasma sheet, *Geophys. Res. Lett.*, **36**, L14105, doi:10.1029/2009GL038881.
- Xing, X., L. Lyons, Y. Nishimura, V. Angelopoulos, D. Larson, C. Carlson, J. Bonnell, and U. Auster (2010), Substorm onset by new plasma intrusion: THEMIS spacecraft observations, *J. Geophys. Res.*, **115**(A10), A10246.
- Xing, X. L. Lyons, Y. Nishimura, V. Angelopoulos, D. Larson, C. Carlson, J. Bonnell, and U. Auster (2010), Substorm Onset by new Plasma Intrusion: THEMIS Spacecraft Observations, *J. Geophys. Res.*, in press.
- Yizengaw, E., Moldwin, M.B., 2005. The altitude extension of the mid-latitude trough and its correlation with plasmopause position. *Geophysical Research Letters* **32**, L09105.
- Zhang, J.-C., R. A. Wolf, S. Sazykin, and F. R. Toffoletto (2008), Injection of a bubble into the inner magnetosphere, *Geophys. Res. Lett.*, **35**, L02110, doi:10.1029/2007GL032048.
- Zou, S., L. R. Lyons, C.-P. Wang, A. Boudouridis, J. M. Ruohoniemi, P. C. Anderson, P. L. Dyson, and J. C. Devlin (2009a), On the coupling between the Harang reversal evolution and substorm dynamics: A synthesis of SuperDARN, DMSP, and IMAGE observations, *J. Geophys. Res.*, **114**, A01205, doi:10.1029/2008JA013449.
- Zou, S., L. R. Lyons, M. J. Nicolls, C. J. Heinselman, and S. B. Mende (2009b), Nightside ionospheric electrodynamics associated with substorms: PFISR and THEMIS ASI observations, *J. Geophys. Res.*, **114**, A12301, doi:10.1029/2009JA014259.
- Zou, S., L. R. Lyons, L. R., M. J. Nicolls, and C. J. Heinselman (2009c), PFISR observations of strong azimuthal flow bursts in the ionosphere and their relation to nightside aurora. *J. Atmos. Terr. Phys.*, doi:10.1016/j.jastp.2008.06.015.
- Zou, S., M. Moldwin, L. Lyons, Y. Nishimura, and M. Nicolls (2010), Relationship between ionospheric trough, convection and field-aligned currents during substorms: GPS-TEC and radar observations, *CEDAR workshop*, Boulder, CO, USA.

A Comprehensive and Continuous Record of Ground Based Space Weather Observations

Justin Mabie – University of Colorado , CIRES at NOAA

Executive Summary:

Magnetometers, Ionosondes and other instruments have been collecting data for many decades covering a large geographical distribution. These instruments provide a reliable and continuous baseline for monitoring the space environment. Specification of the state of the space environment using these data can only be pushed back as far as a sufficient volume of data is available and can be ingested by computational routines. This provides serious limitations on our ability to understand variability on time scales of a solar cycle and longer, and limits our ability to determine the frequency of occurrence of important phenomenon. In the interest in understanding the Earth's climate system and how it has long term consequences on the space environment, the complete set of ground based space weather observations must be modernized. The recommendations of this white paper provide an extremely high 'bang for the buck' and should be implemented before the data becomes at risk of loss.

1) Introduction

This 'white paper' is focused on the need to have a complete data record in order to understand space weather as far back in time as is possible. This effort should be focused on making sure that operation of ground based instruments remains ongoing and that the long term records are not discontinued due to issues that are unrelated to their scientific value. Secondly, the available data record for these observations has not been modernized in a way that is readily available for computational techniques. Modernization of this data should be taken in 2 steps. First, all data dating back to the International Geophysical Year should be modernized in order to provide the most comprehensive data record we have dating back more than five decades. Next, the remaining data should be modernized back as far as possible for longer term studies.

2) Digitizing Magnetograms and Ionograms

Magnetometers have been collecting data since the mid 19th century and ionosondes have been in continuous operation since the early 20th century. These data provide an extensive observation record of the space environment. Digitization of these data carries a very high return for relatively little cost. This effort should be conducted in a concerted, well defined, and comprehensive manner.

Because this data is on outdated media, the first step is to transfer the data records onto a medium that can be read by a computer. The Ionosonde and Magnetometer records that date back to IGY are almost entirely available on 35mm film or microfiche. Prior to IGY, some of the data can be found on film, but much of it is either in the form of paper records (including magnetograms, ionograms and hand written values) or in publications issued by the original observer. The format of these various pre-IGY records are more diverse than the post-IGY film records and will require a separate effort from the film record modernization.

Step 1 – Getting the data on a consistent and readable media

The NOAA Climate Data Modernization Project (CDMP) is an example of a small scale effort to migrate these data into a computer readable format. Currently, Ionosonde and Magnetometer data held on film are converted into digital image files. This process takes the data out of the warehouse and makes it available to the scientific community. The funding rate of this project is not sufficient to accomplish the main goal of obtaining a comprehensive data set, and the data are not convenient to use for specification of the space environment without further digitization. A more concerted effort is needed to fully implement this process.

Step 2 – Defining the traces with artificial intelligence software

The traces on the images must be extracted from the data. Once the data is in a digital image format, recognition techniques can be applied. Currently the technology has been developed for comprehensive recovery of magnetograms and seismograms. This technology can be extended to extract trace values from ionosondes as well. The finished product once these artificial intelligence procedures are applied is a digital record of the actual measurements made at the instrument.

Step 3 – Quality Control of the data

Once the measurements as taken by the instruments have been recovered digitally, traditional quality control techniques should be applied. All digitized data should immediately be subjected to standard QC methods. Once quality control is complete, a complete data record ready for ingest will be available to our community.

Other Data Sets

In addition to Magnetometer and Ionosonde data, other ground based space weather observations that have been widely dismissed and in many cases orphaned entirely should be modernized.

Other space weather datasets that are currently waiting for attention include sunspot drawings, all-sky images of aurora and airglow, solar radio burst data, solar H-alpha imagery, solar magnetograms, solar flare observations, solar brightness and magnetic discontinuities, airborne magnetic surveys, magnetic repeat data, and geomagnetic variometer data.

Conclusion

Technology has developed sufficiently to bring the cost of modernizing historical data to a small fraction of what it cost to collect and maintain the data over the years. As a community we can not

credibly claim to have an understanding of the long term processes that occur in our space environment without having access to the full suite of data available to us. By the end of this decade, we must have completely modernized our space weather environmental record so we can have the most extensive and reliable coverage available with a consistent and accurate baseline.



Figure 1: Magnetograms on 35mm film stored away in a cabinet in a dark corner of a warehouse

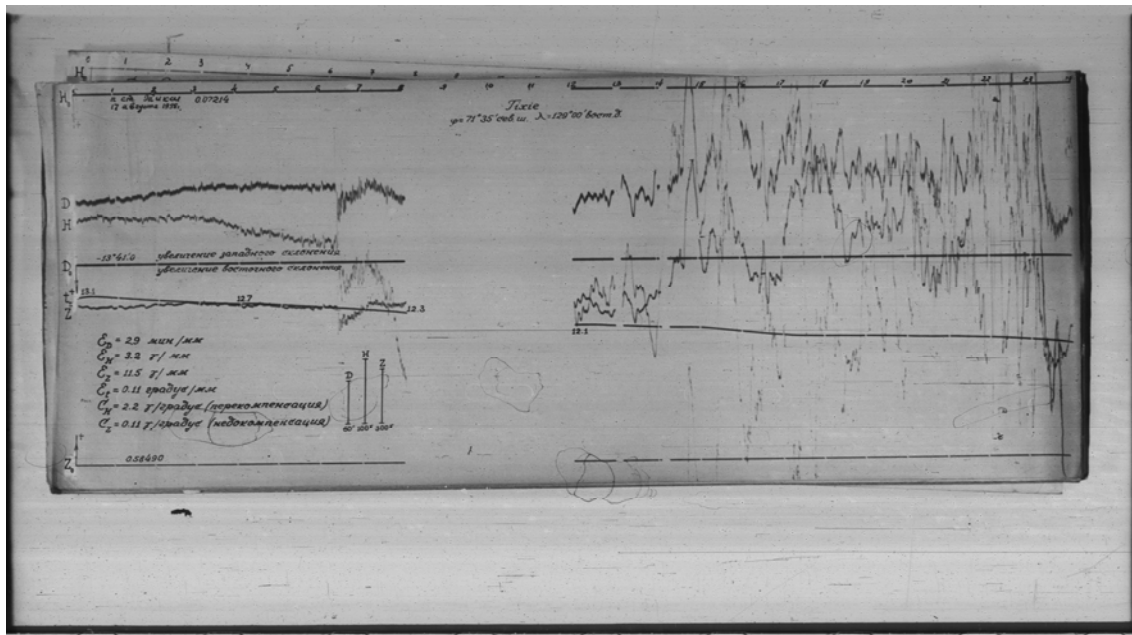


Figure 2: A digital image of the original magnetogram collected at Tixie Bay observatory in August 1958. This magnetogram can now be accessed by the community, but is not easily interpreted by data assimilation routines.

Format	IAGA-2002						
Source of Data							
Station Name	BOULDER						
IAGA CODE	BOU						
Geodetic Latitude	40.133						
Geodetic Longitude	254.767						
Elevation	1650						
Reported	DHZF						
Sensor Orientation							
Digital Sampling	1 minute						
Data Interval Type	Filtered 1-minute (00:30 - 01:29)						
Data Type	definitive						
DATE	TIME	DOY	BOUD	BOUH	BOUZ	BOUF	
1978-05-08	00:00:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:01:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:02:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:03:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:04:00.000	128	10748.00	21375.00	51318.00	99999.00	
1978-05-08	00:05:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:06:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:07:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:08:00.000	128	10748.00	21375.00	51318.00	99999.00	
1978-05-08	00:09:00.000	128	10748.00	21373.00	51318.00	99999.00	
1978-05-08	00:10:00.000	128	10748.00	21373.00	51318.00	99999.00	
1978-05-08	00:11:00.000	128	10748.00	21373.00	51318.00	99999.00	
1978-05-08	00:12:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:13:00.000	128	10748.00	21374.00	51318.00	99999.00	
1978-05-08	00:14:00.000	128	10748.00	21374.00	51318.00	99999.00	

Figure 3: Quality controlled magnetometer data from Boulder Observatory May 8, 1978 in IAGA02 format. This data is ready for assimilation!

A science mission concept to actively probe magnetosphere-ionosphere coupling

Elizabeth MacDonald, Joe Borovsky, Brian Larsen, Eric Dors
Space Science and Applications, Los Alamos National Laboratory
Unlimited Release: LA-UR 10-07644

Description: This mission concept describes how directly mapping magnetic field lines from a magnetospheric satellite to their ionospheric footpoints using an on-board electron emitter and ground imaging techniques can answer long-standing fundamental questions of magnetosphere-ionosphere coupling

Introduction

This white paper is submitted to the *Panel on Atmosphere-Ionosphere-Magnetosphere Interactions (AIM)* for the Heliophysics Decadal Survey. We summarize a mission concept aimed at unambiguously relating ionospheric signatures to the causative magnetospheric conditions and magnetospheric regions and boundaries to their ionospheric signatures. This has been an open question since in situ observations first began to reveal the physical processes more than 50 years ago. Active experiments have shown promise throughout the years for resolving field line mapping ambiguities. In order to truly understand magnetosphere – ionosphere coupling we need to go beyond present limited methodologies that generally relate low altitude phenomena to high altitude phenomena through similarity or empirical models. The format of this paper is to first expand upon the scientific problem and then outline a potential feasible solution.

Next year, scientists will convene in Fairbanks, Alaska, for an AGU Chapman Conference on the Relationship Between Auroral Phenomenology and Magnetospheric Processes. From the conference description: “Some of the most interesting and pressing questions in space physics relate to the origin of auroral structure and dynamics, and the nature of acceleration and pitch angle scattering in the case of discrete and diffuse aurora, respectively. ... The conference will provide an opportunity to present the latest results from analyses of experimental data (including space-borne, ground-based and co-ordinated data), simulation and theory, addressing various aspects of the aurora, in order to connect our knowledge of auroral morphology and mechanisms to candidate physical processes in the magnetosphere capable of powering and structuring the aurora on Earth and other planets.” The topics to be addressed by this conference include the relationship between auroral and magnetospheric dynamics, fields, and gradients. In other words, one goal is to try to determine what it is in the magnetosphere that causes aurorae. Recently, the THEMIS mission has addressed global substorm processes through coordinated multi-point ground and in situ observations in a two dimensional plane [Angelopoulos, 2008]. However, lack of knowledge of field line mapping in the transition region between dipolar and highly stretched field lines leads to major unknowns in understanding the linkages between auroral ionospheric and magnetospheric processes.

The fact that these questions are still under active research and debate is testimony to the major challenge of trying to relate phenomena observed in the Earth’s ionosphere

with phenomena occurring tens of thousands of km away in space. The fundamental difficulty in attributing magnetospheric causality to auroral phenomena is (and always has been) the lack of definitive knowledge of the mapping between one and the other. In magnetospheric and ionospheric physics, the magnetic field is the conduit for the transmission of particle and electrodynamic communication. Therefore, while we have many ground and space observations of aurora and many in situ observations of magnetospheric properties and dynamics, relating them to each other requires knowing the magnetic connectivity between them. The challenge here is that the global magnetic field is strongly affected by currents flowing within the magnetosphere itself and is highly dynamic. This is particularly true in the region of space that is the probable source of most auroral phenomena.

Over the years considerable progress has been made in understanding the global structure of the magnetospheric magnetic field and its variability. Global MHD simulations [e.g., Raeder et al., 2001] demonstrate the strong distortions from the underlying dipolar geometry and how those distortions change based on the properties of the upstream solar wind. Since the early '80s, a class of empirical global field models derived from analysis of large quantities of in situ magnetic field measurements [e.g., Tsyganenko and Usmanov, 1982; Tsyganenko, 1989; Tsyganenko and Sitnov, 2007] has provided a quantitative description of the average global field topology under various conditions. These empirical models, parameterized by various relevant physical quantities such as solar-wind properties, show dramatic variations in the mappings from a given point in the magnetosphere to the ionosphere.

The empirical models have for the last decade served as the primary tool for relating ground- and space-based observations. However, efforts to compare the modeled field with various observations have revealed not only systematic differences [e.g., Fairfield, 1991; Donovan et al., 1992; Jordan et al., 1992; Peredo et al., 1993; Pulkkinen and Tsyganenko, 1996; Thomsen et al., 1996], but strikingly wide variability in magnetic field measurements for an individual event. A glance at the studies mentioned above demonstrates that the instantaneous variation from the model (or the statistical average) field vector is typically a substantial fraction of the average. Thus, as noted in the 2003 Decadal Survey for solar and space physics [NRC, 2003], "the actual instantaneous configuration of the magnetosphere is not expected ever to resemble any of its average states." Indeed, at any given time the magnetic field measured in situ can deviate significantly from that predicted by these various models. Thus, the mapping from ground to space may also be significantly different from the model expectations. For instance, THEMIS investigators report uncertainties in mapping auroral arcs at onset to be larger than 10 Earth radii [Substorm Expansion Focus Group Report, GEM conference, 2010].

These difficulties with using global empirical models to map between the magnetosphere and the auroral ionosphere have led numerous researchers to develop a range of techniques to try to determine the connectivity observationally. Using precipitating charged particle measurements made at low altitudes by polar-orbiting DMSP satellites, and noting the spectral similarities with particle distributions observed previously in various high-altitude regions of the magnetosphere, Newell and Meng and colleagues have mapped extensively the statistical ionospheric occurrence pattern of different regions of the magnetosphere [e.g., Newell and Meng, 1988, 1992, 1994].

Other authors have used low-altitude particle signatures to identify the magnetospheric source location for ionospheric phenomena such as traveling convection vortices (TCVs) [Moretto and Yahnin, 1998] and discrete aurorae [Yahnin et al., 1997] and to monitor the tail current sheet structure [Sergeev, 1995; Sergeev and Gvozdevsky, 1995; Sergeev and Kubyshkina, 1996]. Still other studies [e.g., Weiss et al., 1996] have compared low-altitude and high-altitude electron distributions during nominal conjunctions to identify times of best match, presumably when the satellites were on roughly the same field line. Recently, using THEMIS observations, Nishimura and collaborators have shown that individual chorus elements can be mapped to specific patches of pulsating aurora by comparing wave intensity to auroral luminosity [Nishimura et al., 2010]. The precision of this technique is very promising but its application to one specific phenomenology may limit its utility to global mapping issues.

Such studies have helped greatly to illuminate the regional connectivity between the ionosphere and the magnetosphere and have provided the crucial link enabling the global association between certain ionospheric conditions and the conjugate magnetospheric region. However, such regional mappings are generally unable to specify the location and conditions within a given magnetospheric region that is magnetically connected to the ionospheric footprint. Even in the case of boundary crossings, where the mapping is well localized in one dimension (normal to the boundary), there is great uncertainty about the exact mapping along the boundary. Hence, it is only by rough inference that specific magnetospheric features can be related to specific ionospheric features through their proximity to various regional boundaries.

The ability to definitively identify the magnetic field mapping from over large distances stands in the way of a full, unambiguous understanding of which conditions in the magnetosphere give rise to various ionospheric signatures. One remedy is to use charged particles as field-line tracers, emitting them in one region and detecting them in the other. Combined with appropriate local diagnostics of ionospheric signatures and magnetospheric conditions, such a capability would enable us to definitively connect the physics and the signatures, which is the real prerequisite for magnetosphere-ionosphere coupling studies. Insight gained from this definitive mapping has the potential to “enlighten a space age of measurements” enabling past and future measurements to be placed into the correct context between the ionosphere and the magnetosphere. The mapping allows determination of who drives who, and a vastly improved understanding of the linking between these two regions.

Mission Concept

One feasible technique for determining the magnetic connectivity would be to fly an electron beam generator on a magnetospheric satellite, which would periodically emit an intense beam of electrons parallel to the magnetic field. Such electrons would follow the field down to the ionosphere, where they could be detected via their optical or radar signature. Combining local magnetospheric diagnostics on and near the spacecraft with auroral and convective signatures at the ionospheric location of the beam spot would produce an unambiguous linkage between the ionospheric signatures and the causative conditions in the magnetospheric source region.

This mission is envisioned to include an active electron injector in space, with a plasma contactor to neutralize the payload and permit ambient measurements immediately before and after the injected electron beam pulsing. In situ observations of plasma and electric and magnetic waves are considered critical. Also critical are electron drift measurements and those of ion composition; both missing on THEMIS. The orbit can be highly elliptical with apogee chosen to allow optimal long conjunctions over the ground detection sites. Though this mapping technique is broadly applicable for numerous regions, the highest priority orbit would explore the dynamic under-sampled transition region field lines in the nightside magnetotail approximately 1-3 Earth radii beyond geosynchronous orbit.

The highest risk items for this mission are the present paucity of space flight qualified electron injectors and the uncertainty of seeing the electron beam spot. The former risk can be addressed head-on; the best strategy is to merge the expertise of groups building compact electron accelerators for terrestrial vacuum applications with those of high power space engineering applications. The latter risk can be reduced through modeling efforts, injector optimization, and redundant beam detection efforts. Both optical and radar based detection can resolve a beam of reasonable power with a sufficient signal to noise ratio. Recent technological advances in numerous relevant areas (camera sensitivity and noise, radar arrays, plasma contactors, and electron injectors) enable a highly feasible mission with lower resources.

Relatively modest spacecraft and instrumentation with high heritage make this a relatively low-cost mission, perhaps even within the Explorer program budget cap. If larger programs and launch vehicles are available, other mission configurations (e.g. multiple vehicle missions) may become viable. A complete cost benefit analysis of the additional science to be gained by multiple in situ spacecraft has not been performed.

This technique can be applied to key questions in other areas of the magnetosphere. Numerous other complementary efforts are identified, e.g. ground-based and suborbital campaigns. The THEMIS mission has placed greater urgency upon these questions. Active experiments have a rich history of important contributions to the field of space physics and are the only direct ways to resolve questions of field line mapping. A series of electron beam rocket experiments led by Winckler demonstrated the feasibility of emitting and imaging an electron beam in space to investigate primarily ionospheric physical mechanisms [Winckler, 1992]. As the spiral of knowledge advances, revisiting active experiments holds a key to finally closing fundamental coupling questions.

References

- Angelopoulos, V., The THEMIS Mission, *Space Science Rev.*, 141, 5, 2008.
- Donovan, E. F., G. Rostoker, and C. Y. Huang, Regions of negative Bz in the Tsyganenko 1989 model neutral sheet, *J. Geophys. Res.*, 97, 8697, 1992.
- Fairfield, D. H., N. A. Tsyganenko, A. V. Usmanov, and M. V. Malkov, A large magnetosphere magnetic field database, *J. Geophys. Res.*, 99, 11319, 1994.
- Jordan, C. E., J. N. Bass, M. S. Gussenhoven, H. J. Singer, and R. V. Hilmer, Comparison of magnetospheric magnetic-field models with CRRES observations during the August 26, 1990, storm, *J. Geophys. Res.*, 97, 16907, 1992.
- Moretto, T., and A. Yahnin, Mapping travelling convection vortex events with respect to

- energetic particle boundaries, *Ann. Geophys.*, 16, 891, 1998.
- Newell, P. T., and C. I. Meng, The cusp and the cleft boundary-layer: Low-altitude identification and statistical local time variation, *J. Geophys. Res.*, 93, 14549, 1988.
- Newell, P. T., and C. I. Meng, Mapping the dayside ionosphere to the magnetosphere according to particle-precipitation characteristics, *Geophys. Res. Lett.*, 19, 609, 1992.
- Newell, P. T., and C. I. Meng, Ionospheric projections of magnetospheric regions under low and high solar-wind pressure conditions, *J. Geophys. Res.*, 99, 273, 1994.
- Nishimura, Y., et al., Identifying the Driver of Pulsating Aurora, *Science*, 330, 81, 2010.
- Peredo, M., D. P. Stern, and N. A. Tsyganenko, Are existing magnetospheric models excessively stretched?, *J. Geophys. Res.*, 98, 15343, 1993.
- Pulkkinen, T. I., and N. A. Tsyganenko, Testing the accuracy of magnetospheric model field line mapping, *J. Geophys. Res.*, 101, 27431, 1996.
- Raeder, J., et al., Global simulation of the Geospace Environment Modeling substorm challenge event, *J. Geophys. Res.*, 106, 381, 2001.
- Sergeev, V. A., Energetic particles as tracers of magnetospheric configuration, *Adv. Space Res.*, 18, 161, 1995.
- Sergeev, V. A., and B. B. Gvozdevsky, MT-Index: A possible new index to characterize the magnetic configuration of magnetotail, *Ann. Geophys.*, 13, 1093, 1995.
- Sergeev, V. A., and M. V. Kubyshkina, Low altitude image of particle acceleration and magnetospheric reconfiguration at substorm onset, *J. Geomag. and Geoelec.*, 48, 877, 1996.
- Thomsen, M. F., D. J. McComas, G. D. Reeves, and L. A. Weiss, An observational test of the Tsyganenko (T89a) model of the magnetospheric field, *J. Geophys. Res.*, 101, 24827, 1996.
- Tsyganenko, N. A., A magnetospheric magnetic field model with a warped tail current sheet, *Planet. Space Sci.*, 37, 5, 1989.
- Tsyganenko, N. A., and M. I. Sitnov, Magnetospheric configurations from a high-resolution data-based magnetic field model, *J. Geophys. Res.*, 112, A06225, 2007.
- Tsyganenko, N. A., and A. V. Usmanov, Determination of the magnetospheric current system parameters and development of experimental field models based on data from IMP and HEOS satellites, *Planet. Space Sci.*, 30, 985, 1982.
- Weiss, L. A., M. F. Thomsen, G. D. Reeves, and D. J. McComas, *J. Geophys. Res.*, submitted, 1996., An Examination of the Tsyganenko (T89A) Field Model Using a Database of Two-Satellite Magnetic Conjunctions, *J. Geophys. Res.*, 102, 4911, 1997.
- Winckler, J. R., Controlled experiments in the earth's magnetosphere with artificial electron beams, *Rev. Mod. Phys.*, 64, 3, 1992.
- Yahnin, A. G., V. A. Sergeev, B. B. Gvozdevsky, and S. Vennerstrom, Magnetospheric source region of discrete auroras inferred from their relationship with isotropy boundaries of energetic particles, *Ann. Geophys.*, 15, 943, 1997.

A North American Thermosphere Ionosphere Observation Network

Jonathan J. Makela

Associate Professor, University of Illinois at Urbana-Champaign

Aaron Ridley

Associate Professor, University of Michigan

John W. Meriwether

Professor, Clemson University

Andrew J. Gerrard

Associate Professor, New Jersey Institute of Technology

Abstract

The Earth's ionosphere/thermosphere/mesosphere (ITM) region is driven by inputs from above and below, as well as from the polar and equatorial regions, on a wide variety of scales. Changes in the ITM region, characterized as space weather, can adversely affect our nation's increasingly critical space-based infrastructure. Successful forecasting of these space weather effects will not be feasible until an improved understanding of this complex system has been gained. Parameters of especial import that are not well characterized include measurements of neutral winds and temperatures, as well as ionospheric currents. Historically, individual ground-based instruments to make measurements of these parameters have been deployed at isolated locations without any coordination. Within the past 5-10 years, a recognition of the need for a more systematic approach to ground-based ITM observations has motivated researchers to tie instrument chains together to provide a more global coverage of the ITM region, as called for in the 2003 NRC Decadal Study Report through the idea of the Distributed Array of Small Instruments (DASI) concept. Here, we argue that a network called NATION (North American Thermosphere Ionosphere Observation Network) comprising of key optical and radio science instruments, but focused on instruments capable of measuring key ITM parameters, would enable groundbreaking research by the entire upper atmospheric community. It is important to note that all avenues of research that will be enabled by NATION cannot be foreseen at the outset. Just as the original intent of the global GPS networks so commonly used in today's ionospheric studies was observing surface dynamics, we are confident that the development of NATION will lead to new research areas not envisioned at this time. We describe a subset of instruments that provide a complementary set of currently undersampled measurements of crucial ITM parameters. We also describe a deployment scenario that would balance the desire of maximizing coverage of the ITM over North America and having sites that are reasonably easy to access and maintain.

.

Introduction

The thermosphere is a crucial boundary between the Earth's atmosphere and space. Orbiting satellites and debris are strongly affected by the lifting of the atmosphere that results from thermospheric weather [e.g., *Bruinsma et al.*, 2006; *Fuller-Rowell et al.*, 1994], which can cause unexpected perturbations in their orbits, leading to potential collisions. For example, on March 12, 2009, astronauts on the International Space Station had to take shelter because a piece of orbiting debris passed too close. On February 10, 2009, an Iridium satellite was destroyed in a collision with a defunct Russian satellite. These types of events can be mitigated with better specification of upper atmospheric conditions, leading to more accurate orbit determination. In addition, space weather due to the development of ionospheric irregularities can cause disruptions of engineering capabilities we depend on, such as GPS and radio transmissions [e.g., *Lambour et al.*, 2003; *Jadav et al.*, 2005; *Basu et al.*, 2001].

Despite the importance of this region of the atmosphere in our technology-based society, we still lack a basic understanding of the climatology and dynamics of many of the parameters, most importantly neutral winds, critical to being able to predict the state of the ionosphere/thermosphere/mesosphere (ITM) system. The most robust climatological neutral wind model, the recently revised Horizontal Wind Model (HWM07; *Drob et al.*, 2008), while improving over the previous HWM93 (*Hedin*, 1996), still lacks a large enough dataset in the thermosphere to accurately model the thermospheric wind pattern at low- and mid-latitudes. For understanding the space weather effects of mid-latitude irregularities, one needs to have spatial coverage of electron density, as well as knowledge of parameters that can cause irregularity development, most notably neutral winds, electric fields, and temperatures during periods, by definition, not indicative of what is provided by climatological models.

The lack of measurements of wind, temperature, and other parameters with sufficient spatial coverage has hampered progress in improving our understanding of the ITM. Over the past 35 years, several satellite missions have provided wind measurements (e.g., AE, DE, UARS, TIMED). However, satellite measurements alone cannot resolve the spatial-temporal scales of many interesting phenomena and ITM-focused satellite programs are, unfortunately, few and far between. As a result, the characterization of the thermospheric state is greatly enhanced through the use of ground-based remote sensing devices.

In past years, ground-based Fabry-Perot interferometer (FPI) observatories have been deployed at single sites, providing highly localized measurements of neutral temperature and winds. Similarly, imaging systems and GPS equipment have been deployed at isolated stations. These single-site deployments have garnered a wealth of information on the myriad of phenomena occurring in the ITM at mid-latitudes as well as data on ITM variability. However, these observations also make it clear that the scale of these phenomena is larger than what can be studied from a single location. *Meriwether* [2006] described the space weather community's need for thermospheric wind measurements on a denser scale and advocated for a FPI network that would help satisfy this need. Indeed, the National Academy DASI report highlighted this point as a primary justification for networks of small geophysical sensors [*Foster et al.*, 2006]. Unfortunately, there are few concrete examples of the DASI concept being implemented, due to difficulties in funding the development, deployment and operation of such ambitious networks.

However, the few examples of sufficiently dense networks of an individual type of instrument making coordinated measurements shows the strength of the DASI concept. SuperDARN has utilized many radars to measure the ionospheric flow over an extremely large area, providing maps of ionospheric convection at high latitudes. The GPS network in the US, Japan, and over the whole globe has provided valuable information on how the ionosphere behaves during storms and, in some sense, has become one of the most important data sources for studying spatial-temporal dynamics of the global ionospheric system. Most recently, the THEMIS array of ground-based magnetometers and white-light all-sky imagers (which cover Canada) have changed the way auroral substorm research is being conducted through its coordinated observations.

Here we advocate for the creation of a North American Thermosphere Ionosphere Observation Network (NATION) of ground-based remote sensing devices that would cover the North American continent, with a potential deployment as depicted in Figure 1, and provide the ability for real-time specification of spatial-temporal variability of important ITM parameters, including neutral winds, temperatures, currents, and irregularities. The specific instrumentation types that we advocate provide a relatively complete set of measurements

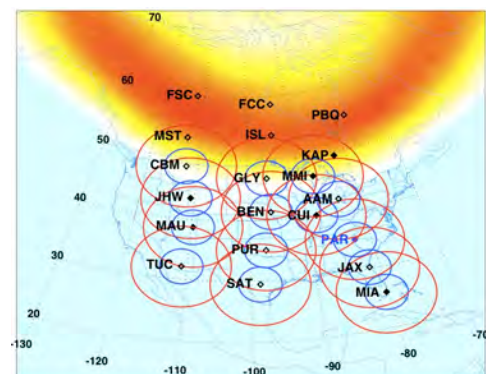


Figure 1: Potential locations of initial nodes in an ITM-focused DASI network.

needed to quantitatively describe the ITM system dynamics when coupled with existing infrastructure, such as the SuperDARN radars and the CORS GPS network. The instruments are a Fabry-Perot interferometer to measure the neutral winds and temperatures, a wide-angle imaging system to measure the large-scale structure of the ITM, a dual frequency GPS receiver to measure line-of-sight total electron content and irregularity characteristics, and a flux-gate magnetometer to measure the geomagnetic activity and currents. The NATION science goals would also be enhanced by the measurements provided by the broad azimuthal coverage of plasma density profiles, ion and electron temperature fields, as well as line-of-sight ion drifts that the Millstone Hill radar can provide when operating in a mode that covers the entire region of the northeastern continental USA and eastern Canada. This enhanced capability would be especially applicable during the upcoming solar maximum period of 2012-2013.

Enabled Research Activities

While there are many different research areas that could be addressed by such a network in instrumentation, we focus on three specific examples below.

Traveling Atmospheric Disturbances

Strong energy inputs at high latitudes and on the dayside can cause massive Joule, particle, or photon-induced heating. This heating can launch sound waves that propagate vertically and gravity waves that propagate horizontally [e.g., Rees *et al.*, 1984; Lu *et al.*, 1995; Deng *et al.*, 2008]. The horizontally propagating wave structure is typically called a traveling atmospheric disturbance (TAD). Both high-latitude heating events and large solar flares can cause TADs. Modeling results have shown that disturbances in the polar region propagate roughly equatorward with the sound speed plus the local wind speed (which can either be in the same or in an opposite direction to the wave velocity).

TADs are important because they serve to dissipate the energy infused into a concentrated region of the upper atmosphere. The efficiency of this diffusion, as contrasted with processes such as radiative cooling and conduction, is not well understood. The balance between these processes determines how much the global atmosphere will lift, affecting satellite drag, chemical composition, ionospheric densities, etc. Further, it is unknown how the waves interact with each other and with the background atmosphere and ionosphere, especially near the equator. Modeling results clearly show the TADs propagating from pole to pole and back again, which is in clear contradiction to the data [e.g., Sutton *et al.*, 2006]. Having more observational evidence on the temperatures and winds (instead of just mass density) will greatly assist modelers in determining why the models do not predict the correct structure.

NATION will allow researchers to examine the wave structure in the temperature and winds as the TAD propagates from the auroral zone to low latitudes (or towards midnight in the case of a solar flare-generated TAD) and will lead to a better characterization of the physics responsible for TADs. In theory, one could apply a single latitudinal chain of instruments to study TADs, but these results would be in many ways similar in approach to that of satellite passes, i.e., one single cut through a 3D structure. With multiple chains, researchers can study the propagation at different local times and longitudes, determining whether the propagation is completely north/south, or if there is a longitudinal component to the wave structure. Also, to study flare-generated TADs, the chain needs to be longitudinal, since the suspected propagation is from the dayside towards midnight.

Meso-Scale Traveling Ionospheric Disturbances

Nighttime Meso-Scale Traveling Ionospheric Disturbances (MSTIDs) in the nighttime mid-latitude ionosphere are typically aligned from the northwest to southeast in the northern hemisphere, propagate to the southwest at a velocity of $\sim 100 \text{ ms}^{-1}$, and have a wavelength of 100-300 km. Observations, notably from multi-imaging networks, have tracked MSTIDs from higher latitudes (55°; Ogawa *et al.* [2009]) down to the bright equatorial anomalies. More recent work in the Pacific and S. American sectors has

tracked the structures all the way to the magnetic equator [Makela *et al.*, 2010] where it appears to be a seed of post-midnight equatorial spread-F during geomagnetically quiet conditions [Miller *et al.*, 2009].

While the electrical properties of MSTIDs are well understood, thanks to extensive measurements of the electric fields obtained from satellite and radar measurements, the physics responsible for their development is not as clear. The Perkins instability [Perkins, 1973] is an oft-cited mechanism, but the growth rate is dreadfully low. Seeding of the Perkins instability by gravity waves is an enticing possibility as well, and this has been often invoked as a way to enhance the low growth rate [e.g., Miller *et al.*, 1997]. However, recent evidence from the Japanese sector suggests that the waves are actually generated in the auroral region [Ogawa *et al.*, 2009] and propagate from that region to mid-latitudes. It is unclear, then, why the waves would be seen to propagate at their characteristic angle (NE to SW), rather than radially from the auroral region, as is the case for TADs. One hypothesis is that waves that survive into the mid-latitudes are only those that survive Joule damping, which is weak in the Perkins' direction due to the polarization electric field perpendicular to the wave's $\mathbf{k}$ vector (Seyler and Kelley, personal communication).

This suggests that an array of instruments spanning from high-latitudes throughout North America would observe the preponderance of wave alignments using optical and GPS techniques. Only waves with the Perkins' alignment should be observed in the south, while a mixture of wave vectors would be expected at the higher-latitude stations. In addition, simultaneous measurements of the background neutral winds would test the hypothesis that the polarization electric fields are caused by winds blowing across the gradients in Pedersen conductivity in the MSTID bands. The necessary measurements for this test would be accomplished with NATION and would be an exciting collaboration with the new mid-latitude SuperDARN array currently being deployed that could provide the background electrodynamic conditions.

Geomagnetic Disturbances at Mid-Latitudes

During active times, and especially storm times, the injection of large amounts of plasma into the inner magnetosphere distorts the geomagnetic field. When the plasma is first injected, the fresh ion population tends to drift from midnight towards dusk and further around to noon, where some of it is lost out of the dayside magnetosphere. The result of this westward ion drift is the formation of a partial ring current. Because the horizontal (westward) current starts around midnight and ends around noon, closure currents exist out of the ionosphere near midnight and into the ionosphere near noon. In the ionosphere, these sub-auroral field-aligned currents cause strong electric fields and ion flows to develop both on the dayside and nightside, creating sometimes massive changes in the ion density structure at mid-latitudes [e.g., Lambour *et al.*, 2003; Basu *et al.*, 2001]

Early satellite measurements of these electric fields and plasma flows revealed the striking existence of narrow bands of strong westward drift (poleward $\mathbf{E}$) that is called the subauroral polarization stream (SAPS). This plasma drift channel extends across the nightside from dusk to early morning for all Kp greater than 4, with a typical peak amplitude of $> 900 \text{ ms}^{-1}$ in the premidnight sector [Foster and Vo, 2002]. The SAPS region is typically 5° in latitude. Embedded within the SAPS region are narrow ($\sim 1^\circ$) subauroral ion drifts (SAID), which are local enhancements of the poleward electric field/westward plasma flow within the broader SAPS channel.

The NATION instrument cluster will provide a wealth of information on how the mid-latitude thermosphere responds to these strong changes in the ion flow at mid-latitudes. While the flows typically exist in regions of low conductivity, they exist for long periods of time and are quite large. This means that the neutral winds may be substantially altered by these electric fields. Further, with the massive heating that exists during storm times, there will be an interesting interplay between the equatorward winds driven by this heating and the westward winds driven by the ion flows. With the overlap of the

SuperDARN radars, along with TEC and magnetometer measurements, the forcing of the mid-latitude nightside winds can be studied in a way that has never been done before.

Other Scientific Areas

The three scientific questions detailed above do not exhaust the list of science topics that would be addressed by the coordinated data set provided by NATION and its interplay with already-existing networks of instruments. Other important issues that would be addressed by NATION include the midnight temperature maximum, mid-latitude thermosphere-ionosphere coupling to the mesosphere, neutral wind control of ion dynamics, advection of chemical constituents in the ITM, and gravity wave and tidal wave affects upon the thermosphere. Finally, recognizing the benefits of instrument clusters, the deployment of NATION would develop the infrastructure on which other instruments could be deployed as the need and interest arise in the community.

Description of Potential Instrumentation

As described above, in order to increase our ability to predict the occurrence of mid-latitude space weather, as well as come to an increased understanding of the basic fundamental processes in the ITM requires knowledge of the spatial-temporal distribution of electron density, electric fields, and neutral winds and temperatures. Much information on the electron densities can be obtained from the CORS GPS network, although higher spatial resolution is required for detailed irregularity studies. The electric fields can be inferred from the developing mid-latitude SuperDARN network. This leaves the need for development of a similar observing network for the neutral temperatures and winds, smaller-scale structure, and ionospheric currents. Below, we describe four instrument types that would provide measurements of the needed parameters to address the science topics discussed above.

Fabry-Perot Interferometer

To provide estimates of the neutral winds and temperatures in the ITM, narrow-field Fabry-Perot interferometers (FPIs) with a pointing capability to look in different directions can be deployed. Recent advances in FPI designs have improved the capabilities of these instruments by utilizing the high quantum efficiency of CCDs as the imaging device. These imaging FPI systems capture multiple orders of the interference pattern created by the interferometer, each of which can be individually analyzed and then combined to produce accurate estimates of winds and temperatures. Current systems have been shown to produce estimates with accuracies on the order of $5\text{--}10\text{ ms}^{-1}$ and 20 K for reasonable (5 minute) integration times at mid-latitudes. These accuracies are acceptable to address the science questions detailed above. The FPIs could use a filter tray or wheel mechanism to switch between different scientific filters. The most important filters for the FPIs would be those targeting the 630.0-nm emission for thermospheric studies and the 557.7-nm emission for mesospheric studies.

Wide-Angle Imaging System

In order to study the spatial-temporal characteristics of the ITM, wide-angle imaging systems equipped with a filter wheel for observations of multiple emissions can be employed. The systems must be wide-angle so that the combination of imaging systems in the network provide a near continuous spatial coverage of the thermosphere when combined. The emissions to be studied are quite dim, especially in contrast to the auroral emissions studied by the THEMIS array, for example. This requires longer time-integration for individual exposures in order to collect an appreciable signal. In order to increase the signal-to-noise ratio further, a high quality CCD camera cooled to very low temperatures would be used to reduce the effects of thermal noise.

Dual Frequency GPS Receivers

Dual-frequency Global Positioning System (GPS) receivers can be used to estimate the TEC from the receiver to any of the GPS satellites. If a receiver is capable of making high-rate (50-Hz) measurements of the received signal power from the satellites, small-scale irregularities in the ionosphere can be studied in

great detail. Although the US is already fairly densely populated by GPS receivers in the CORS network, receivers collocated with the instruments at individual NATION nodes would allow for direct comparisons without problems caused by geometry effects of non-collocated instruments. In addition, most of the CORS receivers are not capable of making high-rate scintillation measurements, so the deployment of scintillation-capable receivers would create a new space-weather monitoring capability across North America.

Magnetometer

Magnetometers measure the currents in the ionosphere, which can be utilized to understand the electrodynamic coupling between the magnetosphere and ionosphere. Coupled with the SuperDARN measurements of the ion flows and the Fabry-Perot measurements of the neutral winds, the magnetometer measurements provide a quantitative assessment of the coupling between the ionosphere and thermosphere. Fluxgate magnetometers have been utilized successfully for this type of study and they can be deployed at each of the NATION nodes.

Deployment and Operation

The network should be deployed so that the instruments will provide coverage of the ITM over a sizeable portion of North America. Additionally, the sites should be close enough so that the wind measurements from individual sites can be made in common volumes, allowing the calculation of vector winds at individual locations. All sites would be connected to the Internet, to allow for real-time transmission of data for use in Space Weather operational products. An initial deployment could consist of three latitudinal chains, as shown in Figure 1. Such a deployment would allow for observing geometries favorable to addressing the science questions detailed above. Many magnetometer sites exist across the US and Canada in remote locations that have favorable conditions for deploying sensitive instrumentation, so the majority of the NATION sites can be stationed at existing magnetometer sites, but it will be inevitable that new sites will have to be commissioned.

Cost Estimation

Based upon a deployment of 20 nodes, each fully instrumented with state-of-the art instruments, a rough cost estimate for the equipment is of the order \$4.5M. Deployment costs are estimated at approximately \$300K. Operation costs would be approximately \$25K/site/year of operation.

References

- Basu, S., et al., Ionospheric effects of major magnetic storms during the international space weather period of September and October 1999: GPS observations, VHF/UHF scintillations, and in situ density structures at middle and equatorial latitudes, *J. Geophys. Res.*, *106*, 30389, doi:10.1029/2001JA001116, 2001.
- Bruinsma, S., J. M. Forbes, R. S. Nerem, and X. Zhang, Thermosphere density response to the 20-21 November 2003 solar and geomagnetic storm from CHAMP and GRACE accelerometer data, *J. Geophys. Res.*, *111*, 6303, doi:10.1029/2005JA011284, 2006.
- Deng, Y., A. D. Richmond, A. J. Ridley, and H.-L. Liu, Assessment of the non-hydrostatic effect on the upper atmosphere using a general circulation model (GCM), *Geophys. Res. Lett.*, *35*, L01104, doi:10.1029/2007GL032182, 2008.
- Drob, D. P., et al., An empirical model of the Earth's horizontal wind fields: HWM07, *J. Geophys. Res.*, *113*, A12304, doi:10.1029/2008JA013668, 2008.
- Foster, J. C., and H. B. Vo, Average characteristics and activity dependence of the subauroral polarization stream, *J. Geophys. Res.*, *107*, 1475, doi:10.1029/2002JA009409, 2002.
- Foster, J., et al., *Solar-Terrestrial Physics: A Workshop*, Chapter Distributed Arrays of Small Instruments for Solar-Terrestrial Research: Report of a Workshop, National Academy of Science, National Academies Press, Washington, DC, ISBN-10: 0-309-10154-9, 2006.

- Fuller-Rowell, T. J., M. V. Codrescu, R. J. Moffett, and S. Quegan, Response of the thermosphere and ionosphere to geomagnetic storms, *J. Geophys. Res.*, 99, 3893, 1994.
- Hedin, A. E., et al., Empirical wind model for the upper, middle and lower atmosphere, *J. Atmos. Terr. Phys.*, 58, 1421–1447, doi:10.1016/0021-9169(95)00122-0, 1996.
- Jadav, R. M., K. N. Iyer, H. P. Joshi, and H. O. Vats, Coronal mass ejection of 4 April 2000 and associated space weather effects, *Planet. Space Sci.*, 53, 671, doi:10.1016/j.pss.2005.01.002, 2005.
- Lambour, R. L., A. J. Coster, R. Clouser, L. E. Thornton, J. Sharma, and T. A. Cott, Operational impacts of space weather, *Geophys. Res. Lett.*, 30, 36, 2003.
- Lu, G., A.D. Richmond, B.A. Emery, and R.G. Roble, Magnetosphere-ionosphere-thermosphere coupling: Effect of neutral winds on energy transfer and field-aligned current, *J. Geophys. Res.*, 100, 19,643, 1995.
- Makela, J. J., E. S. Miller, and E. R. Talaat, Nighttime medium-scale traveling ionospheric disturbances at low geomagnetic latitudes, *Geophys. Res. Lett.*, doi:10.1029/2010GL045922, in press.
- Meriwether, J.W., Studies of thermospheric dynamics with a Fabry-Perot interferometer network: A review, *J. Atmos. Solar-Terr. Phys.*, 2006.
- Miller, C. A., W. E. Swartz, M. C. Kelley, M. Mendillo, D. Nottingham, J. Scali, and B. Reinisch, Electrodynamics of midlatitude spread F. 1. Observations of unstable, gravity wave-induced ionospheric electric fields at tropical latitudes, *J. Geophys. Res.*, 102(A6), 11,521– 11,532, 1997.
- Miller, E. S., J. J. Makela, and M. C. Kelley, Seeding of equatorial plasma depletions by polarization electric fields from middle latitudes: Experimental evidence, *Geophys. Res. Lett.*, 36, L18105, doi:10.1029/2009GL039695, 2009.
- Ogawa, T., N. Nishitani, Y. Otsuka, K. Shiokawa, T. Tsugawa, and K. Hosokawa, Medium-scale traveling ionospheric disturbances observed with the SuperDARN Hokkaido radar, all-sky imager, and GPS network and their relation to concurrent sporadic E irregularities, *J. Geophys. Res.*, 114, 2009.
- Perkins, F., Spread F and ionospheric currents, *J. Geophys. Res.*, 78(1), 218–226, 1973.
- Rees, D., R. W. Smith, P. J. Charleton, F. G. McCormac, N. Lloyd, and A. Steen, The generation of vertical thermospheric winds and gravity waves at auroral latitudes-I. Observations of vertical winds, *Planet. Space Sci.*, 32, 667, doi:10.1016/0032-0633(84)90092-8, 1984.
- Sutton, E.K., J.M. Forbes, R.S. Nerem, and T.N. Woods, Neutral density response to the solar flares of October and November, *Geophys. Res. Lett.*, 33, L22101, doi:10.1029/2006GL027737, 2006.

Global Ionospheric Storms

Anthony J. Mannucci

Jet Propulsion Laboratory, California Institute Of Technology

THEME: Atmosphere-Ionosphere-Magnetosphere Interactions

EXECUTIVE SUMMARY

We discuss Global Ionospheric Storms as an important sub-field of study within solar and space physics. Significant science questions are discussed as well as methods to address them. Progress requires a robust observational approach that includes thermosphere/ionosphere imaging combined with in-situ measurements from space and ground-based measurements from globally distributed instruments. Because of the large-scale changes that occur over time scales of minutes, instruments operating continuously are needed to form the core observations. Small satellite constellations are also well suited to the study of GIS to characterize local time dependence of the ionospheric response.

1. Scientific Challenge

Understanding the storm-time ionosphere is fundamental to solar and space physics. Significant scientific attention and resources have been devoted to this topic over the last few decades, yet it remains fresh. Knowledge has been gained over the years, certainly a triumph for our field, yet this knowledge remains insufficient to predict how the ionosphere will behave during the next storm with a high level of confidence. We consider here “ionospheric behavior” in its broadest sense: behavior that spans time scales from minutes to days, and spatial scales from hundreds of meters to thousands of kilometers. The concerted effort of the past decades has provided an appreciation of the various physical phenomena that play a role in the ionospheric response to storms, yet we cannot pin down with confidence what will happen in any given event. The remaining gaps in our understanding are due to limited observations that are not up to the task of constraining the scientific hypotheses.

The concerted effort of scientists has provided a description of “qualitative” storm-time behavior, and our language used to describe these events is similarly qualitative: “positive phase” storm (increased electron density), “negative phase” (decrease), “prompt” (within minutes to a few hours) or “delayed” (response is delayed 12 hours or more). Yet, we do not have the ability to predict reliably which qualitative description will apply for a given storm, much less predict its quantitative magnitude. When considering the possible variations over a range of spatial and temporal scales, the limitations of the current state of knowledge are starkly revealed.

Recent research into the geospace response to coronal holes serves as a means to highlight today’s uncertainties. The “first cause” of an ionospheric storm is a change in solar wind conditions sunward of Earth, at the magnetopause. One expects that scientific descriptions of storm-time behavior must be closely tied to the categories of solar wind variation that cause the storm. The research literature is not particularly clear on this point. For example, although coronal holes as a heliospheric phenomenon have been known for at least three decades, the distinct ionospheric response to coronal holes is only now being characterized. Ionospheric storm phenomena have often been organized in terms of geomagnetic indices, e.g. the 3-hour Kp index. The problem is that a given Kp value can arise from a wide variety of solar wind conditions. Thus attempts to predict storm behavior based on Kp will suffer from ambiguities in working backwards from Kp to the solar wind causative factors.

Storm time response has also been described in terms of specific solar wind quantities (interplanetary magnetic field, plasma dynamic pressure or speed) that trigger the ionospheric response. For example, important high latitude drivers of storms are organized in statistical models according to the solar wind speed and magnetic field components (e.g. Weimer model). This naturally leads to a storm time description that is tied to the solar wind drivers. Yet, existing descriptions are insufficient because

they often ignore how the solar wind variation is itself organized according to phenomena originating near the Sun. A standard categorization of geoeffective solar surface phenomena identifies “coronal mass ejections” and “coronal holes.” In the interplanetary medium, these phenomena lead to “interplanetary coronal mass ejections” (ICME) and “high speed streams” (HSS). At 1 AU (at Earth), the solar wind characteristics of ICMEs and HSSs are generally quite distinct. Now consider that the ionospheric response depends critically on the detailed behavior of the solar wind, including correlations among the different solar wind variables (magnetic field, speed, density, temperature, etc.) and the time scales over which these variables change. It is natural to categorize storms as “HSS storms” and “ICME storms”, because these two distinct solar wind phenomena probably help organize the solar wind variation. Yet, the ionospheric characteristics of HSS storms are only now being explored as a distinct phenomenon, lagging by several decades the knowledge of HSS existence. This lag is due to limited physical understanding of ionospheric storms, and the fact that it has been a struggle simply to identify the general characteristics and basic physical processes at play during storms, much less understand how those processes are tied to the solar wind.

A particularly fruitful area of research is the spatially localized responses (~100-200 km) that occur in the midst of a major global scale ionospheric perturbation. The last decade has witnessed a significant focus on localized storm-enhanced density effects that appear at sub-auroral latitudes. Quasi-imaging observational approaches based on dense networks of ground-based GPS receivers have revealed these features, but considerable controversy exists over the physical processes at work. This leads to an important question: are poorly understood fundamental physical processes involving plasma and ion-neutral physics playing a role in the formation and evolution of stormtime structures? The easy answer to this question is “yes”, yet the scientific questions regarding these localized processes have yet to be crisply identified.

2. Phenomena

Figure 1 can help motivate the idea of a global ionospheric storm (GIS), although it should not be interpreted as to restrict what a global ionospheric storm might be. This figure shows the large plasma redistribution and creation that occurred rapidly after storm onset for the Halloween storm of October 2003, as observed by total electron content (TEC) measurements above the CHAMP satellite orbiting at 400 km altitude. The following questions remain to be answered about such “superstorm” events:

- What are the relative roles of winds and electric fields in creating the storm time response?
- To what degree is the acting electric field solely of magnetospheric origin, or could there be a dynamo component also (the latter would generally counteract the magnetospheric component during daytime).
- What physical processes lead to the large electric field magnitudes that were estimated for this event (> 4 mV/m at the equator)?
- What is the impact of plasma transport across latitudes?
- What determines the latitudinal extent and structure of the prompt ionospheric response? What is the longitudinal profile of the response? Is there a dependence of the response on the UT onset of the storm? What is the local time response of the storm? What role do seasons play? If multiple storms occur in succession (e.g. October 29 and 30, 2003) is there a significant preconditioning effect? Will preconditioning increase or decrease the ionospheric response?
- How and why do significant mid-latitude plasma gradients form, and what are physical processes at play?
- What is the extended (multiday) response of the thermosphere and ionosphere to intense ICME generated storms. What tends to suppress the response (e.g. November 10, 2003; see Mannucci et al., 2009)? What determines time scales for recovery? How does the recovery phase vary globally? What role is played by the plasmasphere during the storm recovery phase.
- What is the influence of solar EUV flux on the stormtime response?

Of course, a major overarching question is: what solar wind characteristics determine the observed responses (e.g. Figure 1)?

Figure 1 illustrates the daytime GIS response. Nighttime phenomena have recently been discovered over Puerto Rico and Japan, which show the existence of mesoscale propagating irregularity waves and areas of ionospheric depletions with widths of a few hundred kilometers. These mid-latitude features are a few thousand kilometers in length with strong electron density gradients occurring over dimensions of 10 km. Nighttime features during superstorms deserve further study.

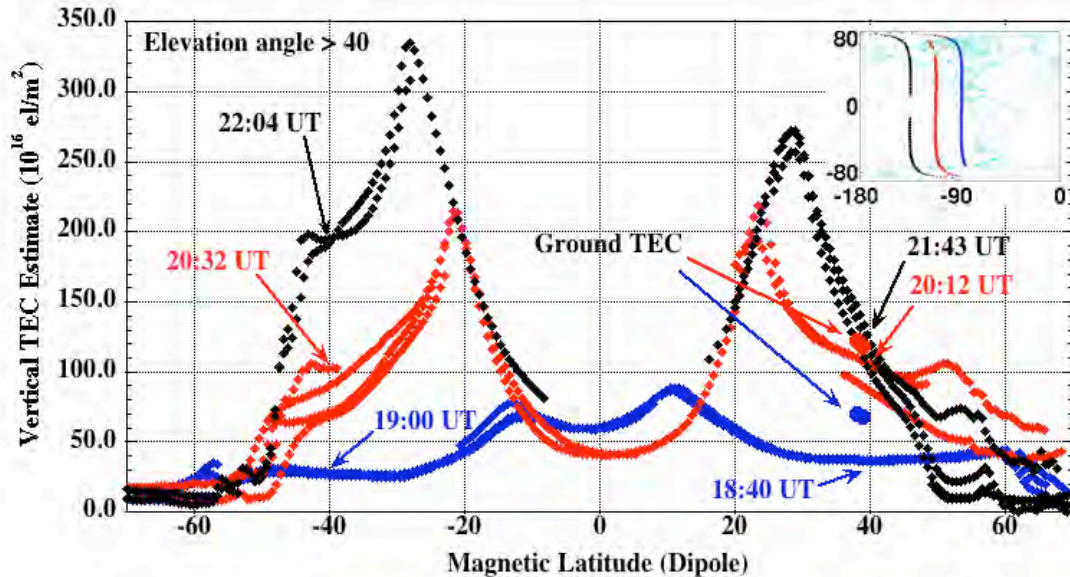


Figure 1: Global scale response to a super storm. The plotted quantity is total electron content (approximately vertical) above the CHAMP satellite orbiting at 400 km altitude. Three daytime passes are shown. The local time near the equator for all three passes is 1300. The blue trace precedes the storm. The red trace occurs after a large B_z southward excursion has reached the magnetopause. The black trace is the next pass. The ground track of the satellite is shown in the upper right. From Mannucci et al., 2005.

There have been relatively few reports of the characteristics of middle-latitude auroras that accompany superstorms. The characteristics and consequences of such large auroral zones should be studied. The enlarged auroral oval could significantly change the physics of penetration electric fields of magnetospheric origin. In fact, the entire structure of the magnetosphere is likely altered significantly during these events. What is the characteristic of “monster auroras” and do they modify the ionosphere on global scales? Little is known. In the boundary region between the auroral zone and middle latitudes, strong gradients form in electron density, conductivity, and temperature, likely altering significantly the global scale thermospheric circulation and ionosphere-magnetosphere interactions

Middle Latitude Structuring

A remarkable feature of GIS is the structures that form at middle to low latitudes that have yet to be explained. Unusually large (factor of 10 TEC increase) localized midnight electron density enhancements at mid-latitudes have been reported. One of these events occurred over Florida during the local nighttime of October 30-31, 2003 [Mannucci et al., 2005; Datta-Barua et al., 2008]. See Figure 2. Preliminary evidence suggests that other superstorms have produced similar events. The following science questions should be addressed:

- 1) What is the electron density structure and time evolution for these events?
- 2) What types of geospace conditions and prior storm conditions are necessary for these enhancements to occur?

- 3) What are the occurrence statistics of these events? Do they occur in areas other than Florida? Are there always conjugate enhancements in the Southern Hemisphere?
- 4) What are the physical mechanisms that lead to these nightside enhancements? Is there any connection between these events and equatorial ionospheric phenomena? Why do the events co-rotate and appear to remain earth-fixed over several hours?

The question arises of where such mid-latitude structuring occurs during superstorms. It is certainly the case that the effects of global ionospheric storms are easier to observe over North America, where GPS receiver chains are quite dense. Satellite observations are needed to answer the global question. Figure 3 shows a pass of the JASON satellite over the Pacific Ocean and the corresponding vertical TEC measured below the satellite by the altimeter system, for the October 30, 2003 superstorm. The ground track of the JASON satellite is shown in the insert. During this global ionospheric storm mid-latitude TEC values increased by up to a factor of ten (point A) and elevated TEC values extend poleward beyond 50° south latitude where there is an extraordinarily sharp gradient in TEC (point B). JASON transited local times of roughly 1100 LT to 1400 LT during this pass. The black trace shows a typical quiet time electron density distribution for the same local time as the storm time pass, and a similar longitude. This example demonstrates a high degree of mid-latitude structuring during the storm, the cause of which is unknown.

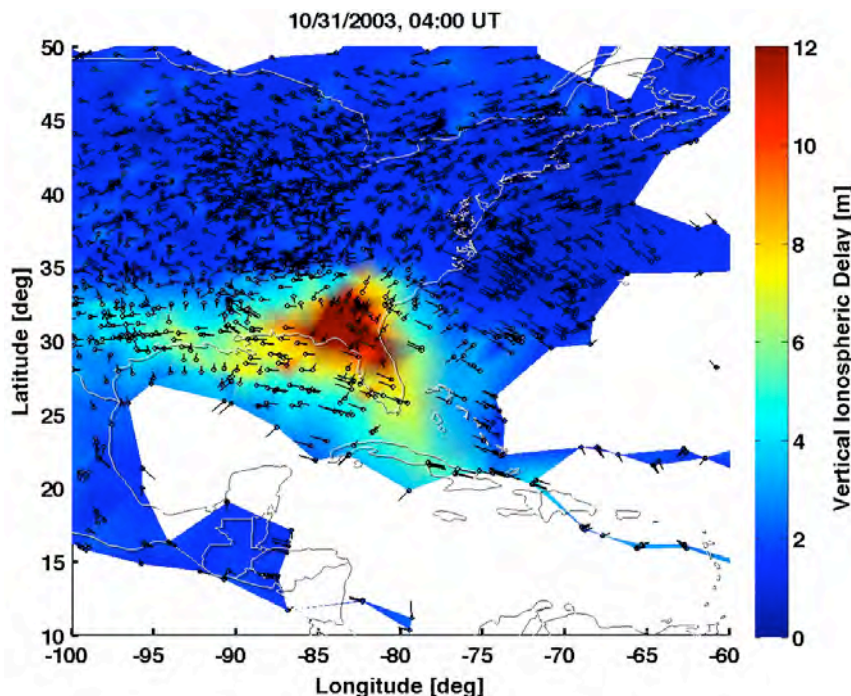


Figure 2. TEC over the eastern US and Caribbean at 2300 LT. GPS measurements (black circles with a ray pointing back toward the GPS receiver) from about 400 stations, converted to an equivalent vertical TEC assuming a 500 km ionospheric shell, and linearly interpolating TEC between nearest measurements. One meter of ionospheric delay corresponds to 6.13 TECU. [after Datta-Barua, 2004].

Plasma Gradients

Generally nature abhors strong gradients and tries to eliminate them with a combination of diffusion, instabilities, and mixing. Yet uncharacteristically large ionospheric electron density/TEC gradients are obvious in previous figures. For example, Figure 3 shows a change of nearly 200 TECU in one degree of latitude. Figure 4 illustrates how the ionosphere can change significantly from the viewpoint of three nearby GPS receivers in the US (Pittsburgh, Hawk Run, PA and USNO in DC). The

plot is position error caused by excess ionospheric delay above the receiver. The large change in range error occurs rapidly at each site, but is clearly distinct among the sites. This suggests sharp TEC gradient scale lengths of a few tens to hundreds of km and that the gradients are moving horizontally relative to the site locations. Large gradients have been observed to persist over the US heartland for 2-3 hours and are 30 to 100 times more intense than gradients produced by diurnal solar illumination.

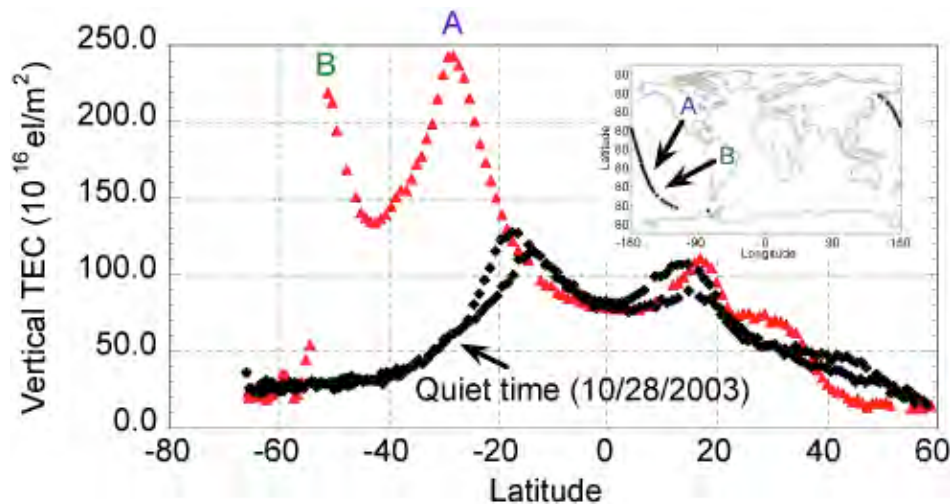


Figure 3. VTEC as measured from the JASON S/C during the October 30, 2003 super-storm. The large VTEC value of nearly 250 TECU was measured over the South Pacific at about 1400 LT.

In contrast to both high and low latitude ionospheric gradients that exist during quiet times, there is limited scientific understanding of what produces and maintains these very large mid-latitude gradients, how often such sharp gradients occur, their full spatial extent, or their temporal evolution. Previous research has provided some hints at the physics involved in generating ionospheric gradients, which are, broadly speaking, the result of either enhanced plasma production, enhanced plasma erosion, or a combination of the two. For example, upward plasma drift caused by electric fields can increase the total plasma quantity during daytime conditions as solar production is maintained and plasma loss is reduced at higher altitudes (Figure 1). Alternatively, significant plasma erosion is thought to occur along the edges of the mid-latitude ionospheric trough through heating associated with fast plasma drifts [Schunk et al., 1976]. This gradient might be associated with the boundary of the high-latitude convection pattern, or due to SAPS electric fields [Foster et al., 2002].

Global (geostationary) imaging is needed to characterize the creation and stability of hemispheric-scale severe density gradients during magnetic storms. Imaging will determine how these structures are interconnected across latitudes and whether they interact with the different geophysical phenomena that they may connect with, such as the equatorial anomaly, auroral oval, plasmaspheric drainage plumes, the high-latitude convection pattern, and inner magnetospheric penetrating electric fields.

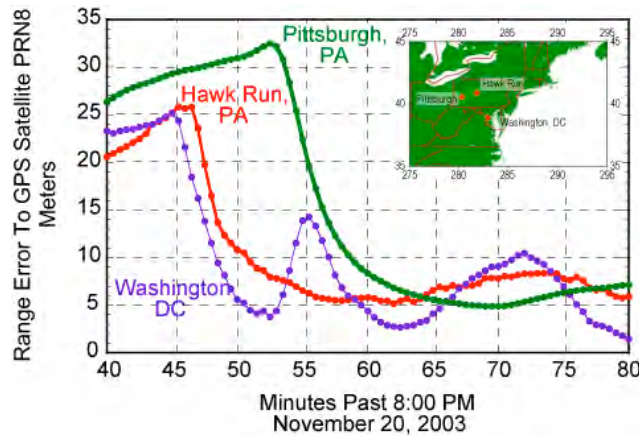


Figure 4. GPS position error and three nearby sites in the US during the November 2003 superstorm. The position error is caused by ionospheric delay. The rapid changes with time indicate significant ionospheric gradients drifting rapidly above each site.

3. Systems Science

The large increase in energy input leading to Global Ionospheric Storms also creates coupling between different regions of geospace that might be obscured during more quiet times. The AIM community has been increasingly focused on the “systems science” perspective as reflected in the CEDAR Strategic Plan (CEDAR, 2010). ***The focus on systems science is needed because a strongly interacting system cannot be described or understood by studying its components in isolation from the larger system.*** An example of system interactions that are hypothesized to occur during GIS is shown in Figure 5. The diagram schematically illustrates how low-to-middle latitude plasma transport, caused by penetrating magnetospheric electric fields and solar EUV production, leads to increased plasma outflow into the magnetosphere, in turn producing ring current energization and additional electric field penetration as well as plasmaspheric changes and changes to the dayside reconnection rate. This complex web of interactions is only in the initial stages of being understood.

4. Missions And Concluding Remarks

Unraveling the complex phenomena that constitute Global Ionospheric Storms requires a robust observational approach combining space-based observations with globally distributed ground-based observations. This scientific problem is certainly worthy of a flagship-class mission, although alternative strategies are certainly possible. It is necessary to combine full Earth-disk imaging from geostationary orbit if possible, with in-situ observations that measure the forces undergoing dynamic changes. In-situ sampling of the boundary region between the auroral and middle latitudes as the auroral zone expands rapidly is a high scientific priority. Small satellite constellations are well suited to GIS science also because of the need to sample multiple local times simultaneously. Ionosphere focussed missions should be combined with magnetospheric and solar wind data to address interactions at the systems level.

5. Acknowledgement

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The author acknowledges with heartfelt gratitude the influence of Paul M. Kintner, who is responsible for promoting Global Ionospheric Storms as a field of scientific inquiry. Professor Kintner chaired the Mission Definition Team of the Geospace Storm Probes mission (report in 2002).

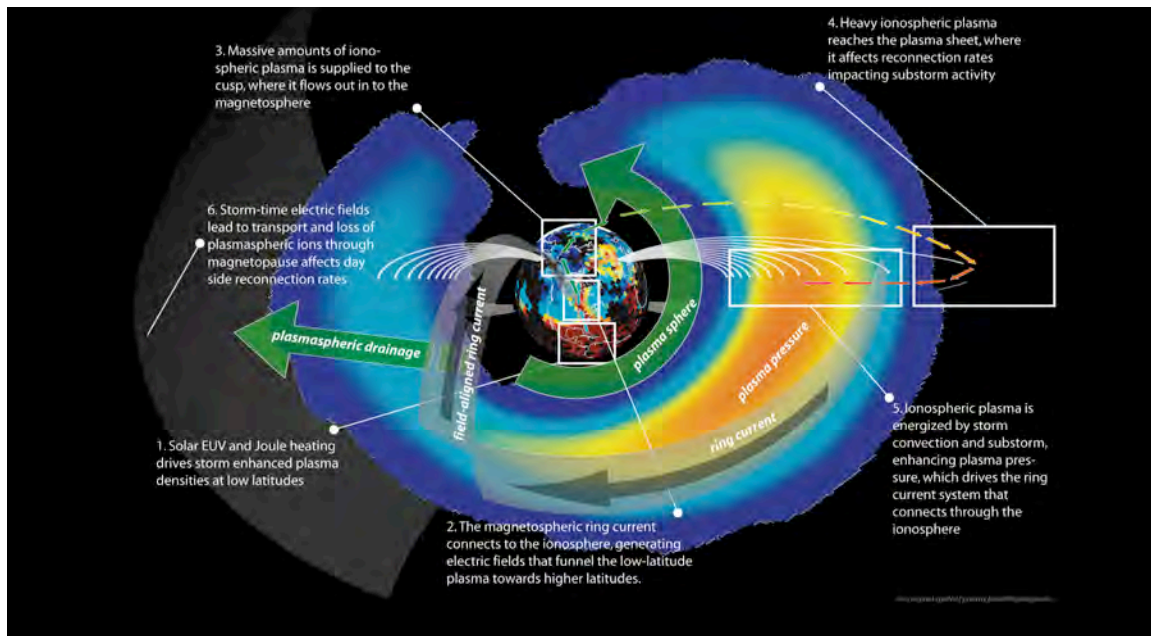


Figure 5. The systems view of outflow and ionosphere-magnetosphere interactions during a Global Ionospheric Storm. What are roles of magnetospheric dynamics, plasmaspheric dynamics and structuring? From Tim Fuller-Rowell's presentation of the Living With A Star Focused Science Team, AGU Fall Meeting, 2008.

6. References

- "CEDAR: The New Dimension", version 9.2 dated October 2010. See: http://cedarweb.hao.ucar.edu/wiki/index.php/Community:CEDAR_Strategic_Plan.
- Datta-Barua, S. (2004), Ionospheric Threats to Space-Based Augmentation System Development, *Proceedings of the Institute of Navigation Satellite Division Meeting*, Long Beach, California: September 21-24.
- Foster, J.C., P.J. Erickson, A.J. Coster, J. Goldstein, and F.J. Rich (2002), Ionospheric signatures of plasmaspheric tails, *Geophys. Res. Lett.*, 29(13), 1623, doi:10.1029/2002GL015067.
- Mannucci, A. J., Datta-Barua, S., Walter, T., Komjathy, A., Sparks, L., Tsurutani, B.T. (2005), Anomalous Nighttime Plasma Structure in the Recovery Phase of a Superstorm, AGU Fall Meeting Abstracts, A275.
- Mannucci, A. J., B. T. Tsurutani, B. A. Iijima, A. Komjathy, A. Saito, W. D. Gonzalez, F. L. Guarnieri, J. U. Kozyra, and R. Skoug (2005), Dayside global ionospheric response to the major interplanetary events of October 29–30, 2003 "Halloween Storms," *Geophys. Res. Lett.*, 32, L12S02, doi:10.1029/2004GL021467.
- Mannucci, A. J., B. T. Tsurutani, M. C. Kelley, B. A. Iijima, and A. Komjathy (2009), Local time dependence of the prompt ionospheric response for the 7, 9, and 10 November 2004 superstorms, *J. Geophys. Res.*, 114, A10308, doi:10.1029/2009JA014043.
- Schunk, R.W., P.M. Banks, and W.J. Raitt (1976) "Effects of electric fields and other processes upon the nighttime high-latitude F layer," *J. Geophys. Res.*, 80, 3271- 3282, 1976.

Research To Operations: Continuous Improvement

Anthony J. Mannucci

Jet Propulsion Laboratory, California Institute Of Technology

THEME: Atmosphere-Ionosphere-Magnetosphere Interactions

EXECUTIVE SUMMARY

The last few years have witnessed an explosion of new space weather operational capabilities, serving civilian and military users. Tremendous progress has been made implementing operational capability that can be traced back to requirements. All too often, however, the requirements exceed what is possible to achieve. Despite the fact that requirements are not being met in a strict sense, the deployed operational capabilities are clearly providing value to users. Operational effectiveness benefits from research targeted towards improving capability. Drawing an analogy to terrestrial weather, this white paper suggests a path forward that, over time, will lead to steadily improving operational capabilities in space weather.

1. Description

The existence of a requirement does not imply the requirement can be met. The space weather community (civilian and military) has made enormous strides in developing operational space weather capabilities of value to users. These capabilities have been developed based on a set of user requirements. However, the complex nature of space weather often implies that the implemented capability does not meet the stated requirements. Deploying an operational capability that does not meet requirements is still useful and of significant value, compared to there being no capability at all. In those cases where requirements are not yet met, there is a need to continuously improve the operational capability so that over time the gap will narrow between performance and stated requirements. A program of continuous improvement will succeed if there is a research community oriented towards producing research results that lead to operational improvements.

We advocate the development of a research community that is focused on operational effectiveness. Such a community has been developed in terrestrial weather. It is possible to do so for space weather as well.

Using the terrestrial weather community as a guide, we describe elements needed to create a robust community of researchers focused on continuously improving operational capability over time. Figure 1 shows a graph provided by the European Center For Medium Range Weather Forecast (ECMWF), the leading European weather prediction center. The figure shows the improvements to forecasting effectiveness that have occurred over the past thirty years. A standard forecasting metric (anomaly correlation of geopotential height at 500 hPa pressure level, global average) is plotted as a function of time, for different forecast periods (3-day to 10-day) and for Northern and Southern hemispheres. This figure demonstrates that the operational capability is continuously

being improved. Such improvement is needed when the delivered operational capability does not yet meet all the requirements.

The ability to produce such figures is of enormous programmatic value and provides a justification for continued funding to improve operational capability. Two key elements are needed to produce such a figure: 1) developing and maintaining at least one metric that is applied consistently over time; and 2) constant monitoring of forecast quality using that metric to assess operational capability.

We believe the following elements are important also: *open access to the metric value, and documentation of the methods used for forecasting*. Providing researchers with easy access to the forecast metric provides those researchers with the opportunity to propose new ideas that improve the forecasts. If the research community understands the methods and models underlying forecast effectiveness, that community can propose ideas to improve effectiveness. Conversely, if the methods and models underlying forecasting performance are not widely known or disseminated, a robust research community with a stake in performance improvement will not develop.

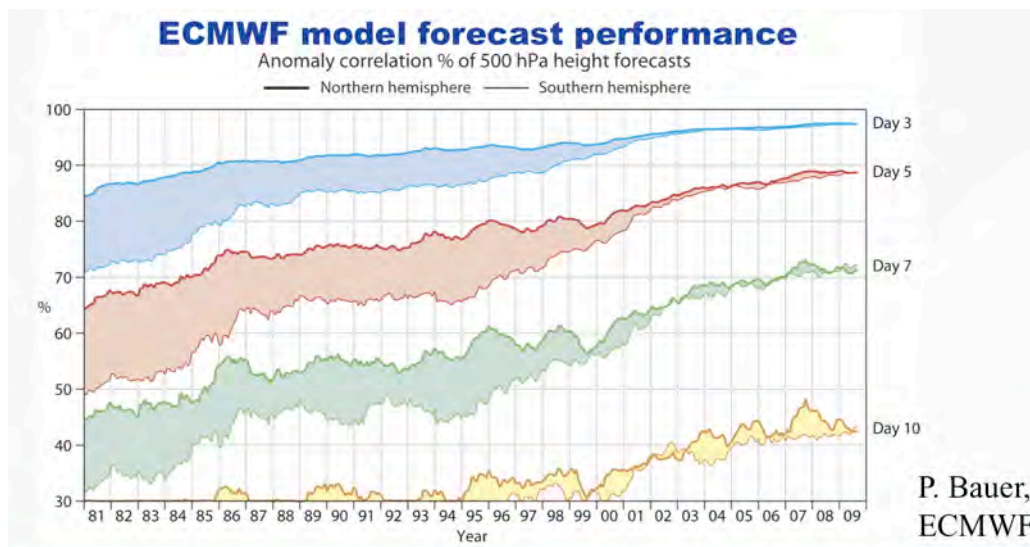


Figure 1. Improvements to weather forecasting effectiveness over time.

2. Concluding Remarks

The National Space Weather Program Assessment (National Space Weather Program, 2006) conducted by the Office of the Federal Coordinator for Meteorological Services and Supporting Research reached the following set of recommendations to “strengthen the science-to-user chain”:

- Maintain and strengthen both targeted and strategic space weather research
- Enhance emphasis and resources for transition of models to operational users
- Increase the private sector role in supplying products and services.

We believe that item 2 above, transitioning models to operations, is receiving sufficient attention. The recommendations suggested in this white paper are meant to strengthen

item 1: targeted space weather research. We have suggested two key elements needed to strengthen research targeted towards operational effectiveness. These two elements are: access to a forecasting metric applied consistently over time, and knowledge dissemination of the models and methods used for forecasting.

3. Acknowledgement

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

4. References

National Space Weather Program (2006), “Report of the Assessment Committee for the National Space Weather Program”, FCM-R24-2006, June 2006.

GNSS¹ Geospace Constellation (GGC): A Cubesat Space Weather Mission Concept

Anthony J. Mannucci, Jeff Dickson, Courtney Duncan, Ken Hurst
Jet Propulsion Laboratory, California Institute Of Technology

THEME: Atmosphere-Ionosphere-Magnetosphere Interactions

EXECUTIVE SUMMARY

Investment in radio occultation science using GPS receivers in low-Earth orbit is bringing enormous scientific benefit to the Atmosphere-Ionosphere-Magnetosphere community. Recent technology developments permit highly miniaturized receivers that provide high quality science data for ionospheric studies. Such receivers are compatible with the Cubesat form factor, raising the possibility of swarms of Cubesats providing unprecedented global coverage of ionospheric electron density measurements throughout the E and F-region ionosphere and up into the plasmasphere. In this white paper, we describe the science and technology needed to achieve this result. ***We recommend technology investment in miniaturized GPS receivers for ionospheric remote sensing.*** Such Cubesat swarms, either as a dedicated constellation or an ad-hoc constellation deployed via launches of opportunity, would be a profoundly useful resource for advancing AIM science, providing plasma measurements to accompany other measurements of the forces that influence ionospheric structure such as solar EUV, thermospheric winds and composition, and ionospheric electric fields.

1. Science

Applying GPS technology to ionospheric measurement has produced an explosion of new science results. The reason is relatively simple: GPS has provided a quasi-imaging capability for ionospheric total electron content. This capability has revolutionized our field because it allows scientists to visualize ionospheric structure on a range of spatial scales, from 100s to 1000s of km, globally and continuously. This has produced discoveries and new science in a range of areas, from extreme ionospheric storms to long-term trend studies and studies of ionospheric irregularities.

The earliest GPS work used receivers on the ground. Starting in 2001, with launch of the CHAMP and SAC-C satellites, a continuous stream of ionospheric data became available from GPS receivers in low-Earth orbit. Using the radio occultation (RO) technique, first pioneered on GPS/MET (1995-1997), the community has demonstrated the feasibility of retrieving electron density profiles from the total electron content measurements obtained from the receivers. Electron density is retrieved using an “onion peeling” (Abel transform) retrieval approach. See Figure 1. While the accuracy of unaided retrievals will vary, it is clear there is a wealth of scientific information in the profiles and in the rawer total electron content measurements obtained in limb viewing

¹ GNSS: Global Navigation Satellite System

geometry. As an added bonus, GPS measurements from ground and space are now being used as input data to assimilative models for research and operations, bringing a new capability to space weather that is analogous to the numerical weather prediction capabilities existing in terrestrial weather.

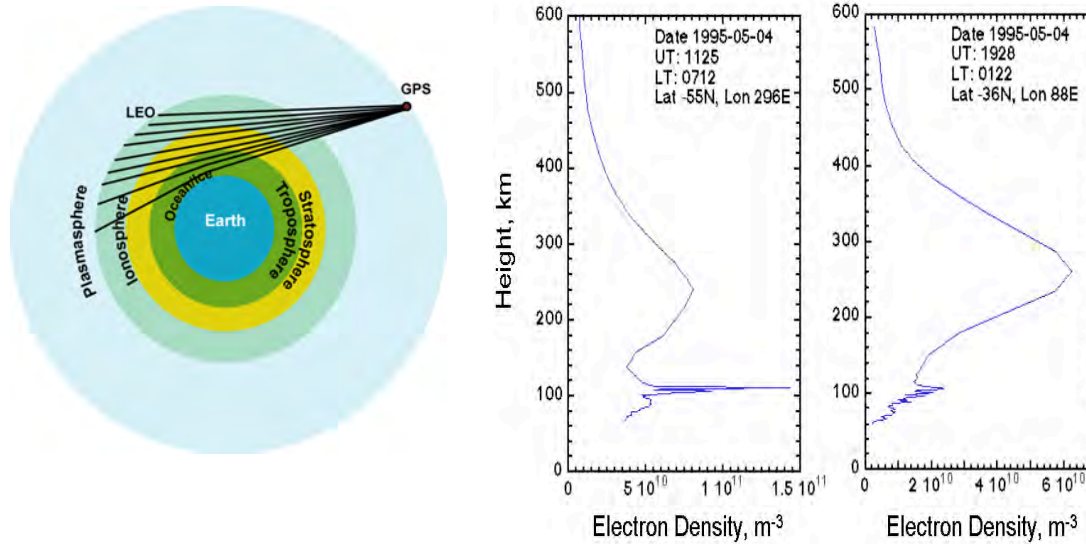


Figure 1. Radio occultation acquisition geometry (left). Retrieved electron density profiles from GPS/MET (right) should show strong E-layer signatures. The GPS receiver is onboard the Low Earth orbiting satellite (LEO) tracking the transmissions from the GPS satellite (GPS).

It is clear the scientific community will benefit from increasing numbers of orbiting GPS receivers. Increased numbers of receivers increases the spatial and temporal resolution of the quasi-images that are obtained. To quantify this benefit, we refer to Figure 2 [Schreiner, 2009]. This shows the number of occultations (electron density vertical profiles) that will occupy a $500 \times 500 \text{ km}^2$ region on average per day from the COSMIC-2 constellation of 12 satellites. Depending on latitude, the 7-10 available profiles per day available represents a historic advance in global coverage for ionospheric studies. However, considering the importance of mesoscale structure, the strong diurnal variations in the ionosphere that significantly affect its vertical structure, and the rapid changes with UT occurring during disturbed periods (e.g. TEC doubling and lifting of a few hundred km in 2-3 hours) it is clear that additional satellites would provide significant scientific benefits enabling new science investigations. We believe that deploying 100-200 receivers on orbit will produce the next quantum advance in scientific capability and new science.

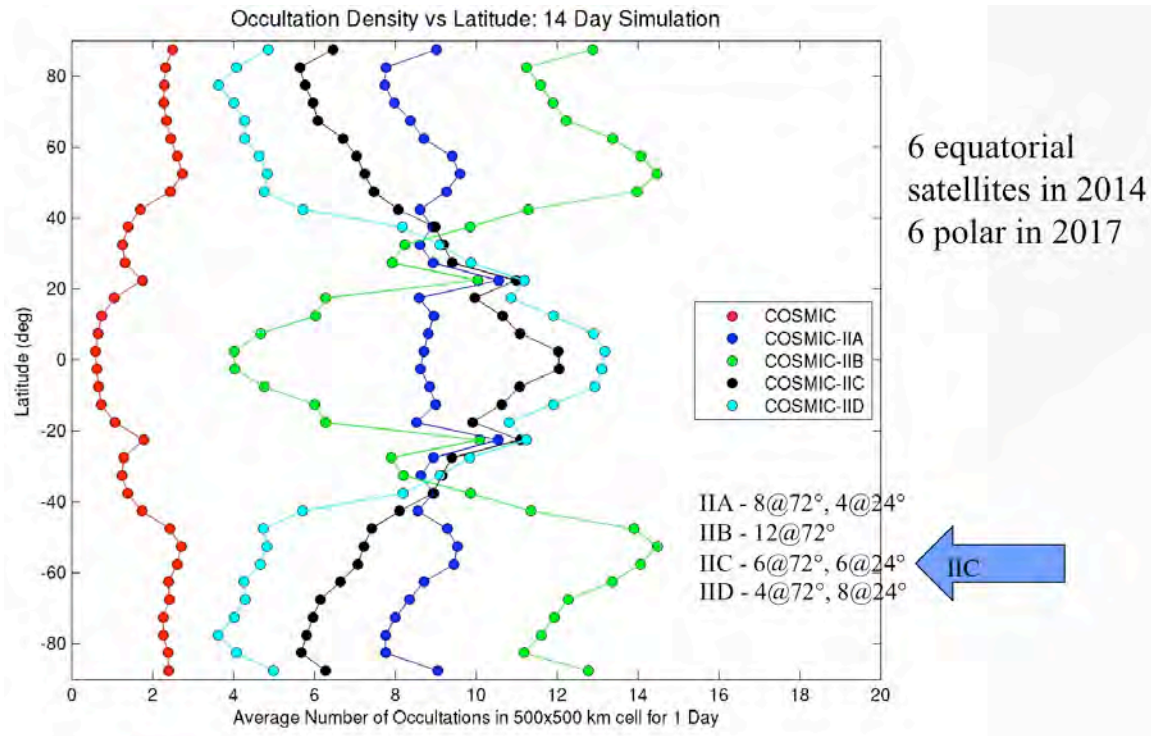


Figure 2. Number of occultations per 500x500 km² region, by latitude, for candidate COSMIC-2 constellations. From a presentation by W. Schreiner, UCAR, presented at the OSSE COSMIC Workshop, Taiwan, August 2009.

2. Technology

The enabler of new science is the possibility of deploying a sufficient number of GPS receivers in low Earth orbit so that the quasi-imaging capability that has been so successful from ~1000 ground-based receivers is brought to the vertical dimension (altitude). Reducing the power, size and mass of ionospheric radio occultation instruments offers the best hope of increasing the resolution of the images and increasing the probability of maintaining the data continuity through time, sampling a wide variety of solar cycle and atmospheric conditions.

Technology is within reach to produce a low power (<3W) low cost (< \$50K each in quantity) instrument for ionospheric remote sensing from space. We base our assessment on the availability of commercial receivers such as the Novatel OEMV-3 board. Table 1 has relevant engineering data for this board. The Novatel receiver board is compatible with a Cubesat power budget and form factor.

Table 1: Novatel OEMV-3 GNSS Receiver

Characteristic	Value
Power	2.1 W
Mass	75 g
Dimension	85x125x13 mm

The Novatel GNSS (Global Navigation Satellite System) receiver has a number of features well suited for our targeted science measurements. Unlike most commercial GNSS receivers, which only produce position outputs based on single (L1) frequency tracking, the Novatel produces the carrier and code phase observables necessary for science. It supports all current and planned GNSS signals such as Galileo and modernized GPS L2C (new codeless² second frequency), L5 (new codeless third frequency), as well as tracking of the current encrypted L2 signal. Access to the receiver version that supports operation in space is possible by NASA centers such as JPL, including the capability to track at high altitudes and high speed.

Radiation testing on Novatel receivers indicates they are suitable for a Cubesat orbit and mission lifetime (Markgraf et al., 2004). In a technology demonstration effort, we recommend leveraging commercial off the shelf parts to keep costs low.

Figure 3 is a sketch of the Cubesat concept. The following characteristics of the system are favorable for Cubesat also:

Spacecraft accommodation: Pointing requirements are modest: 5-10 degrees knowledge and control suffices due to the use of hemispherical patch antennas in the RAM and anti-RAM velocity directions. RO requires three-axis stabilization. Cubesats are rapidly evolving to achieve such stability.

Antenna: two patch antennas could be used (one fore/RAM and one aft/anti-RAM), at ~100g each with dimensions 40x40x10 mm. Cables might add an additional 100-300g of mass, depending on configuration.

² “Codeless” refers to the fact that the signal is not encrypted, available for civilian use.

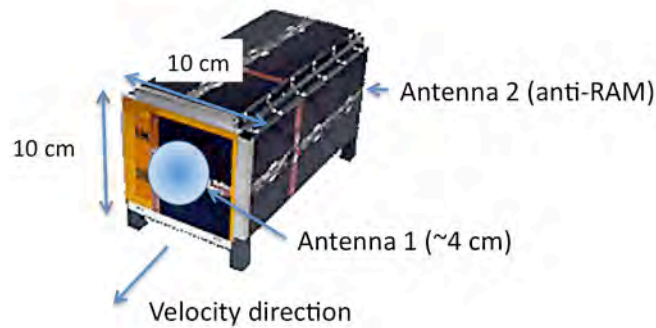


Figure 3. Schematic of the Cubesat with radio occultation capability.

3. Conclusions And Recommendations

We recommend technology development to enable GNSS receivers compatible with the Cubesat form factor for remote sensing of the Earth's ionosphere and plasmasphere. A radio occultation receiver should be considered for deployment on all future NASA and NOAA space weather and Earth science missions (see National Research Council, 2007 for a similar recommendation).

4. Acknowledgement

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

5. References

- Markgraf, M. and O. Montenbruck, Total Ionizing Dose Testing of the NovAtel OEM4-G2L GPS Receiver, DLRGSOC TN 04-04, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, 2004.
- National Research Council (2007) "Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond", *Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future*, ISBN: 0-309-66900-6.
- Schreiner, W. et al., (2009) "Simulations of COSMIC Follow-On Sounding Distributions and Data Latency for OSSE Studies," presented at *Workshop on OSSE for the FORMOSAT-3 / COSMIC Follow-on Mission*, Taiwan, September, 2009.

Estimating The Forces That Drive Ionosphere And Thermosphere Variability: Continuous Data And Assimilative Modeling

Anthony J. Mannucci, Xiaoqing Pi, Mark Butala, Phil Stephens, Brian D. Wilson, Attila Komjathy, Byron A. Iijima, Vardan Akopian, Miguel Dumett
Jet Propulsion Laboratory, California Institute Of Technology

THEME: Atmosphere-Ionosphere-Magnetosphere Interactions

EXECUTIVE SUMMARY

Ionosphere-thermosphere science has long been hampered by a lack of measurements of the underlying forces that determine plasma structure. These forces include electric fields, thermospheric composition, winds and solar EUV irradiance. All of these forces are dynamic quantities and of great importance to the underlying science. The rapid proliferation of electron density and total electron content measurements from ground and space-borne GPS receivers and other instruments suggests a great potential in the following indirect approach to estimating these forces: developing a model-based approach to retrieving the forces from the electron density structure and dynamics. The research community has just begun to explore such an approach, which addresses significant science questions previously out of reach. ***We recommend community investment in developing such “inversion” approaches. After further development, the capability should be maintained and implemented systematically on an ongoing basis, providing scientific insight into the causes of ionosphere and thermosphere variability.***

1. Description

Efforts to model daily thermospheric weather and storm-time effects have been hampered by a lack of direct measurements on global scales of neutral composition and density, winds and electric fields. In contrast, ionospheric measurements of electron density and integrated electron content from ground GPS networks, space-based GPS receivers, and space-based UV airglow instruments are a growing source of continuously available global data. The community is poised to develop a data assimilation methodology for estimating thermospheric composition and density, winds, and electric fields from measurements of electron density and total electron content. By combining the strengths of both thermospheric and ionospheric datasets to retrieve thermospheric parameters and their ionospheric effects, one can generate thermospheric characterizations with improved global distribution and temporal resolution. These improved thermospheric datasets will significantly increase scientific understanding of coupled geospace system and benefit the development and validation of coupled, first-principles thermosphere-ionosphere models.

Inversion of thermospheric and electrodynamic drivers from electron density data leverages assimilative modeling approaches developed over the last decade [Pi et al., 2003; Schunk et al., 2004]. The physics-based assimilative modeling approaches start

with a climatological representation of the three-dimensional (3D) electron density distribution based on climatological representations of the underlying driving forces such as neutral composition and density, winds and electric fields. Using the global GPS and other data sets, differences are calculated between the actual electron density and that of the physics-based model. The differences between measured and modeled electron density are the result of differences between the climatological and actual driving forces, assuming that the physics-based models correctly model the relevant physical processes. Thus, the present era characterized by a plethora of electron density measurements presents an opportunity to determine the actual driving forces, by adjusting drivers in the physics-based models so that the adjusted model better fits the data. The research community has experimented with sophisticated mathematical and computational methods to perform the adjustment, yielding the potential to develop an efficient system for estimating the driving forces [Pi et al., 2003; Scherliess et al., 2009; Luan and Solomon, 2008].

We recommend that the community develop a driver estimation capability leveraging progress made in the last few years in modeling and data assimilation. Additional resources include observation system simulations and coupled thermosphere-ionosphere models (e.g. TIMEGCM) in addition to the assimilative model (e.g. GAIM). Coupled models such as TIMEGCM provide a physically self-consistent and plausible representation of the ionosphere-thermosphere state. Model output can be used to test and examine driver estimation algorithms and processes. Once the effectiveness of thermospheric retrievals has been demonstrated using simulations, data assimilation experiments for estimating thermospheric and electrodynamic parameters should be conducted using real measurements from densely-distributed global and regional networks of GPS receivers, spaceborne GPS radio occultation sensors, and ultra-violet airglow emissions from the TIMED GUVI instrument and similar satellite sensors. Validation of the driver retrievals can be conducted with data collected from incoherent scatter radar facilities and satellite measurements of the drivers where they exist (e.g. electric fields from C/NOFS and neutral composition from TIMED).

Our description here is somewhat simplified because model error in the form of “missing physics”, which certainly exists particularly over the smaller scales (200 km or less), is a confounding factor. Yet, we expect this will not be a dominant error for studies that focus on larger-scale (1000+ km) driver averages. Also, we assume that the inversion process is run over a sufficiently long period (at least several days) to ensure that initialization error is not a significant factor. As model physics improvements are made through research, so will the inversion approach improve as these model improvements are transferred to the data assimilation capability.

The reason for pursuing this approach is that direct measurements of thermospheric quantities such as composition and winds are far more sparsely distributed than measurements of ionospheric electron density which have improved global distribution and are more widely available. Our goal is to use assimilative modeling to leverage ionospheric data sets and generate thermospheric and electrodynamic data sets with improved global distribution.

After the proposed methodology becomes established, the development and validation of coupled thermosphere-ionosphere first principles models will significantly benefit from large quantities of ionospheric observations, such as those available from the global positioning system satellites, and space-based UV radiance sensors. The development of these sensors is happening far faster than development and deployment of sensors of the underlying thermospheric properties and electric fields. Clearly direct measurement of the drivers is desirable if possible. In addition development of these techniques could fill gaps in missions that may exist.

The research described in this white paper is an excellent opportunity to combine geophysical modeling with advanced mathematical research. Thus, it is well aligned with a recent National Science Foundation research thrust “Collaboration in Mathematical Geosciences (CMG)”, whose purpose is to “enable collaborative research at the intersection of mathematical sciences and geosciences, and to encourage cross-disciplinary education.” Driver estimation as described here is not unique to the ionospheric realm. The approach fits within the category of the “joint inference problem”, where *state* and *parameter* estimation is attempted simultaneously. In this case, *state* refers to the ionospheric electron density, which is well measured. *Parameter* refers to the drivers. As reviewed in Kantas et al., [2009], many approaches to this joint inference problem have been considered, often involving the Ensemble Kalman Filter [Evensen, 2009; Evensen et al., 2007; Evensen, 2007; Chitralekha et al., 2010]. Further study could yield a set of approaches well suited to the application described here that also provide insight into the general mathematical issues applicable to a wide variety of related problems across geophysics and space physics.

2. Conclusions And Recommendations

We have identified a fruitful research direction that combines assimilative modeling techniques with the growing volume of ionospheric and thermospheric data to provide a new modeling capability of great scientific and practical value to the Atmosphere-Ionosphere-Magnetosphere (AIM) community. This capability produces estimates of the underlying physical drivers of ionospheric plasma structure. Incorporating direct measurements of the drivers themselves should be done where possible. This new capability should be developed on an ongoing basis to fill the gap that is likely to exist in measurements of the drivers. On the other hand, measurements of ionospheric structure from GPS and other sources have a long term future of continuous operation. The capability described here will bring benefits to the AIM (AIM=?) community over the long term, deriving maximum scientific information content from the plasma structure measurements.

3. Acknowledgement

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

4. References

- Chitrlekha, S. B., J. Prakash, H. Raghavan, R. B. Gopaluni, and S. L. Shah, "A comparison of simultaneous state and parameter estimation for a continuous fermentor reactor," *Journal of Process Control*, vol. 20, pp. 934--943, 2010.
- Evensen, G., "The ensemble Kalman filter for combined state and parameter estimation," *IEEE Control Systems Magazine*, vol. 29, pp. 83--104, 2009.
- Evensen, G., J. Hove, H. C. Meisingset, E. Reiso, K. S. Seim, and Ø. Espelid, "Using the EnKF for assisted history matching of a North Sea reservoir model," in *Proc. SPE Reservoir Simulation Symposium*, Houston, TX, 2007, SPE 106184.
- Evensen, G., *Data assimilation: The ensemble Kalman filter*, Berlin, Germany, Springer-Verlag, 2007.
- Kantas, N., A. Doucet, S. S. Singh, and J. M. Maciejowski, "An overview of sequential Monte Carlo methods for parameter estimation in general state-space models," in *Proc. 15th IFAC Symposium on System Identification, SYSID 2009*, Saint-Malo, France, 2009, pp. 6--8.
- Luan, X. L., and S. C. Solomon (2008), Meridional winds derived from COSMIC radio occultation measurements, *J. Geophys. Res-Space Phys.*, 113(A8), A08302 doi:10.1029/2008ja013089.
- Pi, X. Q., C. M. Wang, G. A. Hajj, G. Rosen, B. D. Wilson, and G. J. Bailey (2003), Estimation of $E \times B$ drift using a global assimilative ionospheric model: An observation system simulation experiment, *J. Geophys. Res-Space Phys.*, 108(A2), Artn 1075.
- Scherliess, L., D. C. Thompson, and R. W. Schunk (2009), Ionospheric dynamics and drivers obtained from a physics-based data assimilation model, *Radio Science*, 44, Artn Rs0a32, Doi 10.1029/2008rs004068.
- Schunk, R. W., et al. (2004), Global Assimilation of Ionospheric Measurements (GAIM), *Radio Science*, 39(1), Artn Rs1s02, Doi 10.1029/2002rs002794.

Ionospheric forcing from below: Effects of lightning

R. A. Marshall, Boston University

Summary

The ionosphere experiences forcing from above and from below, the latter in the form of gravity waves, tides, and lightning. The effects of lightning on the ionosphere are studied in a strong community worldwide, but are often overlooked in the context of the global ionospheric system. The extension of current single-discharge studies to global effects of lightning are required to quantify its effect on the ionosphere.

Introduction

Lightning forms a fundamental component of the Atmosphere-Ionosphere-Magnetosphere system. Each lightning flash discharges up to 10 GJ of energy, a significant fraction of which is emitted as radio waves. These radio waves reflect from, are absorbed in, and are transmitted through the ionosphere, thus affecting all three A-I-M regions. Some amount of energy is deposited in the ionosphere through collisional heating. A smaller, but not insignificant, fraction is transmitted into the inner magnetosphere, where it propagates as whistler-mode waves. This energy is deposited in the magnetosphere through Landau damping with radiation-belt electrons, but in addition, some of those electrons are precipitated into the atmosphere, thus causing a second, indirect contribution to the ionosphere and atmosphere.

It is unknown at this time how much of the lightning energy is deposited in the various regions, or the significance of this energy on a local and global scale compared to other ionospheric contributions, such as gravity waves, tides, and solar wind forcing. While the quantification of each of these contributions is equally necessary to understanding the global energy balance in the ionosphere, lightning seems to be the most often overlooked, and the state of lightning-ionosphere interactions research is just reaching the point where this estimate can be made.

Outstanding Problems and Approach

An accurate quantitative description of the global lightning-ionosphere interaction will require two separate questions to be addressed and then integrated:

1. A description of the lightning interaction with the ionosphere on a per-stroke basis. Recently, a satellite observation from the ISUAL instrument was used to estimate the direct electron density enhancement in the D-region due to a lightning strike [Mende *et al.*, 2005], which found a doubling of the electron density at 90 km altitude. Recent modeling results [Marshall *et al.*, 2009] have corroborated this estimate. These models require significant improvement,

however, and more observational measurements are needed. The calculations then need to be extended to an estimate of energy deposition. These results need to be evaluated as a function of lightning parameters, such as peak current, rise time, and location.

2. Improved global measurements of lightning and TLE occurrence are required to make global extrapolations. Currently, quantitative, US lightning data can be used with global flash measurements from the LIS and OTD satellite data, which does not provide information on lightning type or intensity. Recently, however, a global ground-based detection network has been established by Vaisala, Inc., known as GLD360. This global, quantitative data set can be used for global extrapolations of lightning-ionosphere interactions. As for TLE occurrence, continuous monitoring of TLEs over a large area is necessary to make reliable, quantitative estimates of their occurrence rates.

The combination of these two measurements will allow global estimates of the energy deposition due to lightning in the ionosphere. These estimates can then be used for diurnal, seasonal, latitudinal, and long term trends.

Requirements

1. Improved lightning-ionosphere interaction models are required to make estimates of the interaction.
2. More measurements of elves are required to compare and provide ground-truth to model estimates. Ground-based observations of elves have recently matured using photometric arrays [Marshall *et al.*, 2008; Newsome *et al.*, 2009], which can estimate the intensity in multiple wavelength bands with the sensitivity to catch all events. This instrument, known as PIPER, is uniquely capable of detecting even the weakest TLEs, while distinguishing sprites from elves by their spatial characteristics. A systematic distribution of these instruments (like a DASI array) should be developed to make continuous measurements of elves over a large region. New satellite instruments are being launched in the next few years in Europe to measure elves, but there are currently no US-based space borne measurements.
3. Use of the global lightning dataset will be required to make global extrapolations of the per-stroke estimates.
4. Communication with the magnetosphere, thermosphere, and atmospheric communities will need to be fostered, so that the different communities understand each others' contributions to the energy balance.

Ground-based Space Weather Instrument Suites

R. A. Marshall, Boston University

Summary

We propose that the geoscience community foster a class of instrumentation that involves multi-instrument ground-based suites, similar to satellite instrumentation. Each location on the ground will host an integrated suite of instruments, which are duplicated at different sites. The integrated instrument suite will simplify development and the distribution of data to scientists.

Introduction

With the notable exception of Cubesats, satellite missions typically encompass a suite of instruments, which are integrated for power, data handling and processing, and telemetry, making data retrieval and analysis, and specifically cross-instrument comparisons, that much simpler. Ground-based instruments, however, are most often single instruments, without any inherent connection to other instruments. High quality data analysis, however, requires that cross-instrument comparisons and correlations be made to properly characterize and contextualize the data. In the past decade, there has been a slow progression towards integrated suites of ground-based instruments; in particular, the AGO stations in Antarctica and the THEMIS ground-based array have begun what should be a developing trend. Those sites contain copies of multiple instruments; however, within each site, they still do not integrate as well as they could. We propose that the ground-based observational community prioritize the development of large suites of integrated instruments, which work together towards a complete characterization of the ionospheric environment.

Our goal here is not to propose or develop such an instrument suite, but to demonstrate a few of the possibilities. A possible high-latitude instrument suite might involve an all-sky camera and a magnetometer, similar to the THEMIS sites, as well as a low-cost spectrograph, photometer system, ELF/VLF receiver, and other passive, low-maintenance instruments.

These instrument suites should be developed in a multi-institutional setting, similar to satellite development, taking advantage of different institutions' strengths. The copies-of-instruments approach drives down development cost, and integration at sites lowers the costs associated with maintenance, computing, and data archival and retrieval. Co-location and precise timing simplifies comparisons of data.

Neutral Winds in the Upper Atmosphere

Carlos Martinis- Boston University

Introduction:

Neutral winds play a key role in different processes in the upper atmosphere that involve electro-dynamics and neutral-plasma coupling. At equatorial and low latitude regions the behavior of winds drives the motion of plasma drifts, via E and F region dynamos. Current neutral wind models do not properly represent the conditions in the upper atmosphere. This is a serious limitation to attempts to obtain a clear picture of the dynamics and electrodynamics of the upper atmosphere. For example, the influence of the behavior of neutral winds in the occurrence of ionospheric irregularities in the equatorial region is known to be very important. These irregularities produce serious Space Weather effects on radio navigation and communication systems.

Measurement Objectives:

Determine neutral winds from ~80 to ~300 km at different local times and latitudes.

The task of measuring neutral winds at such low heights makes potential satellite missions very difficult. Other options can include rocket and/or ground-based measurements:

- * Short-lived satellite mission in circular orbit at low altitudes (~300 km) that can make *in-situ* measurements.
- * Satellite mission in elliptic orbit that would allow make *in-situ* measurements as well as remote sensing measurements, including a Fabry-Perot Interferometer (FPI), Michelson Interferometer, or LIDAR instruments to perform remote observations. If the mission performs only remote sensing measurements (i.e., no need to have the satellite orbit at low heights), its lifetime would be extended.
- * Rocket measurements from several launches at different local times.
- * Installation of ground-based distributed arrays of FPIs (DASI concept) would allow the continuous nighttime determination of winds in the thermosphere (via 630.0 nm emissions) and mesosphere (via 557.7 nm emissions). The DASI concept should not be limited to a single type of instrument. Complementing measurements could be obtained by adding MST radars, all-sky imagers, LIDARs, etc, to regional networks that have already instrumentation deployed (South America, Arctic, North America, Antarctic).

Summary:

Modeling of ITM parameters has achieved a good level of maturity, both in the global (or large) and regional (or small) scale. But neutral winds, one of the main inputs of these models for different processes, are poorly represented. This is due to the lack of

observations of neutral winds between 80 and ~300 km in a global scale. Long term space-based observations, via *in-situ* or remote sensing of the thermosphere (and mesosphere) are difficult to achieve. Satellite missions able to measure E and F region winds would fill gaps existing in our understanding of the dynamics of the upper atmosphere. Distributed arrays of Fabry-Perot interferometers would allow make continuous nighttime measurements covering large areas at a relatively small cost. Rocket missions involving several simultaneous launches could also be beneficial for the understanding of low altitude winds. In summary, upper atmosphere neutral winds can be measured using ground- or space-based observations. This is a task that the community needs to tackle in order to understand the behavior of upper atmosphere processes.

NRC Solar and Space Physics Decadal Survey Concept Paper

Atmosphere-Ionosphere-Magnetosphere Interactions

High Altitude Data Assimilation: Characterizing the Effects of Solar Variability from the Ground to the Thermosphere

John McCormack & Stephen Eckermann
Space Science Division, Naval Research Laboratory, Washington DC

Summary

To improve the observational characterization of the atmospheric response to solar variations, the Sun-Earth research community needs a high-altitude data assimilation system capable of exploiting a broad range of satellite-based and ground-based observations using state-of-the-art assimilation techniques. This system must ultimately be able to produce global 6-hourly gridded meteorological analyses from the ground to the lower thermosphere (~100 km altitude) over a decade or more in order to fully resolve the atmospheric response to both short-term and long-term solar variability.

Introduction

Over the past decade, the vast majority of funding in Sun-Earth studies has been directed towards large space-based observing missions. What has not received as much attention (or resources) is the impact of solar variability on the neutral atmosphere at altitudes up to 100 km. A growing body of evidence now suggests that solar-related variations in the chemical composition and transport throughout the mesosphere and lower thermosphere (MLT) can have a significant impact on the geospace environment. This “forcing from below” is an important consideration for Sun-Earth studies over the coming decade for two reasons. First, improved understanding of the atmospheric response to long-term (i.e. decadal) changes in solar irradiance is crucial for accurate projections of future climate scenarios. Second, detailed knowledge of the atmospheric response to short-term fluctuations (time scales from days to weeks) in solar output is vital for improving predictive models of space weather and its effects on human activity.

To date, our understanding of the atmospheric response to solar variability is based on a combination of observational and modeling results. The observational results come from a rapidly growing body of satellite and ground-based measurements whose spatial and temporal coverage vary widely. The modeling results are based on a number of global atmospheric models with differing levels of complexity. From these two sources, we know that the atmospheric response to solar forcing involves numerous photochemical and dynamical processes over a wide range of altitudes from the surface to the lower thermosphere. However, we are now at a point where solar-climate modeling capabilities have far outpaced the observational capabilities. *What is currently lacking is a comprehensive and self-consistent set of meteorological analyses capable of characterizing the atmospheric response to solar variations extending from the surface to the top of the MLT region.*

To address this deficiency, support is needed to develop and run a global atmospheric data assimilation system (DAS) that can generate synoptic meteorological analyses every 6 hours specifically tailored for solar-atmosphere coupling research. This DAS could be based on an

existing DAS such as the NASA Global Earth Observing System Version 5 (GEOS-5, *Rienecker et al.*, 2008) or the high-altitude version of the Navy Operational Global Atmospheric Prediction System (NOGAPS), known as NOGAPS-ALPHA (Advanced Level Physics-High Altitude, *Eckermann et al.*, 2009). In addition to operational meteorological observations in the 0-50 km altitude range, such a system would also have to make use high-altitude measurements from instruments on NASA research satellites in order to obtain information on the region between 50-100 km. Some examples of current instruments are the Aura Microwave Limb Sounder (MLS) and the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument aboard the Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) satellite. Measurements from these research satellites are currently being assimilated by NOGAPS-ALPHA (Figure 1).

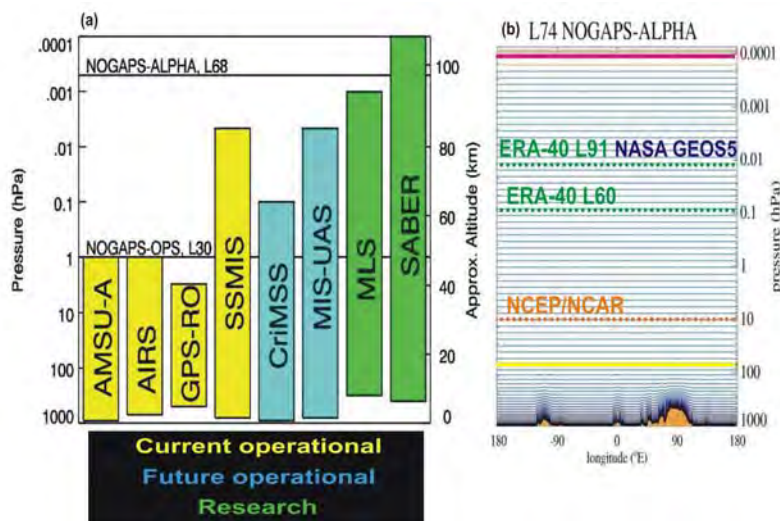


Figure 1. (a) Schematic comparing the top levels of the 30-level operational NOGAPS system and the existing 68-level NOGAPS-ALPHA system. Current operational satellite data measurements assimilated by both systems are in yellow, research satellite data assimilated by NOGAPS-ALPHA in green, future planned instruments in blue; (b) Vertical levels of the proposed L74 NOGAPS-ALPHA assimilation system around 35°N latitude illustrating the hybrid σ -p coordinate that transitions from ground-following to pure pressure levels at 70 hPa (yellow line). Red line indicates L74 top pressure of 9×10^{-5} hPa.

Background

The sun is the primary energy source for the climate system, and therefore changes in solar output will act as a forcing to this system. Over the 11-year solar cycle, observed variations in total solar irradiance (TSI) are $\sim 0.1\%$, while variations in solar spectral irradiance are generally accepted to be 6-8% at ultraviolet (UV) wavelengths, and more than 10% at shorter wavelengths (e.g., *Lean et al.*, 2005). UV irradiance also varies over the 27-day solar rotation period, exhibiting a wavelength dependence similar to the 11-year variations, but with slightly smaller amplitudes. In addition to changes in irradiance, the amount of solar energetic particles (SEP's) reaching the upper atmosphere are another important forcing mechanism.

Impact on the Research Community

A high-altitude DAS based would provide the research community with a unique tool for Sun-Earth studies. This tool could fully exploit information on the atmosphere-mesosphere-ionosphere system from both current and future NASA satellites. Through the use of state-of-the-art data assimilation techniques, the proposed high-altitude DAS would be the best source of ground-to-space meteorological analyses for characterizing the atmospheric response to solar variability. These analyses would aid in the interpretation of increasingly complex model simulations of the physical processes linking solar variability and climate across deep layers of the neutral atmosphere. The following sections discuss this proposed system in greater depth.

The sun emits most of its energy in the visible spectral region, and a large portion of the energy at these wavelengths reaches the Earth's surface without being scattered or absorbed. As a result, changes in TSI are expected to force climate through changes in surface heating, which is sometimes referred to as “bottom-up” forcing. By comparison, the energy emitted by the sun in the UV spectral region is relatively small, and is absorbed primarily in the stratosphere and mesosphere by ozone and molecular oxygen. However, the effects of solar UV variations may produce significant atmospheric climate signals if they are of sufficiently long duration or if nonlinear atmospheric feedback mechanisms are present (e.g., *Gray et al.*, 2010, and references therein).

For example, increased ozone heating via UV absorption at solar maximum warms the equatorial stratopause region (*McCormack and Hood*, 1996; *Frame and Gray*, 2010), which can in turn alter the zonal wind profile through thermal wind balance. Observational studies (e.g., *Hood et al.*, 1993; *Kodera and Kuroda*, 2002, *Frame and Gray*, 2010) show a solar cycle modulation of stratospheric zonal winds that appears first as an increase in the strength of the subtropical jet in the upper stratosphere and lower mesosphere during the solstice periods. Over the following 2-3 months, this initial signal slowly extends poleward and downward into the winter lower stratosphere and troposphere due to interaction with upward-propagating planetary waves. A large and mature body of research now indicates that this “top-down” forcing by quasi-decadal (i.e., 11-year) variations in solar UV output directly impacts a number of key climate variables such as stratospheric ozone (*McCormack and Hood*, 1996, *Soukharev and Hood*, 2006), stratospheric winds and temperatures (*Hood et al.*, 1993, *Kodera et al.*, 2003, *Claud et al.*, 2008, *Frame and Gray*, 2010), major tropospheric circulation patterns such as the Walker circulation and Hadley cells (*Labitzke and van Loon*, 1988, *Haigh et al.*, 2005, *Rind et al.*, 2008), and surface temperatures (*Tourpali et al.*, 2003, *Tung and Camp*, 2008).

Large fluxes of SEP's can also strongly perturb the photochemistry of the polar stratosphere and mesosphere by enhancing the concentrations of odd-hydrogen and odd-nitrogen species that can catalytically destroy ozone. Because these SEP events are more frequent at solar maximum, they may act to reduce ozone amounts at high latitudes by as much as 30% (*Jackman et al.*, 2008). While these events clearly have a short-term (days to months) impact on the middle atmosphere, their role in any long-term climatic changes is currently unclear and merit further study.

Recent findings

While some of the processes relating changes in solar irradiance with changes in the atmosphere are fairly well understood (e.g., the relationship between solar UV and stratospheric ozone), a complete understanding of the underlying physical mechanisms linking solar variability and climate is currently incomplete. This is due to the fact that different wavelengths of solar radiation are absorbed at different altitudes in the climate system, and so a wide range of atmospheric processes are involved from the Earth's surface to the edge of space (~100 km altitude). There has been considerable recent progress in modeling these processes using coupled ocean-atmosphere general circulation models (GCM's) that have a fully-resolved middle atmosphere (the altitude region from 10-100 km) with fully coupled photochemistry (e.g., *Matthes et al.*, 2004; *Matthes et al.*, 2006; *Schmidt et al.*, 2006; *Austin et al.*, 2007; *Marsh et al.*, 2007; *Meehl et al.*, 2009). Quasi-decadal solar UV forcing in these GCM's produces complex climate responses via nonlinear wave-mean flow interactions throughout the entire middle atmosphere due to changes in ozone heating, which can ultimately alter atmospheric circulations near the surface. This “top-down” forcing appears to act in

concert with the “bottom-up” forcing related to changes in TSI to produce observed solar-related climate signals (*Rind et al.*, 2008, *Meehl et al.*, 2009, *Gray et al.*, 2010).

A complicating factor in solar-climate research is that no two solar cycles are identical. In particular, the beginning of solar cycle 24 showed record low levels of solar irradiance since space-based monitoring began over three decades ago. One apparent consequence of this pronounced solar minimum has been the detection of extremely low thermospheric densities (*Solomon et al.*, 2010). Currently, these low densities cannot be explained solely by changes in direct solar heating of the thermosphere, nor can they be accounted for by shrinking of the thermosphere due to increased CO₂ cooling in the lower atmosphere (*Emmert et al.*, 2010). One possibility is that mesospheric densities decreased due to solar-related changes in composition or transport (*Emmert et al.*, 2010). To address this question, improved specification of the atmospheric response to solar variability within the 80-100 km region is needed.

The Need for High Altitude Data Assimilation

Two data sources commonly used for solar-climate studies are the long-term meteorological reanalyses produced by the National Center for Atmospheric Research/National Centers for Environmental Prediction (NCAR/NCEP) and by the European Centre for Medium-Range Weather Forecasting (ECMWF). As Figure 2b shows, the top level of the NCEP/NCAR reanalysis only extends to 10 hPa (~30 km altitude), which is far below the level of maximum ozone heating near the stratopause (50 km), where the effects of solar variability are likely to be largest. The top of the 60-level ECMWF 40-year reanalysis (ERA-40 L60) extends past the stratopause, as does the top of more recent 91-level ECMWF operational analyses. However, neither one of these ECMWF data sets explicitly account for solar variability in their model components, nor do they assimilate any high-altitude (>50 km) ozone or temperature observations such as from NASA or ESA research satellites (*Uppala et al.*, 2005). The existing L68 and proposed L74 NOGAPS-ALPHA systems also have top levels higher than the current GEOS5 system.

In contrast, the high-altitude NOGAPS-ALPHA DAS uses a 74-level (L74) vertical domain extending to the lower thermosphere (Fig. 1b). The shortwave (solar) and longwave (terrestrial) radiative heating- and cooling-rate parameterizations in NOGAPS-ALPHA are better suited to high-altitude data assimilation than either of the schemes used in the NCEP/NCAR or ECMWF systems. For example, NOGAPS-ALPHA employs a 9-band solar radiative heating parameterization that can account for spectrally varying solar irradiance over wavelengths from 125-700 nm plus an infrared cooling parameterization that accounts for non-local thermodynamic equilibrium (LTE) effects (*Eckermann et al.* 2009). Most importantly, NOGAPS-ALPHA’s assimilation of temperature, ozone and water vapor observations in the upper stratosphere and mesosphere provides high-altitude observational information on radiatively significant atmospheric variables that the other reanalyses lack.

Technical Description

This section describes the technical approach for developing state-of-the-art characterization of the neutral atmosphere from 0-100 km. This approach would leverage off of an existing middle atmosphere data assimilation system such as NOGAPS-ALPHA. We begin with a description of the NOGAPS-ALPHA system in its current state and then outline some challenges that must be met in order to extend the system up to 100 km. A key feature of the proposed work is that it builds upon

the framework of an existing state-of-the-art high-altitude DAS that has been fully tested and validated in a variety of recent research applications.

The NOGAPS-ALPHA high-altitude DAS couples the Navy's operational three-dimensional variational (3DVAR) algorithm of *Daley and Barker (2001)* to the NOGAPS-ALPHA global spectral forecast model (GSFM) to produce global synoptic analyses every 6 hours. This is illustrated schematically in **Figure 2**. NAVDAS solves the 3DVAR equation in observation space, i.e.:

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{P}_b \mathbf{H}^T \{ \mathbf{H} \mathbf{P}_b \mathbf{H}^T + \mathbf{R} \}^{-1} [\mathbf{y} - \mathbf{H}(\mathbf{x}_b)]$$

where $\mathbf{x}_a$ is the analysis vector, $\mathbf{x}_b$ is the background vector, $\mathbf{P}_b$ is the background error covariance, $\mathbf{y}$ is the observation vector, $\mathbf{R}$ is the observation error covariance, and the superscript T denotes transpose (see *Hoppel et al., 2008*, for more details). The matrix $\mathbf{H}$ is the Jacobian matrix corresponding to the forward operator $\mathbf{H}(\mathbf{x}_b)$ linearized about the background state vector. The analysis vector consists of the gridded fields of temperature, winds, geopotential height, and humidity. For the applications discussed in this proposal, the analyses are computed using a 6-h update cycle, and $\mathbf{x}_b$ is the 6-hour forecast from the previous update cycle. The innovation $\mathbf{y} - \mathbf{H}(\mathbf{x}_b)$ (also called the observation minus forecast, or O-F) represents the deviation of the forecast from the observations, in observation space. The quantity $\mathbf{x}_a - \mathbf{x}_b$ is the correction vector in model grid space.

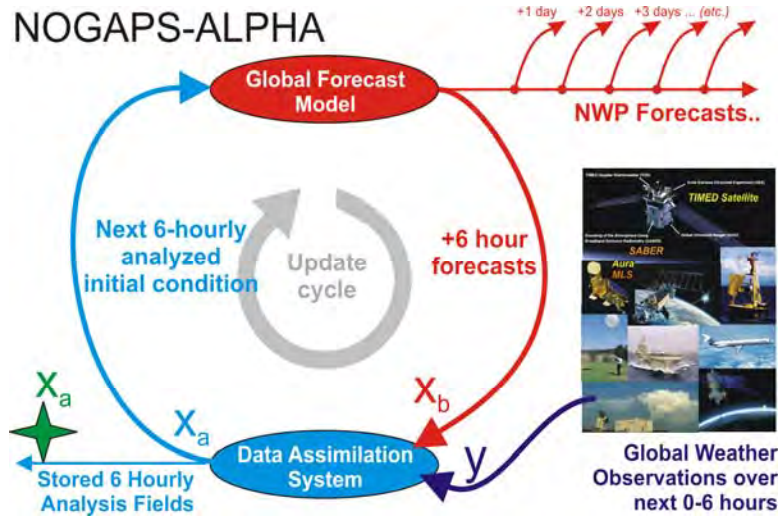


Figure 2. Schematic of the NOGAPS-ALPHA system, where the meteorological analysis ($\mathbf{x}_a$) is updated every 6 hours using observations ($\mathbf{y}$) combined with background fields ($\mathbf{x}_b$) generated with the global forecast model.

Standard output fields include zonal wind, meridional wind, temperature, geopotential height, ozone and water vapor mixing ratios, horizontal divergence, and vertical velocity. The initial version of the NOGAPS-ALPHA DAS is described by *Hoppel et al. (2008)*. The current T79L68 production configuration is described in detail by *Eckermann et al. (2009)*, and forms the basis from which the proposed new system will emerge. This section focuses briefly on some of the salient details here, with full details on each aspect found in *Eckermann et al. (2009)*.

In addition to standard archived operational meteorological observations, NOGAPS-ALPHA assimilates version 2.2 retrievals of temperature and water vapor from the Microwave Limb Sounder (MLS) on NASA's Aura satellite up to ~0.002 hPa (*Lambert et al. 2007; Schwartz et al., 2008*) and MLS ozone up to ~0.02 hPa (*Jiang et al., 2007*). The DAS also assimilates version 1.07 temperature retrievals from the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument on NASA's TIMED satellite (*Remsberg et al., 2008*), with both SABER and MLS temperatures bias corrected prior to assimilation using the method described in section 2.2 of *Eckermann et al. (2009)*. MLS ozone is assimilated only up to 0.02 hPa following the guidance of *Jiang et al. (2007)*, with daytime values only assimilated from

1-0.02 hPa due to the absence of diurnal ozone photochemistry in the GSFM. An existing parameterization of diurnally varying ozone photochemistry recently developed at NRL could be added to the NOGAPS-ALPHA GSFM, thereby allowing assimilation of high-altitude ozone observations during both night and day.

The NOGAPS-ALPHA GSFM uses an Eulerian hydrostatic dynamical core that is spectral in the horizontal and uses energy- and angular momentum-conserving finite differencing in the vertical based on a generalized vertical coordinate. The production DAS configuration of *Eckermann et al.* (2009) uses triangular truncation at wave number 79 (T79). The resulting T79L68 fields have an intrinsic grid point resolution of $\sim 2.25^\circ$ horizontally and ~ 2 km vertically in the middle atmosphere. For the proposed new configuration, T79 horizontal resolution would be more than adequate, but it will be necessary to extend the GSFM upwards by using 74 hybrid σ -p levels (L74) with a top at 9×10^{-5} hPa, retaining the ~ 2 km vertical resolution from ~ 15 -100 km. Shortwave heating and longwave cooling rates due to ozone, CO₂, and water vapor from the surface to the lower mesosphere are currently computed using the method of *Chou et al.* (2001) and *Chou and Suarez* (2002). Additional CO₂ cooling rate calculations are performed using the parameterization of *Fomichev et al.* (1998) for the upper mesospheric region where the effects of non-local thermodynamic equilibrium begin to dominate, with the two cooling-rate calculations blended as in *Eckermann et al.* (2009). Both ozone and H₂O are prognostic variables in the NOGAPS-ALPHA GSFM whose time tendencies in the middle atmosphere are governed by fast parameterizations of gas-phase odd oxygen photochemistry for ozone (*McCormack et al.*, 2006) and methane oxidation and photolytic loss for H₂O (*McCormack et al.*, 2008). Tropospheric H₂O tendencies are also governed by NOGAPS cloud, convection and mixing parameterizations.

Middle atmospheric wind measurement from space remains technologically challenging, and there

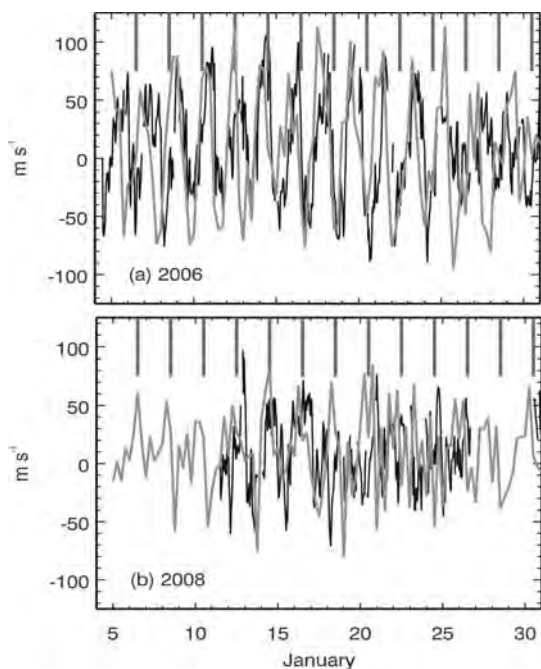


Figure 3. Time series of meridional winds at ~ 88 km over Adelaide (35°S , 138°E) for (a) January 2006 and (b) January 2008 from NOGAPS-ALPHA (black curve) and independent MF radar (gray curve). Vertical lines along x-axis indicate noon local time every 2 days.

are as yet no operational and few research satellite sensors measuring winds at these altitudes. Thus, in common with all other operational DASs, NOGAPS-ALPHA does not at present assimilate wind measurements in the upper stratosphere and mesosphere. Instead, analyzed winds at these altitudes are constrained by the forecast model background and by assimilated temperature measurements via a meridionally-weighted gradient wind constraint built into the off diagonal wind-temperature and wind-geopotential components of the DAS's background error covariance matrix. Both *Eckermann et al.* (2009) and *Stevens et al.* (2010) have shown that the resulting analyzed mesospheric winds in June 2007 compare well with independent meteor wind measurements from ground-based radars at 54°N and 69°N , capturing in particular the amplitude and phase of the dominant tidal wind variations at these locations. *McCormack et al.* (2010) also showed that analyzed NOGAPS-ALPHA meridional winds at ~ 88 km in January 2006 and 2008 accurately captured the amplitude and phase of the dominant quasi-2 day variations seen in independent medium frequency

(MF) radar wind measurements over Adelaide (35°S, 138°E) at these times, as shown in **Figure 3**.

Despite this encouraging agreement in the limited comparisons between NOGAPS-ALPHA mesospheric DAS winds and independent observations cited above, high-altitude DAS winds require more dedicated research. While *Lieberman* (1999) and *Oberheide et al.* (2000) have reported that gradient wind balance often holds in the extratropical mesosphere and lower thermosphere (MLT), it breaks down in the equatorial MLT due to large-amplitude diurnal tidal influences (*Lieberman*, 1999; *McLandress et al.*, 2006) as well as in the extratropical summer MLT in some GCMs due to the effects of nonzonal gravity wave drag (*McLandress et al.* 2006). Large-amplitude diurnal and semidiurnal tides present theoretical challenges for a DAS, both in terms of the 6 hourly analysis window and the DAS algorithms that often assume dominant vortical and limited divergent energy in the atmosphere (see *Eckermann et al.*, 2009). An ideal approach would assimilate MLT wind observations, either from ground based instruments such as MF or meteor radars, or from a space-based instrument such as the TIMED Doppler Interferometer (TIDI, *Killeen et al.*, 2006, *Niciejewski et al.*, 2006).

Relevance

The concepts outlined in this paper would make immediate and high-impact contributions to the Panel theme of atmosphere-ionosphere-magnetosphere interactions by properly accounting for the atmospheric “forcing from below” of the geospace environment. **Specifically, the development of a high-altitude data assimilation system that extends from the surface to 100 km altitude would address a serious deficiency in our collective understanding of how solar variability impacts the composition and dynamics of the neutral atmosphere over both short (weeks to months) and long (years to decades) time scales.**

An improved characterization of the atmospheric state would greatly aid in developing a better understanding of the underlying physical processes linking solar variability to the neutral atmosphere. **This in turn would aid in both our ability to understand and predict the impacts of solar variability on space weather (over short time scales) and climate (over long time scales), both of which would aid in future mission planning and policy making.**

Compared to large space-based observation missions, the costs of developing and running a high-altitude DAS are extremely modest (< \$1M per year). Such a system would be capable of exploiting information from a broad range current and upcoming satellite-based Earth observing missions, and would thus complement existing observational programs. In addition, this atmospheric DAS can, and should, be coupled with more extensive thermosphere-ionosphere-magnetosphere systems in order to provide a seamless characterization of the space weather environment out to 500 km.

Finally, development of a high-altitude DAS for Sun-Earth studies based an existing system such as NOGAPS-ALPHA would provide a high degree of readiness and low degree of risk.

References

- Austin, J., L. L. Hood, and B. E. Soukharev (2007), Solar cycle variations of stratospheric ozone and temperature in simulations of a coupled chemistry-climate model, *Atmos. Chem. Phys.*, **7**, 1693-1706.
- Chou, M.-D., *et al.* (2001), A thermal infrared radiation parameterization for atmospheric studies, NASA Tech. Mem. 104606, 19, 65pp.
- Chou, M.-D., and M. J. Suarez (2002), A solar radiation parameterization for atmospheric studies, NASA Tech. Mem. 10460, 15, 52pp.
- Claud, C., C. Cagnazzo, and P. Keckhut (2008), The effect of the 11-year solar cycle on the temperature in the lower stratosphere, *J. Atmos. Sol. Terr. Phys.*, **70**, 2031-2040.
- Daley, R. and E. Barker (2001), NAVDAS: Formulation and diagnostics, *Mon. Wea. Rev.*, **129**, 869-883.
- Eckermann, S. D., *et al.* (2009), High-altitude data assimilation system experiments for the Northern Hemisphere summer mesosphere season of 2007, *J. Atmos. Sol. Terr. Phys.*, **71**, 531-551, doi:10.1016/j.jastp.2008.09.036.
- Emmert, J. T., J. L. Lean, and J. M. Picone (2010), Record-low thermospheric density during the 2008 solar minimum, *Geophys. Res. Lett.*, **37**, L12102, doi:10.1029/2010GL043671.
- Fomichev, V. I., *et al.* (1998), Matrix parameterization of the 15 mm band CO₂ cooling in the middle and upper atmosphere for variable CO₂ concentration, *J. Geophys. Res.*, **103**, 11,505 - 11,528.
- Frame, T. H. A., and L. J. Gray (2010), The 11-year solar cycle in ERA-40 Data: An update to 2008, *J. Climate*, **23**, 2113-2222.
- Garcia, R. R., *et al.* (2007): Simulation of secular trends in the middle atmosphere, *J. Geophys. Res.*, **112**, D09301, doi:10.1029/2006JD007485.
- Gray, L. J., *et al.* (2010), Solar influences on climate, *Rev. Geophys.*, in press.
- Haigh, J. D., M. Blackburn, and R. Day (2005), The response of tropospheric circulation to perturbations in lower-stratospheric temperature, *J. Clim.*, **18**, 3672-3685.
- Hood, L. L., J. L. Jirikowic, and J. P. McCormack (1993), Quasi-decadal variability of the stratosphere: Influence of long-term solar ultraviolet variations, *J. Atmos. Sci.*, **50**, 24, 3941-3958.
- Hood, L. L. (2004), Effects of solar UV variability on the stratosphere, in *Solar Variability and its Effects on Climate*, Geophysical Monograph Series No. 141, J. Pap and P. Fox, eds., AGU, Washington, DC, 366pp.
- Hoppel, K. W., N. L. Baker, L. Coy, S. D. Eckermann, J. P. McCormack, G. Nedoluha, G., and D. E. Siskind (2008), Assimilation of stratospheric and mesospheric temperatures from MLS and SABER in a global NWP model, *Atmos. Chem. Phys.*, **8**, 6103-6116.
- Jackman, C. H., *et al.* (2008): Short- and medium-term atmospheric constituent effects of very large solar proton events, *Atmos. Chem. Phys.*, **8**, 765-785.
- Jiang, Y. B., *et al.* (2007), Validation of Aura Microwave Limb Sounder Ozone by ozonesonde and lidar measurements, *J. Geophys. Res.*, **112**, D24S34, doi:10.1029/2007JD008776.
- Killeen, T. L., Q. Wu, S. C. Solomon, D. A. Ortland, W. R. Skinner, R. J. Niciejewski, and D. A. Gell (2006), TIMED Doppler Interferometer: Overview and recent results, *J. Geophys. Res.*, **111**, A10S01, doi:10.1029/2005JA011484.
- Kodera, K., and Y. Kuroda (2002), Dynamical response to the solar cycle, *J. Geophys. Res.*, **107** (D24), 4749, doi: 10.1029/2002JD002224.
- Kodera, K., *et al.* (2003), Solar impact on the lower mesospheric subtropical jet: A comparative study with general circulation model simulations, *Geophys. Res. Lett.*, **30**, 1315, doi:10.1029/2002GL016124.

- Labitzke, K., and H. van Loon (1988), Associations between the 11-year solar cycle, the QBO, and the atmosphere. Part I: the troposphere and stratosphere in the northern hemisphere winter, *J. Atmos. Terr. Phys.*, 3, 197-206.
- Lambert, A., et al. (2007), Validation of the Aura Microwave Limb Sounder middle atmospheric water vapor and nitrous oxide measurements, *J. Geophys. Res.*, 112, D24S36, doi:10.1029/2007JD008724.
- Lean, J., G. Rottman, J. Harder, and G. Kopp (2005), SORCE contributions to new understanding of global change and solar variability, *Sol. Phys.*, 230, 27–53, doi:10.1007/s11207-005-1527-2.
- Lieberman, R. S. (1999), The gradient wind in the mesosphere and lower thermosphere, *Earth Planets Space*, 51, 751-761.
- Marsh, D. R., R. R. Garcia, D. E. Kinnison, B. A. Boville, F. Sassi, S. C. Solomon, and K. Matthes (2007), Modeling the whole atmosphere response to solar cycle changes in radiative and geomagnetic forcing, *J. Geophys. Res.*, 112, D23306, doi:10.1029/2006JD008306.
- Matthes, K., et al. (2004), Improved 11-year solar signal in the Freie Universitaat Berlin Climate Middle Atmosphere Model (FUB-CMAM), *J. Geophys. Res.*, 109, D06101, doi:10.1029/2003JD004012.
- Matthes, K., et al. (2006), Transfer of the solar signal from the stratosphere to the troposphere: Northern winter, *J. Geophys. Res.*, 111, D06108, doi:10.1029/2005JD006283.
- McCormack, J. P., and L. L. Hood (1996), Apparent solar cycle variations of upper stratospheric O₃ and temperature: Latitude and seasonal dependences, *J. Geophys. Res.*, 101, 20933-20944.
- McCormack, J. P., et al., CHEM2D-OPP (2006): A new linearized gas-phase ozone photochemistry parameterization for high-altitude NWP and climate models, *Atmos. Chem. Phys.*, 6, 4943-4972.
- McCormack, J. P., D. E. Siskind, and L. L. Hood (2007), Solar-QBO interaction and its impact on stratospheric ozone in a zonally averaged photochemical transport model of the middle atmosphere, *J. Geophys. Res.*, 112, D16109, doi:10.1029/2006JD008369.
- McCormack, J. P., K. W. Hoppel, and D. E. Siskind (2008): Parameterization of middle atmospheric water vapor photochemistry for high-altitude NWP and data assimilation, *Atmos. Chem. Phys.*, 8, 7519-7532.
- McCormack, J. P., S. D. Eckermann, K. W. Hoppel, and R. A. Vincent (2010), Amplification of the quasi-two day wave through non-linear interaction with the migrating diurnal tide, *Geophys. Res. Lett.*, 37, L16810, doi:10.1029/2010GL043906.
- McLandress C., W. E. Ward, V. I. Fomichev, K. Semeniuk, S. R. Beagley, N. A. McFarlane, T. G. Shepherd (2006), Large-scale dynamics of the mesosphere and lower thermosphere: An analysis using the extended Canadian Middle Atmosphere Model, *J. Geophys. Res.*, 111, D17111, doi:10.1029/2005JD006776.
- Meehl, G. A., et al., (2009), *Science*, 325, 1114 – 1118, doi:10.1126/science.1172872.
- Niciejewski, R., et al. (2006), TIMED Doppler Interferometer on the Thermosphere Ionosphere Mesosphere Energetics and Dynamics satellite: Data product overview, *J. Geophys. Res.*, 111, A11S90, doi:10.1029/2005JA011513.
- Oberheide, J., M. E. Hagan, W. E. Ward, M. Riese, and D. Offermann (2000), Modeling the diurnal tide for the Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) 1 time period, *J. Geophys. Res.*, 105(A11), 24, 917–24, 929.
- Rienecker, M.M., M.J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R.D. Koster, R. Gelaro, I. Stajner, and J.E. Nielsen (2008), The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. *Technical Report Series on Global Modeling and Data Assimilation*, 27.
- Rind, D., J. Lean, J. Lerner, P. Lonergan, and A. Leboissitier (2008), Exploring the stratospheric/tropospheric response to solar forcing, *J. Geophys. Res.*, 113, D24103, doi:10.1029/2008JD010114.

- Remsberg, E. E., et al. (2008), Assessment of the quality of the Version 1.07 temperature-versus-pressure profiles of the middle atmosphere from TIMED/SABER, *J. Geophys. Res.*, *113*, D17101, doi:10.1029/2008JD010013.
- Schmidt, H., and G. P. Brasseur (2006), The response of the middle atmosphere to solar cycle forcing in the Hamburg model of the neutral and ionized atmosphere, *Space Science Reviews*, *125*, 345-356.
- Schwartz, M. J., et al. (2008), Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements, *J. Geophys. Res.*, *113*, D15S11, doi:10.1029/2007JD008783.
- Solomon, S. C., T. N. Woods, L. V. Didkovsky, J. T. Emmert, and L. Qian (2010), Anomalously low solar extreme-ultraviolet irradiance and thermospheric density during solar minimum, *Geophys. Res. Lett.*, *37*, L16103, doi:10.1029/2010GL044468.
- Soukharev, B. E., and L. L. Hood (2006), The solar cycle variation of stratospheric ozone: Multiple linear regression analysis of long-term satellite data sets and comparisons with models, *J. Geophys. Res.*, *111*, D20314, doi:10.1029/2006JD007107.
- Stevens, M. H., et al. (2010), Tidally induced variations of PMC altitudes and ice water content using a data assimilation system, *J. Geophys. Res.*, in press.
- Tung, K. K. and C. D. Camp (2008), Solar cycle warming at the Earth's surface in NCEP and ERA-40 data: A linear discriminant analysis, *J. Geophys. Res.*, *113*, D05114, doi:10.1029/2007JD009164.
- Tourpali, K., C. J. E. Schuurmans, R. van Dorland, B. Steil, and C. Brühl (2003), Stratospheric and tropospheric response to enhanced solar UV radiation: A model study, *Geophys. Res. Lett.*, *30*, 1231, doi:10.1029/2002GL016650.
- Uppala, S. M., et al. (2005), The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, *131*, 2961–3012.

The Importance of Thermospheric Winds for Ionospheric Modeling

A White Paper for the Decadal Survey of Solar and Space Physics

Sarah E. McDonald, Douglas P. Drob, John Emmert, Christoph Englert, David Siskind
Space Science Division, Naval Research Laboratory, Washington, DC

Joseph Huba, Jonathan Krall
Plasma Physics Division, Naval Research Laboratory, Washington, DC

Sunanda Basu
Institute for Scientific Research, Boston College, Chestnut Hill, MA

Introduction and Motivation

One of the most fundamental Heliophysical processes is the coupling between charged and neutral particles. In the presence of a magnetic field, drag forces between the charged and neutral gases produce complex electrodynamic interactions. Understanding these interactions was listed as one of the three key research focus areas in the recent Heliophysics Roadmap for enabling space environmental prediction. In the terrestrial environment, these ion-neutral interactions are most vividly manifest in the low latitude ionosphere. Specifically variations in the neutral wind can drive a complex system of ionospheric currents and electric fields, which in turn profoundly influence the structure and composition of the ionosphere. Recent observations have suggested new and previously unknown ionospheric morphologies, which have been linked to neutral wind variability. State-of-the-art coupled thermosphere-ionosphere-electrodynamics models make specific predictions for ion-neutral coupling based upon this variability; however, these theories remain untested by appropriate global observations. A key stumbling block is the lack of neutral winds with sufficient altitudinal and horizontal resolution. Here we present some specific examples where neutral wind measurements are needed to enable reliable ionospheric modeling.

The Wind Dynamo

Winds in the Earth's thermosphere (75 to 500 km) are driven primarily by pressure gradients resulting from temperature differences, with the absorption of solar EUV radiation being the dominant source of heating. Secondary sources of heating arise from collisions between the atmospheric constituents and precipitating energetic auroral particles, and friction resulting from the interaction of neutral gas with magnetospherically driven ion flows at high latitudes. Deposition of energy from the lower atmosphere in the form of upward propagating planetary waves, tides and gravity waves is also important, especially in the lower thermosphere (below 200 km).

The Earth's global wind system contributes significantly to the morphology and variability of the ionosphere. Thermospheric (or neutral) winds blow the ionospheric plasma across the magnetic field, resulting in a separation of charge similar to a dynamo. In the E-region ionosphere (90 – 140 km) the ion-neutral collision frequency is so high that the neutral wind can drag the ions with it, whereas the electrons are constrained to gyrate along the magnetic field and drift transverse to the electric field. Differences in the ion and electron velocities, and small differences in their concentrations are responsible for generating currents and polarization

electric fields in this highly conducting region of the ionosphere. Such electric fields arising from winds in the thermosphere are called dynamo electric fields. Due to the high conductivity along the magnetic field, the E-region dynamo fields map along the magnetic field lines to the F-region (140 – 500 km) ionosphere. Though the E-region dynamo dominates during the day, ion-neutral interactions in the F-region are responsible for an F-region dynamo that becomes important at night when the E-region conductivity is low.

A key feature of the E-region dynamo is the generation of a zonal (east-west) electric field that drives a vertical $\mathbf{E} \times \mathbf{B}$ drift of F-region plasma at the equator, which is one of the most important processes controlling the large-scale distribution of plasma. To first order, the neutral winds blow from the dayside to the nightside leading to upward plasma drift during the day and downward drift at night. The upward plasma drift at the magnetic equator in combination with ambipolar diffusion along the field lines results in regions of enhanced plasma density on either side of the magnetic equator, referred to as the Equatorial Ionization Anomaly (EIA), an important feature of the low-latitude ionosphere.

A major achievement in the development of global modeling of the thermosphere-ionospheric system in the past two decades has been the incorporation of a self-consistent electrodynamic model. This was first achieved with the Thermosphere Ionosphere Electrodynamic General Circulation Model (TIEGCM) (Richmond et al., 1992). Using TIEGCM, Fesen et al. (2000) was able to successfully model the key features of the $\mathbf{E} \times \mathbf{B}$ drift. However, there is still much work to be accomplished in modeling the observed day-to-day, longitudinal and seasonal variations of the ionospheric electrodynamics. In fact, there have been a number of new observations of ionospheric phenomena that suggest that the ionosphere-thermosphere (IT) system is much more interesting and complex than previously thought. Longitudinal variation in the morphology of the low-latitude ionosphere has been linked to non-migrating tides of tropospheric origin (Hagan and Forbes, 2003) that drive the E-region dynamo, thus affecting the F-region ionosphere (Immel et al., 2006). Large pre-dawn density depletions in the equatorial ionosphere have been observed below ~550 km (de la Beaujardiere, 2009) and have been attributed to zonal winds (Huba et al., 2010). Chau et al. (2009) found a unique daytime pattern in the vertical $\mathbf{E} \times \mathbf{B}$ drifts over Jicamarca associated with a largescale meteorological disturbance known as a sudden stratospheric warming (SSW) event in January 2008.

The largest hurdle in understanding these phenomena is knowledge of the winds that drive the E- and F-region dynamos. Below, we highlight four important ionospheric phenomena driven by thermospheric winds that require additional modeling in order to be fully understood. We then summarize the thermospheric winds measurements that have been made to date and highlight a new wind instrument concept.

Pre-reversal Enhancement of the Vertical Plasma Drift

At dusk, when the E-region density decreases, the F-region dynamo becomes more significant. The complex interaction of the E- and F-region dynamos causes an enhancement of the eastward electric field (and upward $\mathbf{E} \times \mathbf{B}$ drift) prior to turning westward. Because of the large upward drifts associated with this prereversal enhancement (PRE) of the electric field, it is a significant factor in the development of ionospheric irregularities (or depletions) that can impact communication and navigation systems. A large PRE can lift the F-region ionosphere to high

altitudes, where recombination is slow. At the same time, the E-region ionosphere is quickly depleted in the absence of photoproduction, which sets up the conditions for a gravitational Rayleigh-Taylor instability to form. Small perturbations in the bottomside density can lead to growth of the instability that results in plasma irregularities (i.e. Equatorial Spread F) and the formation of bubbles, which are structures of depleted density that rise up through the F-region ionosphere. Thus, better understanding of the conditions that lead to the development of the PRE will result in improved irregularity forecasting systems.

A number of theories and models have been developed to explain the PRE and are summarized by Eccles (1998). In each case, the main driver of the PRE is the zonal (eastward) F-region neutral wind, but modeling work by Millward et al. (2001) indicates that the E-region neutral winds can modulate the amplitude of the PRE. There are now various ground-and space-based measurements of the F-region vertical plasma drift that show substantial longitudinal and seasonal variations. For example, Fejer et al. (2008) used ROCSAT-1 drift measurements to develop an empirical model that captures the longitudinal and seasonal variation of the plasma drifts. Figure 1 shows the longitudinal and seasonal variation of the PRE. When such drift models are used in ionospheric models, resulting electron densities are often in good agreement with observations (e.g. Fang et al. 2009). However, models that self-consistently calculate the electric fields from neutral wind models have yet to fully capture the observed variations in the vertical plasma drift. Higher resolution wind models at E- and F-region altitudes are needed to fully understand the drivers of the variation in the PRE as well as to validate the ionosphere models.

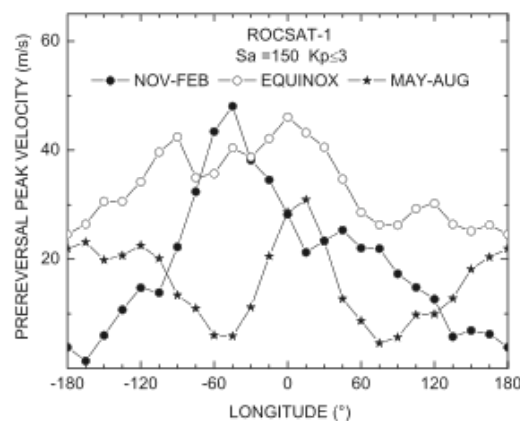


Figure 1. Peak velocity of the PRE as a function of longitude and season. From Fejer et al. (2008).

Ionospheric Irregularity Formation

The strength of the PRE, due to competing dynamo action of winds in the E- and F-region ionosphere, is an important factor in the development of ionospheric irregularities. But the thermospheric winds also play other important roles in the formation and/or suppression of irregularities. Gravity waves have long been suspected to play a role in providing the necessary perturbation to the bottomside of the unstable layer in order to set off an instability (e.g. Fritts et

al., 2009). However, Kudeki et al. (2007) have presented experimental and modeling evidence that F-region eastward neutral winds at sunset play the dominant role in triggering equatorial spread-F. Maruyama and Matuura (1984) predicted that asymmetric density distributions along a flux tube, resulting from interhemispheric F-region winds, can suppress the growth of the Rayleigh-Taylor instability. To test this theory, Krall et al. (2009) used an ionospheric model of ESF (SAMI3/ESF) to show that a constant meridional wind of 60 m s^{-1} stabilizes equatorial spread-F.

Dawn Depletions

Large-scale, dawn density depletions have been recently observed in the equatorial ionosphere during the recent extreme solar minimum conditions (de la Beaujardiere et al., 2009). Simulation results using SAMI3 (Huba et al., 2010) show that the observed depletions can be explained by a post-midnight enhancement of the upward $\mathbf{E} \times \mathbf{B}$ drift as result of strong eastward winds between 150 and 250 km. Huba et al. (2010) showed that SAMI3, which self-consistently solves for the electric fields, yields two different solutions when driven by two different empirical models of the neutral winds: HWM93 (Hedin et al., 1996) and HWM07 (Drob et al., 2008). As shown in Figure 2, a dawn density depletion forms over South American when the HWM07 wind model is used, but that this depletion is not present with the HWM93 model. A comparison of the wind models suggests that the zonal neutral wind is responsible for the depletion. This study highlights the sensitivity of the ionosphere to different wind models, and that the ionosphere is especially sensitive to winds in the E-region.

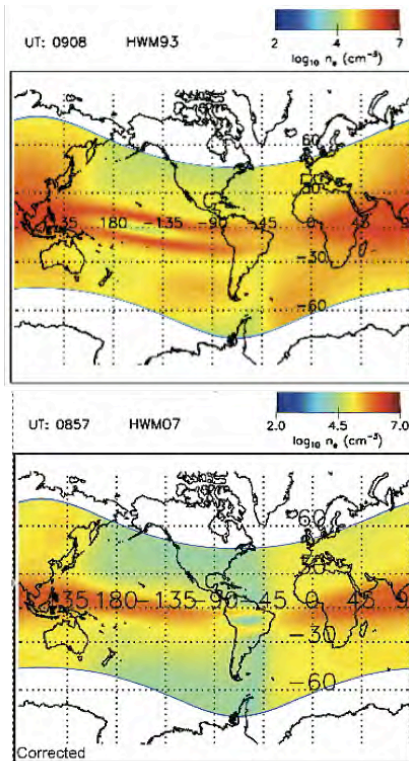


Figure 2. SAMI3 electron density at 0900 UT as a function of geographic latitude and longitude at an altitude of 308 km using (top) HWM93 and (bottom) HWM07. The HWM07 winds produces a density depletion at the magnetic equator over South America. From Huba et al. (2010).

Ionospheric Variability due to Planetary Waves and Tides

Longitudinal variation of the low-latitude ionosphere can in part be attributed to atmospheric nonmigrating tides generated in the lower atmosphere (Immel et al., 2006). Modeling studies have shown that tidal winds can reach significant amplitudes, especially under solar minimum conditions, and that they are significant enough at E-region altitudes to influence the zonal electric fields, and hence the electron densities of the low-latitude F-region ionosphere (Hagan et al., 2007). Recently, Liu et al. (2010) performed numerical simulations to show that planetary waves can cause variability in migrating and nonmigrating tides through nonlinear interactions. The largest changes were found at 100 to 150 km and extended into the upper thermosphere. Their results showed a complex pattern of neutral wind perturbations as a function of time, latitude and longitude.

Wind Measurements

Our understanding of thermospheric winds has historically been limited by a lack of comprehensive measurements. Although tremendous advances in characterizing lower thermospheric (90–120 km) wind patterns have been made as a result of wind measurements by NASA's Upper Atmosphere Research Satellite (UARS) and Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) missions, winds patterns in the middle thermosphere (120–200 km) are still poorly understood. The Wind Imaging Interferometer (WINDII) on UARS is the only major source of wind measurements in this altitude region (Figure 3). While the WINDII data have afforded significant improvements in our understanding of thermospheric wind patterns, this dataset cannot provide an adequate climatological description of dynamo-generating winds. Between 120 and 200 km the data are limited to daytime only, and the 200–300 km nighttime data are limited to latitudes below 45° and only cover late 1991 to 1994. The UARS mission was designed to study the stratosphere, mesosphere, and lower thermosphere (Reber et al., 1993); the upper thermospheric wind measurements were a supplementary benefit of WINDII, the design of which is optimized primarily for the lower-altitude measurements (Shepherd et al., 1993). Offsets and artifacts in the middle and upper thermospheric WINDII profiles, on the order of 50 m/s (Emmert et al., 2002) impede accurate characterization of average winds.

Archival thermospheric wind measurements as a whole suffer from several key gaps. First, very little is known about nighttime winds between 120 and 200 km. An airglow emission suitable for measurement of Doppler shifts has not been identified, so the only wind data in this region comes from in rocket-borne in situ measurements such as chemical releases. These highly accurate measurements have provided enormously valuable insight into small-scale thermospheric dynamics, but their limited spatial and temporal coverage (about 500 profiles since 1998 [Larsen and Fesen, 2009]) preclude them from contributing significantly to knowledge of global-scale patterns. Another important gap is the availability of wind data during severe geomagnetic storms ($K_p \geq 6$). This is partly a result of the relative rarity of such events; only about 2% of wind measurements have been taken under these conditions. Consequently, our empirical understanding of how winds respond to magnetospheric energy input is very coarse (Emmert et al., 2008), particularly with respect to their development and recovery. To address this deficit, it is essential that we continue building up the wind database for at least several more

solar cycles. Finally, because thermospheric wind patterns are so heavily influenced by upward propagating tides that exhibit significant day-to-day variability, for space weather applications it is critical that winds be continuously monitored at multiple local times. In the upper thermosphere, there is currently very little space-based monitoring of winds (TIMED/TIDI began measuring Doppler shifts of 630.0 nm airglow in late 2008), and contemporary wind measurement coverage is far from sufficient to specify and predict the ionosphere-thermosphere system.

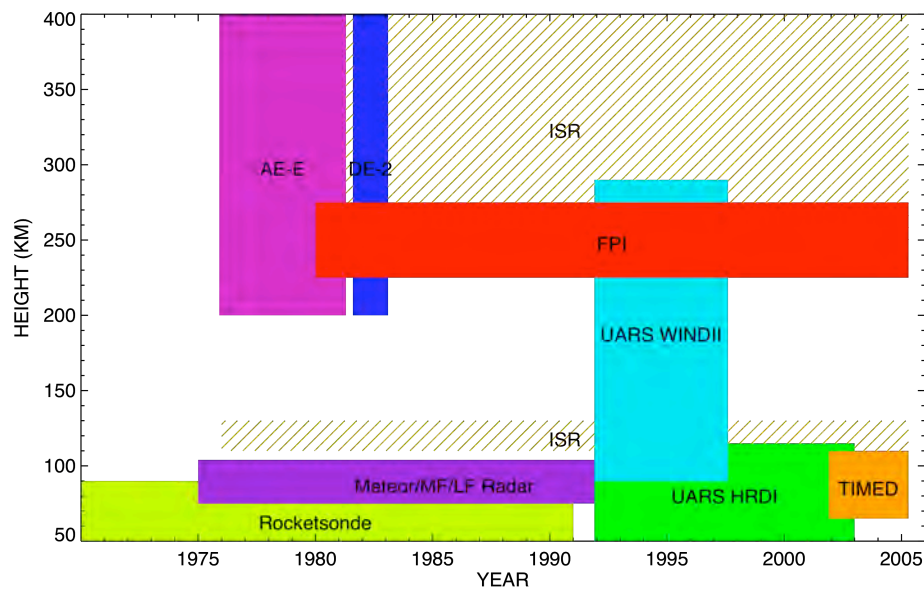


Figure 3. Altitude coverage of major upper atmospheric wind datasets over the past four decades. Most of the datasets shown here are additionally limited in local time and/or latitude. For example the upper thermospheric WINDII measurements are daytime only; ground-based FPI measurements are nighttime only, and the upper thermospheric ISR measurements consist of inferred meridional winds at midlatitudes.

A New Wind Instrument Concept

To provide sufficient input to the ionospheric models, global thermospheric wind measurements with significant spatial resolution are required. Wind profiles encompassing the E- and F-region altitudes are critical. A prime candidate for measuring the thermospheric winds is the Doppler Asymmetric Spatial Heterodyne (DASH) interferometer (Englert et al., 2007). DASH would be capable of measuring thermospheric wind and temperature profiles from low Earth orbit. For the thermosphere, typical measurements could be conducted between 90-300 km altitude with up to 5 km altitude resolution by observing the oxygen red and green lines. With horizontal resolutions of a few hundred kilometers, one can achieve a wind speed precision of less than 5 m/s and temperature precision of less than 5K, depending on the source brightness and altitude resolution. The DASH instrument would also be small and light enough (< 35 kg) to be included even on small missions.

Summary

We have highlighted several important phenomena that impact Earth's ionosphere through the effects of the thermospheric neutral winds. Ionospheric models have reached a level of maturity where improved wind models and/or measurements are needed in order to both validate the models and gain further insight into the physical processes that are being simulated. Thermospheric winds vary considerably with local time, longitude and altitude in the lower thermosphere and have a dramatic effect on the ionosphere. It is therefore critical to promote instruments and new techniques that measure winds with sufficient resolution as well as to develop improved wind models. As a result, our space environment modeling and forecasting capabilities will be greatly enhanced.

References

- Chau, J.L., B.G. Fejer, and L.P. Goncharenko, Quiet variability of equatorial $E \times B$ drifts during a sudden stratospheric warming event, *Geophys. Res. Lett.* 36, L05101, doi:10.1029/2008GL036785, 2009.
- de La Beaujardiere, O., et al., C/NOFS observations of deep plasma depletions at dawn, *Geophys. Res. Lett.*, 36, L00C06, doi:10.1029/2008JA013668.
- Drob, D. P., et al. (2008), An empirical model of the Earth's horizontal wind fields: HWM07, *J. Geophys. Res.*, 113, A12304, doi:10.1029/2008JA013668.
- Emmert, J. T., B. G. Fejer, G. G. Shepherd, and B. H. Solheim (2002), Altitude dependence of middle and low-latitude daytime thermospheric disturbance winds measured by WINDII, *J. Geophys. Res.*, 107, doi:10.1029/2002JA009646.
- Emmert, J. T., D. P. Drob, G. G. Shepherd, G. Hernandez, M. J. Jarvis, J. W. Meriwether, R. J. Niciejewski, D. P. Sipler, and C. A. Tepley (2008), DWM07 global empirical model of upper thermospheric storm-induced disturbance winds, *J. Geophys. Res.*, 113, A11319, doi:10.1029/2008JA013541.
- Fejer, B. G., J. W. Jensen, and S.-Y. Su (2008), Quiet time equatorial F region vertical plasma drift model derived from ROCSAT-1 observations, *J. Geophys. Res.*, 113, A05304, doi:10.1029/2007JA012801.
- Fesen, C.G., R.G. Roble, A.D. Richmond, G. Crowley, B.G. Fejer, Simulation of the pre-reversal enhancement in the low-latitude vertical ion drifts, *Geophys. Res. Lett.*, 27, 1851-1854, 2000.
- Fritts, D.C., et al. (2009), Overview and summary of the Spread F Experiment (SpreadFEx), *Ann. Geophys.*, 27, 2141-2155.
- Hagan, M.E. and J.M. Forbes (2003), Migrating and nonmigrating semi-diurnal tides in the upper atmosphere excited by tropospheric latent heat release, *J. Geophys. Res.*, 108(A2), 1062, doi:10.1029/2002JA009466.
- Hagan, M. E., A. Maute, R. G. Roble, A. D. Richmond, T. J. Immel, and S. L. England (2007), Connections between deep tropical clouds and the Earth's ionosphere, *Geophys. Res. Lett.*, 34, L20109, doi:10.1029/2007GL030142.
- Hedin, A. E., E.L. Fleming, A.H. Manson, F.J. Schmidlin, S.K. Avery, R.R. Clark, S.J. Franke, G.J. Fraser, T. Tsunda, F. Vial and R.A. Vincent (1996), Empirical Wind Model for the Upper, Middle, and Lower Atmosphere, *J. Atmos. Terr. Phys.*, 58, 1421.
- Huba, J.D., G. Joyce, J. Krall, C.L. Siefring, and P.A. Bernhardt, Self-consistent modeling of equatorial dawn density depletions with SAMI3, *Geophys. Res. Lett.*, 37, L03104, doi:10.1029/2009GL041492, 2010.
- Immel, T. J., E. Sagawa, S. L. England, S. B. Henderson, M. E. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton (2006), Control of equatorial ionospheric morphology by atmospheric tides, *Geophys. Res. Lett.*, 33, L15108, doi:10.1029/2006GL026161.
- Krall, J., J.D. Huba, G. Joyce, S. Zalesak (2009a), Three-dimensional simulation of equatorial spread-F with meridional wind effects, *Ann. Geophys.*, 27, 1821.
- Kudeki, E. Akgiray, A., Milla, M.A., J.L. Chau, and D. Hysell, Equatorial spread-F initiation: post-sunset vortex, thermospheric winds, gravity waves, *J. Atmos. Solar Terr. Phys.*, 69, 2416-2427, 2007.
- Larsen, M. F., and C. G. Fesen (2009), Accuracy issues of the existing thermospheric wind models: can we rely on them in seeking solutions to wind-driven problems?, *Ann. Geophys.*, 27, 2277-2284.
- Liu, H. - L., W. Wang, A. D. Richmond, and R. G. Roble, Ionospheric variability due to planetary waves and tides for solar minimum conditions, *J. Geophys. Res.*, 115, A00G01, doi:10.1029/2009JA015188, 2010.
- Maruyama, T., and N. Matuura (1984), Longitudinal variability of annual changes in activity of equatorial spread F and plasma bubbles, *J. Geophys. Res.*, 89, 10903-10912.

- Millward, G.H., I.C.F. Muller-Wodarg, A.D. Aylward, T.J. Fuller-Rowell, A.D. Richmond, and R.J. Moffett, An investigation into the influence of tidal forcing on F region equatorial vertical ion drift using a global ionosphere-thermosphere model with coupled electrodynamics, *J. Geophys. Res.*, 106, 24733-24744, 2001.
- Reber, C., C. Trevathan, R. McNeal, and M. Luther (1993), The Upper Atmosphere Research Satellite (UARS) mission, *J. Geophys Res.*, 98, 10,643–10,647.
- Richmond, A.D., E.C. Ridley, and R.G. Roble, A thermosphere/ionosphere general circulation model with coupled electrodynamics, *Geophys. Res. Lett.*, 19, 601-604, 1992.
- Shepherd, G. G., et al. (1993), WINDII, the Wind Imaging Interferometer on the Upper Atmosphere Research Satellite, *J. Geophys Res.*, 98, 10,725–10,750.

Measuring Energy Inputs and ITM Response Using a Constellation of Small Satellites

Geoff McHarg, Delores Knipp

In 2011 our community celebrates the 50-year anniversary of Dungey's explanation of solar wind-geospace coupling. Years of community effort have extended his ideas to produce physics-based models that link the solar wind to the Ionosphere-Thermosphere-Mesosphere (ITM) system. Currently some of these run as nowcasts at NASA's Coordinated Community Modeling Center (CCMC). Event-based runs of coupled MHD-ITM Global General Circulation Models (GGCMs) indicate that the community is tantalizing close to upstream, solar wind driven short-term forecasting of the ITM state, (<http://ccmc.gsfc.nasa.gov/models/index.php>).

Scientific Rationale: Still, various aspects in the linkage are highly parameterized. The next leap in understanding the ITM regime requires assimilation of global measurements of the state variables and energy inputs. ITM forecasting requires a model containing the correct physics, computing capability that can run the model faster than real-time, and data to serve as initial conditions for the model forecast, Daley [1991]. Sufficient data to 1) assimilate into the models, and 2) verify that model physics is correct are lacking. Knipp et al. [2004] estimated that, on average, the magnetosphere provided 20% of the ITM energy budget (essentially Poynting flux and particle flux) with the remainder coming from shortwave-length solar photons. However, they found that on the occasion of extreme solar wind forcing, the global geomagnetic input exceeded the solar input. Understanding such events is of extreme interest to the scientific and operational community. **We propose that recent advances in satellite and measurement technology afford the possibility to launch a constellation of inexpensive small satellites to provide the measurements required to understand the fundamental physics that will enable ITM forecasting.**

Table1 shows the basic energy inputs and their deposition locations. Response to these inputs is the plasma and neutral density, temperature and winds. Measurement of the energy input, ITM response, and coupling to the lower atmosphere has a long, but discontinuous, history in space physics. Rarely has there been sufficient temporal or spatial coverage to capture the variability or dynamics of the ITM response to solar driving. To date, estimates and measurements of ITM state parameters have come from 1) single satellites (e.g. TIMED); 2) two satellites at different altitudes (e.g. DE-1 and 2); and 3) from assimilation of disparate, and irregularly spaced, global ground- and space-based data. A major interaction between the plasma and the neutrals comes from the neutral winds, Thayer and Semeter [2004]. This interaction is most often inferred because measurements have been lacking.

Data from NOAA's GOES and POES support the production of solar indices for to represent dayside thermospheric EUV and FUV heating [Bowman et al., 2008]. A still missing link is global view of the ITM state from a sufficiently dense network of measurements of the field and particle inputs, and appropriate ionospheric and thermospheric responses. Anderson et al. [2000, 2002] have used the magnetometers on the Iridium satellite constellation to estimate the cross-track magnetic field perturbations due to Birkeland currents. The Iridium constellation has over 70 satellites distributed in 6

orbital planes. Waters et al. [2004] combined these data with ground based radar measurements of ionospheric convection to estimate the imposed electric field, and hence provide the high-latitude Poynting flux into the ionosphere. Recently Anderson et al. [2008] have described an update to this capability, the Active Magnetosphere Polar Electrodynamics Response Experiment (AMPERE), using new flight software, ground systems, and science-data systems to for the Iridium satellites in this constellation

Table 1: Energy input and deposition altitude [Judith Lean, Naval Research Laboratory and Knipp work in progress, 2010]

Source	Energy Flux (W/m ²)	Deposition Altitude (km)	Latitude/ Longitude
Solar Radiation			Dayside
X ray-EUV 5nm–120 nm	0.0032 ± 0.0009	100–500	
EUV-UV 121nm–300 nm	14.9 ± 0.1	30–120	
UV-RF 301nm–10000 nm	1350.5 ± 0.5	Surface	
Stellar Radiation			Global
Total Starlight	0.0000018	Surface–120	
UV 123 nm–135 nm	0.000001	90–120	
Particles			
Solar energetic protons	0.002 ± 0.002	10–90	Polar Cap
Magnetospheric protons	0.001–0.006	100–130	High Latitude
Magnetospheric electrons	0.003–0.03	30–130	High Latitude
Galactic cosmic rays	0.000007	0–90	
Magnetospheric Poynting Flux	0.000015–0.15	100–500	High Latitude
Tidal Forcing from the Lower Atmosphere	0.00005	100–125	Dayside

Figure 1 shows an example of localized neutral density perturbations discovered with the help of data assimilation of ground and space-based electric and magnetic field observations, a satellite accelerometer and a solar wind-influenced global circulation model [Crowley et al., 2010]. The simulated density enhancement roughly matches the CHAMP accelerometer data for the interval, and is consistent with a heretofore unexplained, near-cusp density structure inferred from statistical studies of various satellite drag data bases. The fact that this localized type of ITM response has eluded discovery during the 5 decades of space-based measurements is testimony to the need for a more organized and consistent approach to ITM observations. The regional coherence of such structures and their global ramifications can only be understood and predicted with an adequate coverage/density of ITM data.

We argue that rather than make exquisite measurements of ionospheric and thermospheric energy inputs and responses from one or two satellites, it is better to have a large number of “good enough” measurements from a constellation of satellites commensurate with those derived from Iridium and its follow-on system AMPERE. Sufficient quantity and coverage of “good enough” measurements of the relevant

parameters are needed to constrain the data assimilation schemes that are likely to be the foundation of ionospheric and thermospheric forecast capability.

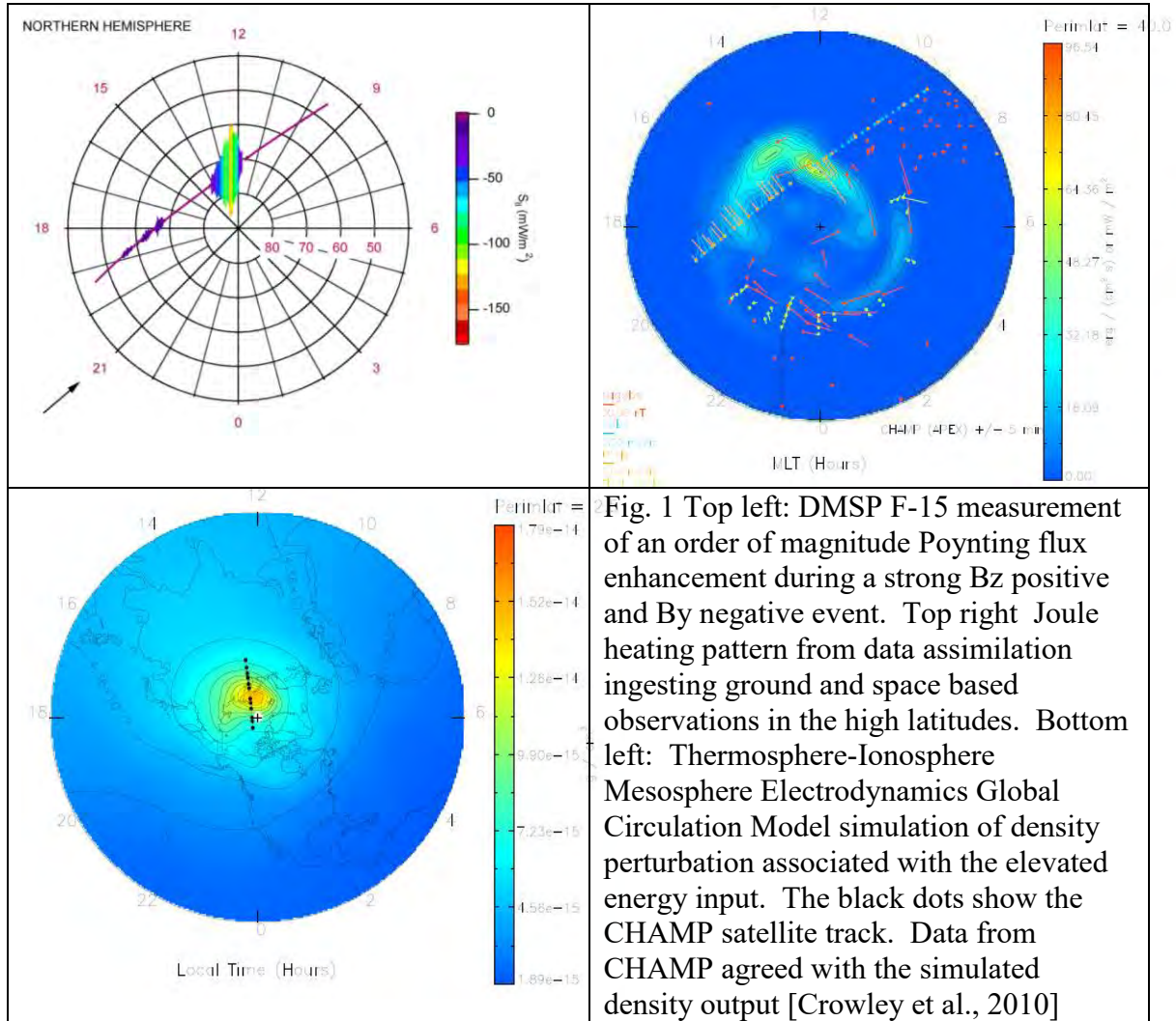


Fig. 1 Top left: DMSP F-15 measurement of an order of magnitude Poynting flux enhancement during a strong Bz positive and By negative event. Top right: Joule heating pattern from data assimilation ingesting ground and space based observations in the high latitudes. Bottom left: Thermosphere-Ionosphere Mesosphere Electrodynamics Global Circulation Model simulation of density perturbation associated with the elevated energy input. The black dots show the CHAMP satellite track. Data from CHAMP agreed with the simulated density output [Crowley et al., 2010]

We estimate the number of satellites needed in the constellation by looking at the temporal and spatial coverage required to resolve known temporal and spatial changes.

Temporal Coverage: Revisit time of any one spot in an orbit plane every 9-10 minutes. This corresponds roughly to the time it takes the ionosphere to respond to dynamic changes in a large geomagnetic disturbance. Using a low-earth-orbit (LEO) period of ~90 minutes, this means we need approximately 10 satellites per orbit plane.

Longitudinal Coverage: At least 6 orbit planes. We admit this is the least well understood requirement. We match the 6 orbit planes of the Iridium constellation since this has been shown to work for the magnetometer measurements. If spaced appropriately we will thus have two orbit planes each in the dayside and nightside, and one each in the dawn and dusk meridian. The constellation requirement is thus at least 60 satellites. The proposed measurements needed are shown in Table 2.

Table 2: ITM Measurements from a Small/Nano Satellite Constellation

Parameter Measured	Reason for measurement	Possible Techniques
Vector electric field	Energy Input	Voltage over distance and ion drift meter
Vector magnetic field	Energy Input	Fluxgate and induction coils
Energetic Particles	Energy Input	Electrostatic Analyzer
Plasma density	Ionospheric Response	Langmuir Probes and Electrostatic Analyzer
Plasma Temperature	Ionospheric Response	Langmuir Probes and Electrostatic Analyzer
Neutral Density	Thermospheric Response	Ionize neutrals, optical methods, and accelerometers
Neutral Temperature	Thermospheric Response	Ionize neutrals and optical methods
Neutral Winds	Coupling Mechanism	Ionize neutrals and optical methods

Our vision is a total system cost commensurate with that of a NASA Small Explorer (SMEX) and Missions of Opportunity proposal. As of 2007, the cost cap for a SMEX proposal was 105M (\$FY2008), (<http://nspires.nasaprs.com/external/viewrepositorydocument/cmdocumentid=85529/SMEX%20Community%20Announcement.pdf>). Thus the cost per satellite including payload must be approximately \$1.5M.

In our estimation, the only competitive approach in this price range is to use small and or nanosatellites. As an example, the National Science Foundation funds nanosatellites called CubeSats for Space Science missions. These are funded at a maximum of \$900k per mission, <http://www.nsf.gov/pubs/2008/nsf08549/nsf08549.htm>. Among these missions are most of the sensors required in the above table.

For instance, the Dynamics Ionosphere CubeSAT experiment hosts an electric field sensor and Langmuir probe. The Firebird mission hosts a solid state detector for relativistic electrons, and the CubeSAT for Ions, Neutrals, Electrons and Magnetic fields (CINEMA) has an in-situ combined ion-electron sensor in the 2-4 keV and along with a magneto-resistive magnetometer. The Colorado Student Space Weather Experiment (CSSWE) hosts a relativistic electron and proton telescope sensor designed to look at 0.5-3 MeV electrons and 10-40 MeV ions.

In addition to the NSF missions, the Department of Defense has sponsored a space weather CubeSat capable of measuring plasma density and temperature, as well as neutral density and temperature and neutral winds. Between the CubeSAT's sponsored by NSF and DoD, all the relevant measurements mentioned in Table 2 have been planned for a CubeSAT mission.

NSF and DoD CubeSAT's are already serving as valuable risk reduction efforts allowing the community to plan an ITM constellation. While the details of such a

constellation are still not clear, further investigation of how to move from the very promising individual experiments on NSF and DoD CubeSATS to a full ITM constellation is clearly warranted.

Educational Rationale: The very prolonged process associated with large space missions severely limits the opportunity for the nation's young Science Technology, Engineering and Mathematics workforce to be involved in developing space hardware and space missions. The constellation concept we propose would expand such opportunities. A consortium of universities and colleges could participate in the hardware and mission design and perhaps mission operations.

References:

- Anderson, B. J., Takahashi, K., and Toth, B. A.: Sensing global Birkeland currents with Iridium engineering magnetometer data, *Geophys. Res. Lett.*, 27, 4045, 2000.
- Anderson, B. J., Takahashi, K., Kamei, T., Waters, C. L., and Toth, B. A.: Birkeland current system key parameters derived from Iridium observations: Method and initial validation results, *J. Geophys. Res.*, 107, A6, SMP 11-1-11-13, 10.1029/2001JA000080, 2002.
- Anderson, B.J., Continuous Global Birkeland Currents from the Active Magnetosphere and Planetary Electrodynamics response Experiment: AMPERE, 37th COSPAR Scientific Assembly, p88, 2008.
- Bowman, B. R., W. Kent Tobiska, F. A. Marcos, C. Y. Huang, C. S. Lin, and W. J. Burke, A New Empirical Thermospheric Density Model JB2008 Using New Solar and Geomagnetic Indices, AIAA/AAS Astrodynamics Specialist Conference August 2008, Honolulu,
- Crowley, G., D. J. Knipp, K. A. Drake, J. Lei, E. Sutton, and H. Lühr, Thermospheric density enhancements in the dayside cusp region during strong B_Y conditions, *Geophys. Res. Lett.*, 37, L07110, doi:10.1029/2009GL042143, 2010
- Daley, Roger, Atmospheric Data Analysis, Cambridge Atmospheric and Space Science Series, Cambridge University Press, 1991.
- Dungey, J. W. , Interplanetary Magnetic Field and the Auroral Zones, *Phys. Rev. Lett.*, 6, 47-49, 1961.
- Knipp, D. J., W. K. Tobiska and B. A. Emery Solar Direct and Indirect Thermospheric Heating Sources for Solar Cycles 21-23, *Solar Physics*, 224:495-505, 2004
- Knipp, Space Weather and the Physics Behind It, Text book in preparation, 2010
- Thayer, J.P. and J. Semeter, The convergence of magnetospheric energy flux in the polar atmosphere, *Journal of Atmospheric and Solar-Terrestrial Physics*, 66 807-824, 2004.
- Waters, C.L., B.J. Anderson, R.A. Greenwald, R.J. Barnes, and J.M. Ruohoniemi, High-latitude pointing flux from combined Iridium and SuperDARN data, *Annales, Geophysicae*, 22: 2861-2875, SRef-ID: 1432-0576/ag/2004-22-2861, 2004.

Synergy Between Large Data Sets, First-Principles and Empirical Models of the Magnetosphere

V. G. Merkin, M. I. Sitnov, A. Y. Ukhorskiy, E. Talaat, J. Gjerloev (JHU/APL)
L. Waldrop (University of Illinois)

This white paper advocates for the need of support for programs that take advantage of both the wealth of currently available and future ionospheric and magnetospheric measurements and the maturity of physics-based models of the ionosphere-thermosphere-magnetosphere system (ITM). To motivate and frame the forthcoming discussion we start with a quote regarding data assimilation in terrestrial weather modeling [1]: “The improvement in skill of numerical weather prediction over the last 40 years... is due to four factors:

- The increased power of supercomputers, allowing much finer numerical resolution and fewer approximations in the operational atmospheric models;
- The improved representations of small-scale physical processes... within the models;
- The use of more accurate methods of data assimilation...
- The increased availability of data...”

It will not be an exaggeration to say that numerical modeling of *space* weather, or more broadly, of the ITM system in general, has reached a state where, according to all four factors mentioned, major advances are possible. Here we would like to specifically concentrate on the last factor, “the increased availability of data”.

Any numerical model inevitably makes approximations to the physics of the phenomenon studied and/or to the specification of initial and boundary conditions owing to the lack of observational data. For instance, the ITM system is largely describable by magnetohydrodynamic (MHD) equations, but there are major domains of the system where the MHD approximation fails, e.g. in the ionosphere, the inner

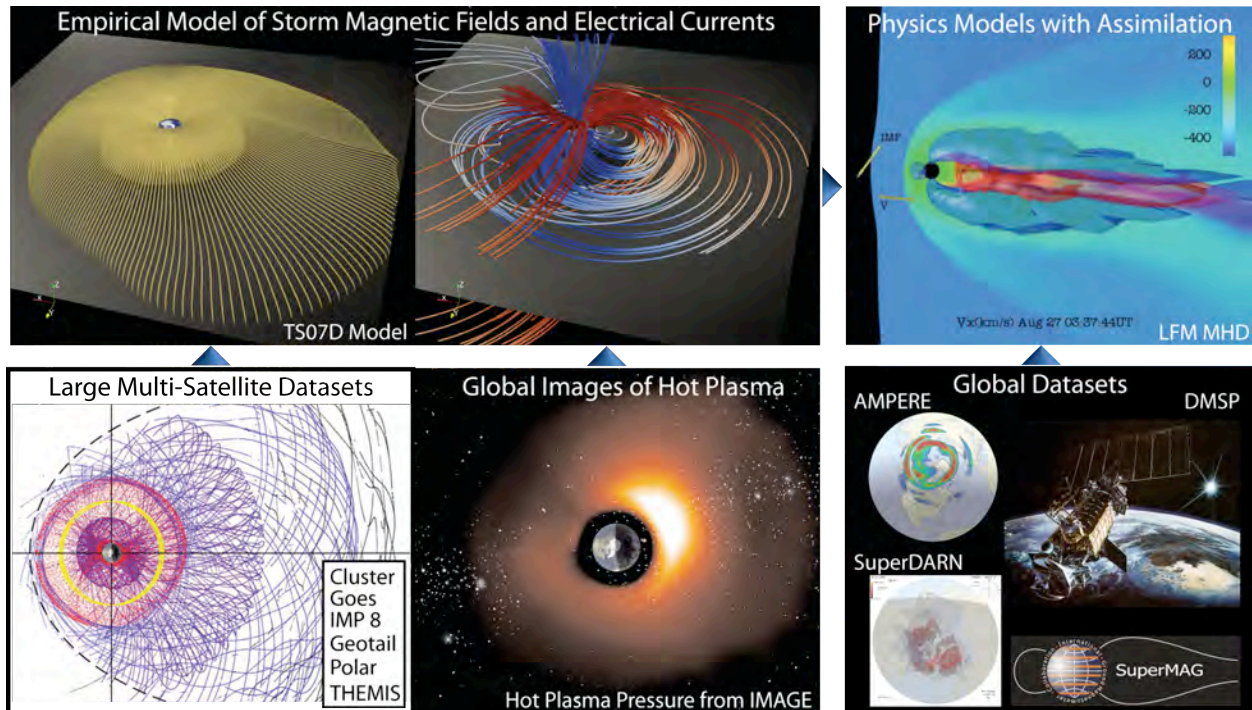


Figure 1. First-principles and empirical models and large data sets can work in synergy to produce a realistic representation of the ionosphere-thermosphere-magnetosphere system. The figures are adopted from: LFM (Lyon-Fedder-Mobarry global MHD model) [17], IMAGE [18], AMPERE [19], DMSP [20], SuperDARN [21], SuperMAG [22].

magnetosphere, as well as in the magnetotail and at the magnetopause, where magnetic reconnection takes place. While the work on inclusion of new physics through improved model resolution and coupling with smaller-scale physics models (the first two factors in the above list) is a viable and crucially important direction (which will undoubtedly be represented by other white papers), here we would like to draw the community's attention to the fact that a large and ever-growing collection of data has been accumulated that can be used to improve the existing models. We provide two examples below:

1. The primary means of cross-scale coupling in the ITM system is through closure of magnetospheric currents. The ionospheric conductivity, a measure of how freely charged particles in the ionosphere can flow past the neutrals, controls this process thereby governing the energy deposition in the ionosphere-thermosphere (via collisions with neutrals) due to Joule heating and particle precipitation and also regulating the electric fields, flows and distributions of plasma at high altitudes in the magnetosphere. The conductivity itself is affected by the state of the complex ITM system: the magnetosphere influences electron density by driving strong ionospheric turbulence and particle precipitation, while the thermospheric winds and temperature regulate the neutral density distribution and thus collision frequencies. Ionospheric conductivity is therefore a primary variable in this complicated chain of diverse physical processes ultimately responsible for coupling of the solar wind-driven fully-ionized magnetosphere to the neutral atmosphere. The sobering fact however is that ionospheric conductivity is highly variable in both distribution and magnitude and perniciously difficult to characterize, because directly measuring ionospheric neutral and electron density profiles, including local turbulent and kinetic modifications, on a global scale remains unfeasible.

For most global MHD models of the Earth's magnetosphere as well as kinetic models of the inner magnetosphere the ionosphere plays the role of the inner boundary condition. Ionospheric conductivity is crucial for this specification as it controls the coupling, via Ohm's law, between the field-aligned currents (calculated by the magnetospheric simulation) and the ionospheric convection electric field (the boundary condition used in the simulation) [e.g. 2]. Historically, the conductivity tensor (both Pedersen and Hall terms) has been calculated using semi-empirical models including the background solar EUV ionization and auroral particle precipitation [e.g. 3]. While the first source is well-specified, the second is associated with large uncertainties. In addition, these conventional models lack conductivity enhancements due to ionospheric turbulence – effects that have been theoretically predicted to be capable of doubling the total height-integrated conductivity during disturbed geomagnetic conditions [4].

The long-standing problem of poor specification of ionospheric conductivity in magnetospheric modeling may finally be mitigated now that there exist multiple observational data sets of ionospheric electrodynamic quantities, e.g. AMPERE (field-aligned currents), SuperDARN, DMSP (convection electric field), and SuperMAG (ground magnetic perturbations). Combination of such comprehensive synoptic data sets from these large and global-scale networks of radars, ground stations and satellites may be applied synergistically to produce improved specifications of ionospheric conductivity, and, in fact, can be assimilated into global magnetosphere, ring current or coupled ITM models. Such efforts will enable new understanding of cross-scale interactions in the ITM system driven by realistic conductivity distributions.

2. Another example comes from the need to evaluate hot plasma pressure, which determines the strength of the ring current responsible for the strongest disturbances in geospace, magnetic storms. Global MHD models underpredict the pressure in the inner magnetosphere, because they do not include an adequate description of energy-dependent drift physics and they underestimate consequences of explosive magnetic reconnection in the tail. Dedicated kinetic ring current models, in turn, may have the right micro-physics but lack global self-consistent magnetic fields and suffer from uncertainties in specification of boundary conditions, whereby the corresponding information is ingested into the model either at geosynchronous orbit, using LANL satellite data on plasma density and temperature [e.g. 5], or further outside, using crude empirical relationships connecting density and temperature in the plasma sheet with the bulk plasma parameters in the solar wind [e.g. 6].

On the other hand, plasma pressure can be reconstructed from extensive measurements of magnetospheric magnetic field, using simple force balance condition $\nabla P = J \times B$ and the abundance of space magnetometer data from many past and present missions. Moreover, modern data-mining and data-fitting techniques allow one to organize these data so as to reconstruct or even predict the detailed empirical picture of the geomagnetic field and underlying currents. In particular, the recent empirical model TS07D [7] combines magnetic field measurements from years of observations by such magnetospheric spacecraft as Cluster, Geotail, Polar, IMP8, GOES 8, 9, 10, and 12, and can be augmented by other observers, e.g. the 5 THEMIS spacecraft. Such a model can conceivably be used to nudge a first-principles magnetospheric model toward a more realistic solution either through a data assimilation scheme or some other way of data ingestion. The quasi measurements of global magnetic field distribution may be used directly or indirectly, through a pressure-reconstruction scheme based on force balance described above. In this case, TS07D can be used as a source of information about hot particle pressure in the inner magnetosphere– the plasma population component underrepresented in first-principles models– and, again, assimilated or ingested to produce stronger inner magnetosphere pressure, ring and Region 2 currents, and thence magnetotail stretching. A similar data ingestion concept could be formulated for hot particle pressure distributions reconstructed from Energetic Neutral Atom (ENA) images of the magnetosphere [e.g. 8] available from past (IMAGE [9]), present (TWINS [10]) and prototype (GEMINI [11]) spacecraft missions.

There are already instances of very successful application of conventional data assimilation methods in the physics of the ITM system, particularly in the ionosphere [12-14] and radiation belts [15, 16]. In other cases, such as the ones we presented above, detailed data assimilation and ingestion schemes and algorithms are yet to be worked out. The goal of this white paper is by no means to introduce such schemes; rather, it is to urge the funding agencies to recognize the fact that the existence of large observational datasets and well-developed empirical and physics-based models make our field ripe for scientific endeavors building on synergy between them.

References

- [1] Kalnay, E. “Atmospheric modeling, data assimilation and predictability”, Cambridge University Press, 3rd edition, p. 2, 2006.
- [2] Wolf, R. A., The Quasi-Static/Slow-Flow/Region of the Magnetosphere. Solar-Terrestrial Physics: Boston College, p. 303, 1983.
- [3] Fedder, J. A., et al. Global numerical simulation of the growth phase and the expansion onset for a substorm observed by Viking. *J. Geophys. Res.*, vol. 100, p. 19083, 1995.
- [4] Dimant, Y. S., and G. M. Milikh, Model of anomalous electron heating in the E region: 1. Basic theory. *J. Geophys. Res.*, vol. 108, p. 1350, 2003.
- [5] Liemohn, M. W., et al., Modeling the ring current response to a sawtooth oscillation event, *J. Atmos. Solar-Terr. Phys.*, vol. 69, pp. 67-76, 2007.
- [6] Zheng, Y., et al., Relationship between Region 2 field-aligned current and the ring current: Model results, vol. 111, A11S06, doi:10.1029/2006JA011603, 2006.
- [7] http://geomag_field.jhuapl.edu/model/.
- [8] Roelof, E. C., et al. Derivation of currents and diamagnetic effects from global plasma pressure distributions obtained by IMAGE/HENA. *Advances in Space Research*, vol. 33 pp. 747, 2004.
- [9] <http://pluto.space.swri.edu/image/>
- [10] <http://twins.swri.edu/mission.jsp>
- [11] <http://www.terpconnect.umd.edu/~drake/DCTownHall/talks/brandt.pdf>
- [12] Schunk, R. W., et al. Global Assimilation of Ionospheric Measurements (GAIM). *Radio Science*, vol. 39, 2004.
- [13] Richmond, A. D., Assimilative mapping of ionospheric electrodynamics, *Advances in Space Research*, vol. 12, pp. 59, 1992.

- [14] Bust, G. S., and G. Crowley, Tracking of Polar Cap Ionospheric Patches using Data Assimilation, *J. Geophys. Res.*, 112, A05307, doi:10.1029/2005JA011597, 2007.
- [15] Kondrashov, D., Y. Shprits, M. Ghil, and R. Thorne, A Kalman filter technique to estimate relativistic electron lifetimes in the outer radiation belt, *J. Geophys. Res.*, 112, A10227, doi:10.1029/2007JA012583.
- [16] Koller, J., Y. Chen, G. D. Reeves, R. H. W. Friedel, T. E. Cayton, and J. A. Vrugt, Identifying the radiation belt source region by data assimilation, *J. Geophys. Res.*, 112, A06244, doi:10.1029/2006JA012196, 2007.
- [17] Wiltberger, M. J., et al., Analysis and visualization of space science model output and data with CISM-DX, *J. Geophys. Res.*, vol. 110, p. 09224, 2005.
- [18] Cover page, *Geophys. Res. Lett.*, vol. 29, 20, 2002.
- [19] <http://multivu.prnewswire.com/mnr/iridium/45153>
- [20] <http://www.stratcom.mil/imagelibrary/space/21/278>
- [21] <http://superdarn.jhuapl.edu>
- [22] <http://supermag.jhuapl.edu>

GEospace Magnetosphere-Ionosphere-Neutral Interaction (GEMINI)

A science mission framework targeting the core of the globally coupled magnetosphere-ionosphere providing continuous data optimized for global model validation

D. Mitchell (JHU/APL), P. Brandt (JHU/APL), E. Donovan (U. Calgary), M-C Fok (GSFC), S. Fuselier (Lockheed), D. Gallagher (MSFC), J. Goldstein (SwRI), J. Kozyra (U. Mich), R. Meier (NRL), S. Mende (UCB), T. Moore (GSFC), P. Newell (JHU/APL), S. Ohtani (JHU/APL), G. Parks (UCB), L. Paxton (JHU/APL), T. Sotirelis (JHU/APL), J. Spann (MSFC), R. Wolf (Rice)

Introduction

The Solar System is comprised of a collection of coupled systems. In fact it is one large system-of-systems in which the Sun is the primary source of energy. Earth, as our home, is the planet we know most about. We have ventured extensively into Earth's immediate surroundings, called Geospace, and have learned how Geospace is intrinsically interconnected over diverse scales of space and time. Over time an armada of spacecraft have revealed extraordinary global phenomena such as geomagnetic storms, with plasma energized in the magnetosphere distorting its magnetic and electric fields leading to strong modification of the radiation belts, and ionospheric storms, in which large regions of Earth's ionosphere are redistributed over the globe. Plasma and fields in both the ionosphere and magnetosphere are coupled with each other and multiple processes (precipitation, Joule heating, particle acceleration, wave generation, mass outflow) are competing simultaneously on global and local scales to produce the global phenomena we observe. Such a system requires *simultaneous global and continuous* measurements of the critical regions of the magnetosphere and ionosphere. By the very nature of coupled systems, observation of the relationships among components is critical to understand and characterize the collective behavior. Imaging is the most effective (and arguably the only) way to carry out such system-level observations. Understanding of the complex interrelationships is impossible without a global view of Geospace. Imaging provides more information than any practical number of distributed single-point measurements and therefore is indispensable for system-level exploration.

In this paper, we outline a science mission framework that targets the core of the globally coupled magnetosphere-

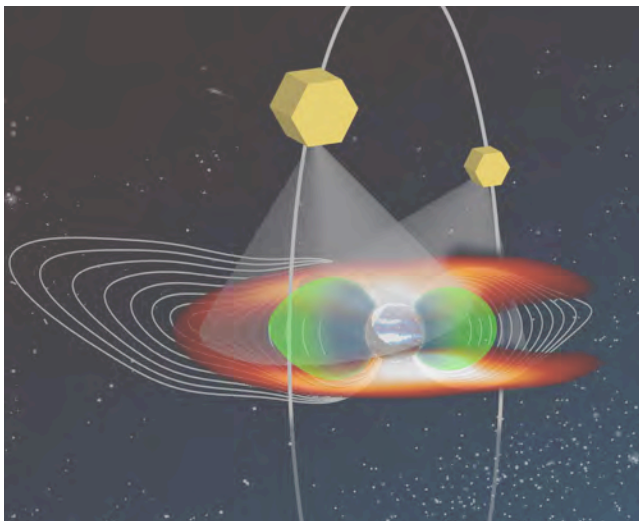


Figure 1: GEMINI targets the core of the coupled magnetosphere-ionosphere system by providing global, continuous 3D images of the ring current (orange), plasma sphere, aurora, ionospheric-thermospheric dynamics and flows. GEMINI obtains the collective behavior of the complex, coupled, interconnected Geospace system and the measurement resolution enabling global model validation and discovery science.

ionosphere system from two platforms in a high circular orbit, providing continuous data optimized for global model validation. A range of outstanding mysteries in space physics drives the choice of observational targets. Planetary magnetospheres are efficient particle accelerators, generating a “ring current” that couples the inner magnetosphere with the ionosphere, and produces field distortions and wave activity leading to dramatic radiation belt intensification. In ionospheric storms, plasma is redistributed globally by magnetospheric as well as solar wind driven currents, auroral precipitation and solar irradiation, disrupting communications and navigation systems. GEMINI targets the ring current, plasma sphere, aurora, and the ionospheric-thermospheric plasma redistribution through continuous imaging of both northern and southern auroral emissions; mid-latitude far ultraviolet (FUV) imaging; stereo extreme ultraviolet (EUV) imaging of the plasmasphere; and stereo energetic neutral atom (ENA) imaging of the ring current and the near-Earth plasma sheet. These measurements are made together with

ground-based radar (both HF SuperDARN and ISR such as AMISR), auroral imaging arrays, and magnetometer chains) as well as other existing space assets such as

IRIDIUM/AMPERE, DMSP, GOES, LANL, GPS, whatever is upstream, and any NASA Explorers or other Earth magnetosphere/ionosphere missions that may fly contemporaneously. With ground-based and LEO data providing detailed information on field aligned currents, ionospheric electron densities, temperatures and flows, GEMINI will provide the means to link the global scale magnetospheric state with these detailed ionospheric conditions.

Global modeling and assimilation is a core part of the GEMINI concept. We have reached a stage where we no longer can understand the behavior of geospace without the use of physical models to retrieve the mechanisms that compete simultaneously on several scales. Meanwhile, the existing global models are continually starved for global data against the models can be validated. Imaging plays a natural and necessary role in providing global validating observations of system level interactions and processes that are not possible otherwise.

Dynamics Explorer made a good beginning to the investigation of this coupling. DE imaging and ionospheric and magnetospheric *in situ* observations gave us a simultaneous view of small-scale physical processes and geospace at the system-level, albeit projected onto the 2D ionosphere. The DE images provided context, and more importantly, they provided our first means of assessing how the smaller scales probed by the *in situ* observations coupled to the larger system. ISTP took this to another level, using multiple coordinated *in situ* probes in combination with auroral imaging. The IMAGE Explorer mission expanded the kinds of imaging available to characterize the global scale of the Geospace system, adding magnetospheric energetic ENA and plasmaspheric EUV imaging to the FUV imaging that had been used on the earlier missions. And the many *in situ* measurements provided by FAST, DMSP, Polar, Geotail, and other missions provided the means of assessing how the smaller-scale processes drove (or were driven by) the larger scale system. Features as fundamental as plasmaspheric drainage plumes (which were still argued about until the IMAGE EUV images proved their existence), the local time distribution of the partial ring current (which was shown by ENA imaging to vary between early morning and dusk from one storm to another), and the substorm acceleration of oxygen in the context of storms (revealed by ENA imaging) became instantly clear using the IMAGE global images. Combined with *in-situ* measurements it provided the global context required to interpret the *in-situ* measurements correctly. For example, was a spacecraft inside or outside of the partial ring current? Where was the plasmapause relative to the spacecraft measuring a particular plasma wave signature?

Mission Objective and Science Questions

GEMINI seeks to determine how the magnetosphere-ionosphere-thermosphere system is coupled and responds to solar and magnetospheric forcing and to provide a long-duration mission framework into which additional science missions can naturally fit.

GEMINI's science questions are chosen to address critical pieces in the global redistribution, acceleration and response in the magnetosphere and IT system as illustrated in Figure 2. We discuss these overarching science questions in terms of the global system and identify "ports" where external science investigations fit in naturally,

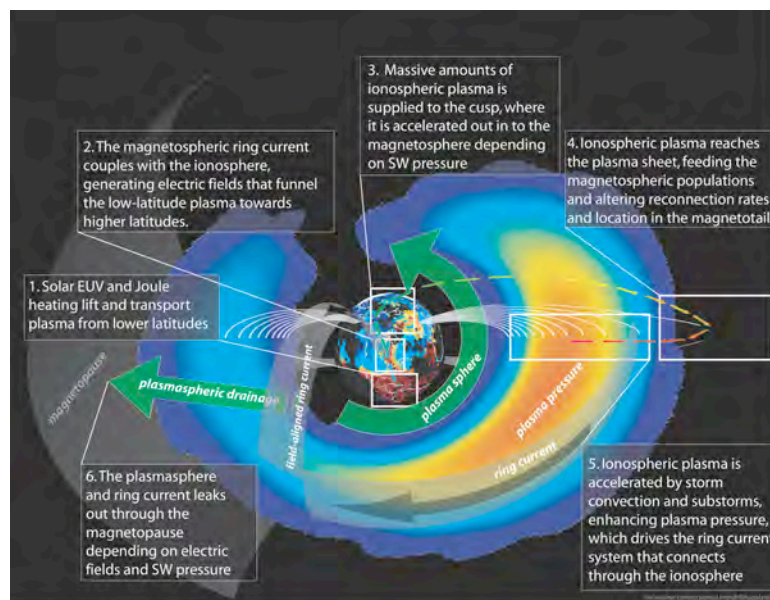


Figure 2. The magnetosphere-ionosphere-thermosphere system is complex. Global scale measurements are critical to placing *in situ* (HEO, MEO, and LEO) and ground-based measurements in an appropriate framework. Global models are strongly constrained by global scale imaging. The combination (global imaging, *in situ*, ground based, and models) will enable understanding the coupling among these interconnected plasmas.

enabled by GEMINI's long lifetime.

1. How does MI-coupling control plasma redistribution in the magnetosphere and ionosphere/thermosphere?

(GEMINI)

Both magnetospheric and solar forcing gives rise to IT plasma redistribution. Plasma accumulation in the afternoon/dusk sector at low latitudes appear to be funneled to higher latitudes that eventually reach the cusp, where most of the ionospheric outflow is seen in statistical studies. Is this snaking pattern a how plasma is supplied to the breathing hole of the Earth's ionosphere to space? The funnel coincides with the Sub-Auroral Polarization Stream (SAPS) channel, which appears almost perfectly reflected in the plasmaspheric drainage plume in the magnetosphere. The SAPS phenomena is believed to be an effect of the closure of the ring current through the low-conductance gap in the ionospheric trough region, which leads to enhanced ionospheric flows in a narrow region. Fundamental plasma physics mysteries are nicely illustrated in this interface region. As magnetospheric currents close through the different and dynamically changing conductances of the ionosphere, the global electric field of the ionosphere is altered and feeds back to the large-scale electric field of the magnetosphere producing macroscopic effects in the location and dynamics of both the ring current and the plasma sphere. Defying expectations that the ring current would shield the inner magnetosphere through these coupled processes, E-field in the equatorial magnetosphere displayed the *opposite* behavior, with surprisingly strong E-fields found very deep in the magnetosphere. The complicated MI-coupling also modifies important boundary regions such as the overlap between the hot and cold plasma of the ring current and plasma sphere, critical for wave growth that can accelerate electrons to high energies. The dense plasma of the plasma sphere is further stretched out in EUV-visible plasmaspheric drainage plumes that reach the dayside magnetopause, where the presence of heavy ions may regulate the dayside reconnection rates.

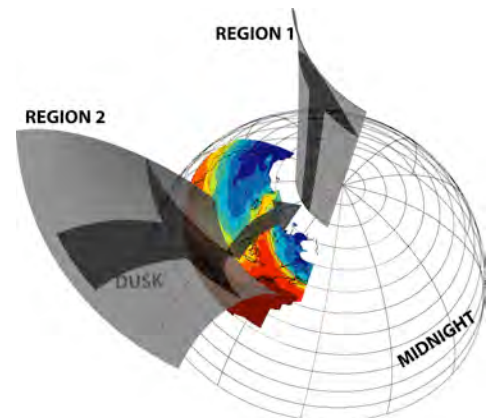


Figure 3: Strong ionospheric flows (SAPS) is only one example of a coupled phenomena that requires simultaneous imaging of both the magnetospheric “driver” and the ionospheric response. The SAPS flow is an effect of magnetospheric currents closing through the low-conductance ionospheric trough region [Brandt et al., 2008].

2. What are the processes governing ionospheric outflow from the polar ionosphere and its subsequent acceleration in the 5000 km gap region that transports it to the magnetosphere?

(External investigations needed)

After ionospheric plasma has been redistributed to higher latitudes, other processes take over that accelerate ions to escape energies. This acceleration region has seen very scarce measurements, impacting our ability to reproduce how ionospheric plasma populates the magnetosphere and leaving a gap in our understanding of when/how the ring current becomes dominated by heavy ions from the ionosphere (O^+). Addressing this critical question requires addition satellites taking measurements in this critical altitude range (5000-10000 km altitude), a good example of how a separate or coordinated mission might naturally fit in to the global science framework of GEMINI.

3. How is hot plasma created, accelerated and transported to the inner magnetosphere to form the driving plasma pressures of the plasma sheet and ring current and how do these couple to the ionosphere? (GEMINI HEO only)

Both ionospheric and solar wind plasma are believed to enter through the flanks of the magnetosphere from where it is transported in to the magnetosphere. While terms such as “storm convection” and “substorm injection” have been used to represent the collective mechanisms that are responsible for transporting and heating plasma to the energies that form the driving plasma pressure, observations show no resemblance to a convective plasma bulk motion with negligible response to solar wind driving [Hori et al., 2006] and no resemblance to “fronts” of substorm injections moving earthward, but rather fin-like structures populating the even the deep inner magnetosphere indicating a critical role of plasma instabilities [Ohtani et al., 2008; Yang et al., 2008; Keika et al., 2010; Nose et al., 2010], not unlike what is found at Saturn and Jupiter. Hot plasma represents a gap when it comes to modeling. Magneto Hydro

Dynamic (MHD) models capture the global dynamic field variations and their response to solar wind driving, but has no representation of hot plasma. So-called kinetic models (typically solving a bounce averaged Boltzmann equation) reproduce the large-scale features of the ring current provided source populations at its outer boundary, but do not reproduce the field variations required for energizing the hot plasma. To successfully couple these two families of models, model validation against global data is not only desired, but also required.

4. How does coupling between the ionosphere and the magnetosphere control inter-hemispheric currents and structure?

Earth's auroral zones map from the open polar cap at high latitudes to as low as $L \sim 4-5$ at low latitudes, during storms. However conjugate imaging of the north and south auroral zones has been only by chance, not design, and so data is sparse and wavelength coverage and spatial resolution poorly matched. GEMINI provides for a long, nearly continuous experiment in which conjugate, full auroral zone images may be analyzed for evidence of direct interhemispheric connection, dynamics confined to one hemisphere only, and north-south asymmetry in precipitation patterns while simultaneously obtaining ring current pressure and composition with stereo ENA imaging and global electric field structure and strength with stereo EUV imaging of the plasmasphere internal structure and flow speeds, drainage plumes, and plasmopause location. An added AMPERE-like capability will enable GEMINI to address how field-aligned currents (FACs) are affected by hemispheric conductance differences.

Mission Implementation

The GEMINI mission uses two identical nadir-viewing spacecraft in an $8 R_E$ circular polar orbit with variable orbital phase separation to enable global and continuous stereo measurements of the 3D distributions of the ring current pressure and plasma sphere. Auroral and airglow imaging provide nearly continuous and, depending on mission phase, conjugate remote observation of the ionosphere and thermosphere, enabling estimations of the precipitating electron energy flux, ionospheric electron density, and thermospheric composition and energy input. UV analysis techniques now exist to rapidly convert airglow emission intensities into geophysical quantities (e.g., precipitating particle average energy and flux, ionospheric Pedersen and Hall conductivity, and thermospheric O/N₂ ratio). Large-scale maps of the spatial, intensity and energy distribution of proton and oxygen precipitation will be made by measuring their small signature Doppler-shifts. This approach enables tracking of dynamical features in the ionosphere, such as polar cap ionization patches, positive and negative storm effects, neutral atmospheric responses to auroral heating, ionospheric scintillations and plasma bubbles.

The important coupling components missing from the high altitude measurements will be obtained using GEMINI's low altitude components, which comprise measurements from a large range of resources including DMSP, IRIDIUM/AMPERE (for global field aligned currents), SuperDARN radar at high and mid latitudes for ionospheric flows and electron density profiles, Millstone Hill and AMISR radar for mid latitude dynamics, GPS TEC maps, and magnetometer and all-sky camera arrays. The result will be global specification of the ionospheric electric field and electric current patterns in both hemispheres, essentially completing the electrodynamic picture at low altitude. A new feature of this mission, enabled by nearly simultaneous observations in both hemispheres, is the ability to investigate the little-understood role of inter-hemispheric asymmetry (due to the tilt of the dipole and rotation axes) on the global system behavior.

Mission Strategy

- Image the 3D magnetospheric distribution and evolution of critical populations:
 - The plasma pressure in the plasma sheet and ring current using energetic neutral atoms (ENA) with sufficient temporal and spatial resolution to retrieve the electrical current system that distorts the magnetic field and that connects through the ionosphere producing the electric field
 - The plasmasphere using extreme ultraviolet (EUV) with sufficient temporal and spatial resolution to retrieve response to the electric field
- Image the ionosphere-thermosphere system using
 - multiple wavelengths of far ultraviolet (FUV) to assess auroral energy input, conductance, IT composition changes and redistribution
 - multiple radars to image the plasma flow and electron density profiles in order to estimate the ionospheric electric field and constrain conductance

- multiple low-flying magnetometers to derive global FAC patterns, such as AMPERE to help constrain ionospheric electric fields
- ground based magnetometers to map magnetospheric coupling regions at low altitude
- Ground based all-sky camera arrays to obtain high spatial and time resolution auroral configuration and precipitation patterns.

Mission Description

- Two High Altitude Spacecraft in $\sim 8R_E$ circular near-polar orbit (continuous and global 3D)
 - 1 ENA, 1 EUV, 3 FUV cameras per S/C
 - nadir pointing with yaw about nadir
- Ground-based radar network to cover the mid- to high-latitude ionosphere
- Existing ground-based magnetometer networks and all-sky camera arrays
- Low-altitude satellite component or existing facilities such as AMPERE and Swarm
- 2 year life time required, but orbit enables a long-duration (~ 10 yr) mission
- No new technology development
- Taurus II provides ample launch capacity; we have book-kept \$100M to \$130M in the mission costs

Attitude Control Requirements (driven primarily by auroral imaging):

- Attitude control 0.1° (3 sigma)
- Attitude knowledge 0.02° (3 sigma) (after the fact)
- Orientation 3-axis with payload sensor axis to nadir

Mission Cost (2 spacecraft HEO component only)

- Original detailed estimate made in 2003, assuming Delta II 7925.
- Current cost estimated using 1.2 (low) to 1.5 (high) inflation across all mission components.
- Cost, FY2010 \$290M to \$405M including 25% reserve, all mission phases;
- For 10 year mission, cost range is \$350M to \$500M

Instruments:

Component [Heritage]	Measurement	Mass kg	Power Watts	FOV (deg)	Envelope (cm)	Bitrate kbps
FUV Spectrographic Imager [IMAGE]	Oxygen-I (1356 Å) Proton (Doppler shifted 1216 Ly- α)	21	2.4	15x15	80x50x30	40
FUV 1	LBH long (170-190 nm)	6	1.5	15x15	32x17x18	20
FUV 2	LBH short (140-160 nm)	6	1.5	15x15	32x17x18	20
ENA [Cassini, IMAGE]	2 - 200 keV H, 10 - 300 keV O 2° ang. res. > 10 keV/nuc. 30 second time resolution	6.1	3	120x120	40x50x50	20
EUV 30.4 nm [IMAGE]	He 30.4 nm	20	20	90x90	50x50x50	10
Electronics		7.5	11	N/A	23x23x23	0.25
Total		66.6	44.4			110.5

HEO Instruments (2 flight, 1 flight spare for each type):

FUV (SI + 2 CAM+ELECTRONICS) \$16400K

EUV \$7800K

ENA \$17200K

GEMINI can be supplemented by any remote or in-situ measurements of plasma, energetic particles, electric and magnetic fields at as many points separated by distances of RE as feasible. Many in situ measurements are generally available from a variety of existing assets (e.g., AMPERE, GOES, LANL Geosynch, GPS, ACE/WIND/SOHO/DMSP (in situ + SSUSI, SSULI)/Swarm, etc.). While ionospheric dynamics can be studied in detail by in situ measurements, valuable complementary tools to obtain medium scale morphology include remote sensing through auroral imaging, ground-based auroral imaging networks and magnetometer networks, as well as radar (both HF, such as SuperDARN including the newer mid-latitude stations, and incoherent scatter, such as AMISR). Global imaging provides the system-level view while high spatial and temporal resolution imaging of more limited regions probe detailed coupling. Simultaneous, conjugate ENA and auroral imaging provides critical information on global ring current evolution and pitch angles, and on how ionospheric conditions differ in the two hemispheres and what those differences can tell us about magnetospheric coupling—in particular, the flow of particles and currents between different parts of the system. Coupling between neutral winds and the ionosphere, serving as either a generator or a load for the magnetosphere, requires measurement of neutral winds at altitudes where neutral-ion collisions are important to transferring momentum between the two populations. This component would require in situ and/or interferometric neutral wind measurements from LEO satellites.

While we have not specified a directly funded low altitude component to address the coupling problem, many such components can be envisioned. Our estimate of the GEMINI HEO mission costs places the mission squarely in the low-cost category. This leaves ample room to add a specifically funded low altitude spacecraft, or even pair of spacecraft. For example, one appropriately instrumented spacecraft in a FAST-like orbit could measure ion outflows from the polar cap, auroral zone, and plasmaspheric refilling (addressing science question 2), while also dipping deep into the mid and high latitude ionosphere at perigee to directly measure ionospheric and thermospheric parameters important to establish conductance, heating, current closure, precipitation, neutral wind speeds, etc. Or, alternatively a pair of LEO polar spacecraft could be placed relatively close to each other in longitude to measure longitudinal ionospheric structure and dynamics at, for example, the dawn and dusk meridians. We have not estimated costs for specific LEO satellites, as we consider GEMINI to be a complete mission without them. However, an even more ambitious mission could include some combination of the LEO examples we have suggested, or others we have not mentioned, and probably at an incremental cost that would remain within the medium size range (\$500M - \$750M).

Modeling

The coupling problem cannot be understood without a strong modeling effort. A cornucopia of quantitative theoretical models now exists that solve basic equations of the inner and middle magnetosphere and their coupling to the ionosphere. Generally, each of these models only include a subset of the relevant physics needed to describe the entire region. For example, some compute energy spectra, pitch-angle distributions, and charge-exchange rates for ring current particles (e.g., Jordanova et al., JGR, 111, A11S10, 2006). Some compute particle drifts self-consistently with electric and magnetic fields, including ionosphere-magnetosphere coupling (e.g., Lemon et al., GRL, 31, L21801, 2004). Global MHD models provide the overall magnetospheric context and some now include ion outflow (e.g., Gloer et al., JGR, 114, A05216, 2009) and coupling to active ionosphere-thermosphere models (e.g., Raeder et al., Space Sci. Rev., 141, 535, 2008). Much effort has been devoted to coupling these models together (e.g., DeZeeuw et al., JGR, A12219, 2004; Moore et al., JGR, 113, A06219, 2008). At this point, it would be easy to configure solar-wind-driven coupled models that could compute nearly all of the parameters to be measured in the GEMINI mission, including parameters measured by the GEMINI spacecraft themselves as well as ground observations and complementary spacecraft programs. Of course, the GEMINI images could help determine which model combination is most accurate. The broad central challenge for modeling now is to figure out how to adjust and upgrade the models to bring them into optimal agreement with observations made at many different points by many different instruments. The global images provided by GEMINI of the plasmasphere, ring current, and innermost plasma sheet will help enormously, because comparison with images will show how well the models are doing throughout a large region of the magnetosphere that is central to the most important magnetospheric events, including substorms and storms as well as sawtooth and steady-magnetospheric-convection events. Ultimately, models constrained by imaging data will lead to better predictive capabilities in which upstream solar wind conditions can be used to forecast conditions throughout the magnetosphere-ionosphere-thermosphere system. This will not be verified without the global scale constraints that a mission like GEMINI can provide.

Evaluation Criteria

High Priority: The basic mission design and science objectives for this GEMINI mission have appeared as specific missions and science concepts in the previous Solar and Space Physics Decadal Survey and the 1997 SEC Roadmap (as SMI), in the 2003 NASA Roadmap (as GSRI), and in both the 2005 and 2009 NASA Heliophysics Roadmaps (as GEMINI).

Contribution to Multiple Panel Themes: GEMINI contributes both to the Solar Wind-Magnetosphere Interactions theme as well as to the Atmosphere-Ionosphere-Magnetosphere Interactions theme of the current Heliophysics Decadal Survey. Coupling and feedback between and within geospace regions (the atmosphere, ionosphere and magnetosphere and the solar wind) are fundamentally important in determining the response of the Earth's near-space environment to the high-energy electrified plasmas and magnetic fields moving outward from the sun. The interplay between geospace components can alter the first order global system response even going so far as to limit the entry of solar wind energy during extreme events. Key science questions for the next decade center on this coupling. A true understanding of Earth's near-space environment requires a new observational approach that provides a systems view of geospace and includes any effects of hemispheric asymmetries on the global response.

Contributes to Important Scientific Questions: The GEMINI science objectives derive from the NASA Heliophysics Roadmap Societal Relevant Science Investigation: How do the magnetosphere and the ionosphere systems interact with each other? And the fundamental questions related to plasma processes that determine how energy and momentum from the solar wind propagate downward through geospace to Earth. Understanding the dominant plasma acceleration and transport processes operative in the Earth's geospace region also has more general applicability. Particle acceleration and particle transport are two of the most fundamental processes in the plasma universe occurring in essentially every region of the heliosphere from the Sun to the heliopause, in the solar wind, in planetary magnetospheres and ionospheres, in cometary tails, and many others. The same processes that have been discovered by Heliophysics missions form a foundation for understanding astrophysical systems.

Complements Other Observational Systems or Programs Available: As discussed above, GEMINI exquisitely complements a multitude of other observational systems, including everything from SDO to upstream solar and solar wind measurements to geosynchronous, GPS, and various LEO spacecraft (both NAS and other agency), as well as many ground based resources (radar, magnetometers, auroral imaging arrays). GEMINI science goals directly benefit from these other systems, and science based primarily on any of these systems is enhanced by what GEMINI can contribute.

Is Affordable: Estimated Cost, FY2010 \$290M to \$405M including 25% reserve, all mission phases; For 10 year mission, cost is estimated \$410M to \$530M

Appropriate Degree of Readiness: No technology developments required. Instrument TRL is high (5 to 9).

Mapping to RFA, Decadal Challenges, and Decadal Missions

This science target primarily falls under research focus areas H2 to understand changes in the Earth's magnetosphere, ionosphere, and upper atmosphere to enable specification, prediction, and mitigation of their effects and F2 to understand the plasma processes that accelerate and transport particles.

This science target is primarily relevant to Decadal Survey Challenges numbers 3, 4, and 5:

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 principals 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.

Spectral Signatures of Geospace Climate Change

Martin G. Mlynczak and Christopher J. Mertens
NASA Langley Research Center, Hampton, VA

Scott Jensen and Roy Esplin
Utah State University/Space Dynamics Laboratory, Logan, UT

Jeng-Hwa Yee
Johns Hopkins University Applied Physics Laboratory, Laurel, MD

Linda Hunt and Jeffrey Mast
SSAI, Hampton, VA

Stanley Solomon and Liying Qian
National Center for Atmospheric Research, Boulder, CO

A White Paper submitted to the Heliophysics Decadal Survey

November 12, 2010

1. Introduction

A fundamental prediction of climate change theory is that the Earth's upper mesosphere and thermosphere will cool over time due to increases in the carbon dioxide (CO₂) abundance in the atmosphere caused by human activity [e.g. *Roble and Dickinson*, 1989]. In the upper mesosphere and thermosphere (which we will call the geospace environment) CO₂ efficiently cools the atmosphere as it essentially radiates in the cool-to-space limit. Increases in the CO₂ abundance caused by fossil fuel burning will increase the amount of infrared emission in the upper mesosphere and thermosphere nearly linearly. This will then increase the cooling of the upper atmosphere, thereby resulting in lower temperatures, all else being equal. This "global cooling" of the geospace environment is in contrast to the "global warming" expected in the troposphere and at the Earth's surface due to increases in CO₂.

Increasing CO₂ concentrations are expected to affect the energy balance, composition, and dynamics of the upper atmosphere. As the geospace environment, particularly the thermosphere, is the first region in which the Sun begins to exert an influence on the Earth system, it is critical to understand changes in the climate of this region. In addition, a cooler upper atmosphere, at a given altitude, results in a lower density over time. This effect has significant implications for space operations ranging from satellite orbit lifetime to the tracking of space junk. Thus, in addition to scientific interests, there are substantial operational reasons for understanding climate change in Earth's upper atmosphere. Evidence already suggests that geospace climate change is occurring [e.g. *Marcos et al.*, 2005].

In this white paper we propose the continuation of a fundamental climate data record of the geospace environment now 9 years extant. Specifically, continued measurement of the infrared limb emission from nitric oxide (NO) and CO₂ with new infrared sensors is recommended. These measurements provide the infrared cooling rates, infrared fluxes, and total infrared radiated power that control the thermal structure and climate from 80 to ~ 275 km. Their observation over the last 9 years by the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument on the NASA Thermosphere-Ionosphere-Mesosphere Energetics and Dynamics (TIMED) satellite has led to a detailed understanding of the natural variability of the geospace environment, which is the first step in understanding and determining climate change. Continuing this time series with measurements beyond the current SABER record will provide unequivocal signatures of climate change in the geospace environment over time. We also describe in this paper a new approach to using these data in the development and detection of spectral signatures (sometimes called "fingerprints") of climate change that offer the potential for earlier detection of geospace climate change than traditional parameter retrievals. Lastly we discuss considerations for continuing these measurements with lightweight derivatives of the SABER instrument and the potential to incorporate these measurements as part of other missions.

2. The SABER Infrared Time Series Record

The upper mesosphere and thermosphere cool to space through infrared radiation emitted by carbon dioxide (15 μm), nitric oxide (5.3 μm) and atomic oxygen (63 μm). Carbon dioxide and nitric oxide compose about 90% of the total infrared power radiated from the thermosphere. As such, these two emission features are fundamental to controlling the climate and energy balance of the geospace environment (80 to 275 km). The CO₂ and NO emissions have been measured for nearly 9 years by the SABER instrument (launched December 7, 2001; science operations commenced January 2002) on the NASA TIMED satellite. The instrument has operated nominally recording approximately 98% of possible science data. The CO₂ emissions are routinely observed up to 140 km and the NO emissions have been observed up to at least 275 km during disturbed conditions.

The SABER infrared record to date is remarkable. Figure 1 shows the time series from January 2002 through October 2010 of daily global infrared power (Watts) radiated from the geospace environment (100 to 250 km) by NO and CO₂. The 60-day running mean is also shown. Sixty days is the time required for the TIMED satellite to precess through all local times. Also shown are the daily and 27-day averages of the F10.7 solar radio flux. Figure 2 displays the same data but from January 2009 to the present in order to show more clearly the behavior at the onset of solar cycle 24.

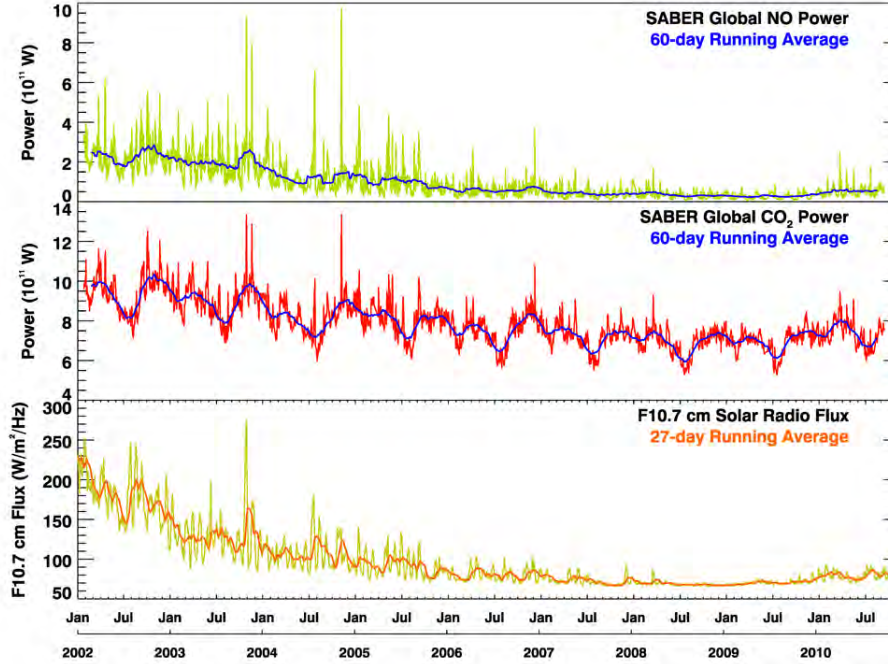


Figure 1. Time series of NO and CO₂ daily infrared power radiated from the thermosphere between January 2002 and October 2010 along with the F10.7 solar radio flux. After *Mlynczak et al.*, [2010a].

Variability in the data on timescales ranging from days to years is evident. The “spikes” in the data are not noise but are in fact the response of the thermosphere to geomagnetic storms. The long-term decline in the infrared power is due to the nominal “11-year” solar cycle, which was very prolonged in cycle 23 through 2008. Variability on annual and semi-annual timescales is visually evident in the CO₂ power. Variability on timescales associated with the solar rotation period and its harmonics is also found in the data [*Mlynczak et al.*, 2008].

To further illustrate the variability of the infrared cooling of the thermosphere we have deseasonalized (essentially a subtraction of the mean power) the NO and CO₂ time series shown in Figures 1 and 2 for each 60-day yaw cycle observed to date, with the results shown in Figure 3. The deseasonalized NO and CO₂ power clearly show the decrease in infrared cooling (radiated power) from the start of the mission in 2002. The decrease continues in both NO and CO₂ until the middle of 2008 at which time a minimum of infrared cooling is reached. Starting in mid- to late 2009 the deseasonalized mean power starts to increase. The increase is indicative of increasing infrared power corresponding to an increase in thermospheric temperatures at the onset of solar cycle 24 [*Mlynczak et al.*, 2010b]. Note that the deseasonalized power is negative some of the time. This is simply a reflection of the fact that the power at some times is less than the mean power. The NO and CO₂ cooling rates shown in Figure 3 have a linear correlation coefficient of 0.985.

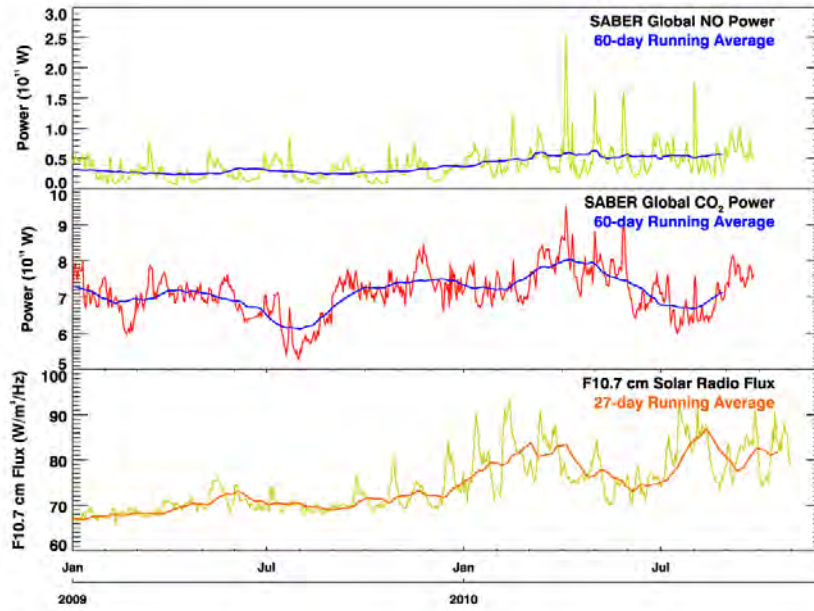


Figure 2. Same as Figure 1 but expanded from January 2009 through October 2010 to show detail. The increase in solar activity is visually evident in the F10.7 solar radio flux and in the NO power 60-day running mean. After *Mlynczak et al.*, [2010b].

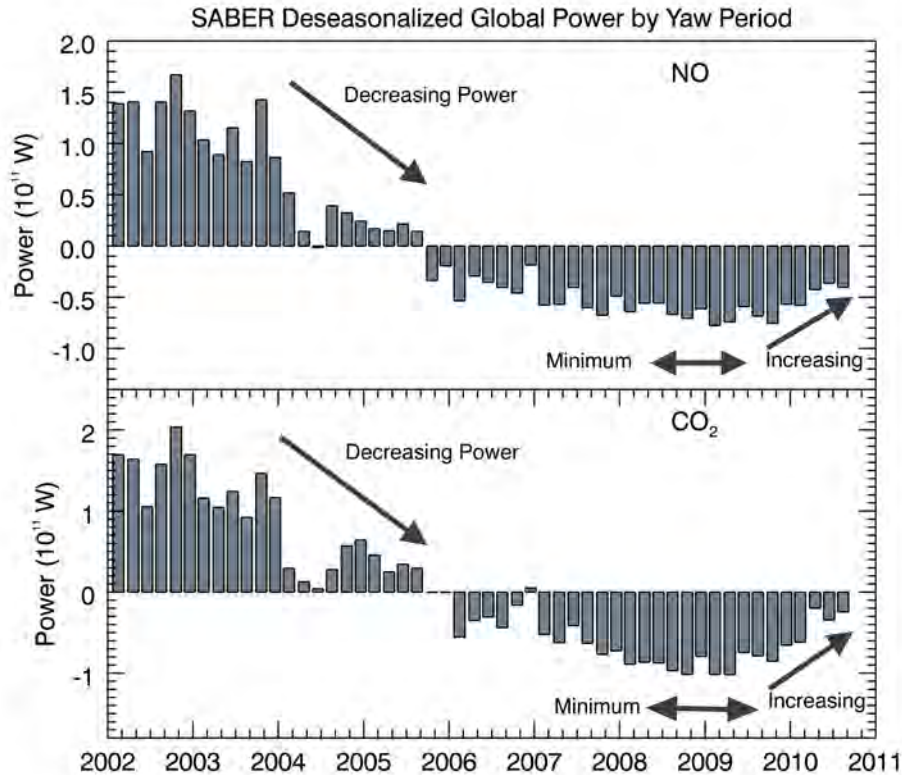


Figure 3. Deseasonalized time series of infrared power from NO and CO₂ in the geospace environment. The data indicate increasing infrared power (increased infrared cooling) since ~ mid 2009, which correlates with increasing solar activity. After *Mlynczak et al.*, [2010b].

From the perspective of measurement of geospace climate parameters, the infrared emissions from which cooling rates, fluxes, and power can be derived are fundamental. The measured infrared limb radiance constrains the derived infrared cooling rates [Mlynczak *et al.*, 2010a]. In addition, these infrared parameters are directly measurable, i.e., they are not generated by inversion of a non-linear radiative transfer equation. Furthermore, the infrared emissions themselves constitute a climate benchmark [e.g., Anderson *et al.*, 2004]. These benchmarks will in time offer a spectral signature associated with climate change in the geospace environment.

3. Infrared spectral signatures of geospace climate change

The climate of the Earth's atmosphere is determined by the energy input from the Sun and the energy radiated from the Earth. Climate change occurs when the energy budget changes due to natural or possibly anthropogenic forcings. These forcings cause the climate system to adjust, sometimes engendering feedbacks that amplify or mitigate the forcings, resulting in a new climate state that is warmer or cooler than the original. The amount of carbon dioxide has been increasing in Earth's atmosphere since the Industrial Revolution began in the late 1700's, providing a climate forcing of anthropogenic origin. Carbon dioxide is radiatively active and is a principal factor in determining Earth's climate from the surface to the thermosphere.

The key question going forward is how to detect geospace climate change. We propose, based on the analysis and understanding of the infrared radiative features observed for the past 9 years, that measurement of the time series of infrared cooling emissions from the thermosphere will provide a distinct spectral signature from which climate change will become obvious and irrefutable, and which can be used to advance modeling and prediction of future climate change.

The measurement of infrared spectral signatures of climate change has been discussed in the literature relating to lower atmosphere (tropospheric) climate for nearly 30 years. Kiehl [1983] and Charlock [1984] noted that distinct spectral signatures would be evident in changes in the infrared radiance emitted by the Earth and its atmosphere, and that these could be observed from orbiting satellites. These spectral signatures could further be associated with the forcing of climate due to the anthropogenic increase of CO₂.

Since the pioneering work of Kiehl and Charlock, a number of other studies have advanced the concepts of spectral signatures of climate change [e.g., Goody *et al.*, 1995; Slingo and Webb, 1997]. Harries *et al.* [2001] showed convincingly the increase in the greenhouse effect due to methane using spectral signatures of infrared radiance spectra taken 27 years apart from satellite spectrometers. More recently, Huang and Ramaswamy [2009] and Huang *et al.* [2010] have shown that distinct infrared spectral signatures associated with specific aspects of climate change (e.g., cloud height, tropospheric water vapor) can be uniquely determined.

Today the concept of using spectral signatures to identify and attribute tropospheric climate change is firmly accepted and is in fact the foundation of the Climate Absolute Radiance and Refractivity Observatory (CLARREO) satellite mission defined in the Earth Science Decadal Survey [National Research Council, 2007]. The CLARREO mission will measure Earth's infrared emission spectrum and the spectrum of reflected solar radiation from the Earth. The primary CLARREO data products are the emitted and reflected radiance spectra. These radiance measurements and the spectral signatures and changes derived from their time series will elucidate the degree of climate change in the troposphere over time. The CLARREO mission is now being developed by the NASA Langley Research Center and at this writing is days away from being confirmed and moved into formulation (Phase A) by NASA.

The ideas relating to the measurement of infrared spectral signatures to determine climate change in the troposphere can be readily and more easily applied to the geospace environment. Spectral separation of the different cooling features is complete in the upper atmosphere as opposed to the troposphere in which water vapor and carbon dioxide emit throughout the infrared

spectrum, thus necessitating the measurement of nearly the entire infrared spectrum for the CLARREO mission. The NO and CO₂ cooling are spectrally distinct (5.3 μm and 15 μm) and are measured unequivocally with little or no interfering emissions [Mlynczak *et al.*, 2010a]. Thus the measurement of these two emission features offers an essentially pure time series of the fundamental processes controlling the climate and energy balance of the geospace environment. These emissions will yield the clear signal of the geospace climate response due to forcing by increases in carbon dioxide. The spectral fingerprint will be manifest as a consequence of increased CO₂ abundance and decreasing temperature. In the long term the increasing CO₂ abundance will result in more measured radiance from the thermosphere despite a modest (~ 10 to 15%) decrease associated with cooler temperatures. In contrast, NO emission will decrease much more rapidly (assuming no change in NO abundances) due to greater temperature sensitivity than that of CO₂. This is a fundamental expectation from infrared radiation physics. The sensitivity of NO to temperature is $\sim \exp(-2700/T)$ while the sensitivity of CO₂ is $\sim \exp(-960/T)$. The greater sensitivity of NO is due to the higher energy associated with its quantum transition at 5.3 μm vs. the energy of CO₂ at 15 μm .

The application of infrared spectral signature determination in geospace climate is an emerging concept and does require more detailed study. We suggest modeling studies especially to investigate the anticipated chemistry changes associated with NO in particular and to determine in more detail the expected signature of NO infrared emission change associated with geospace climate change. Furthermore, the timescales for natural variability and trend detection need to be more fully defined. The approach for doing this with infrared measurements has been developed by Leroy *et al.*, [2008] and the existing SABER data can be used. We emphasize that because the infrared radiance is a direct observable, whose uncertainty is quite small and can be ascertained by on-board calibration, trends in radiance (and derived fluxes and cooling rates) will result in detection and attribution of climate change in much less time than that derived from other remotely-sensed, retrieved parameters, such as temperature or species concentrations. This is because the uncertainty in the radiance is substantially less than in retrieved parameters, and time to detect trends is directly strongly dependent on the uncertainty of the parameter being trended.

4. Instrument Concept

SABER is a filter radiometer that measures infrared radiation in 10 spectral bands. Interference filters in front of the detectors largely define the spectral response of the instrument. In the immediate future a key factor to continuing the infrared climate data record is to have a new instrument overlap the current SABER instrument. Analogous to solar irradiance measurements, differences in instrument calibrations and spectral responses can be best dealt with by having overlap of successive instruments. We note that technology for upper atmosphere measurements has yet to include on-board systems to enable on-orbit *absolute* calibration and spectral response such as those now developed for the CLARREO mission [e.g., Dykema and Anderson, 2006].

With this in mind, the Langley Research Center and the Space Dynamics Laboratory of the Utah State University have been working on a lightweight derivative of the SABER instrument. The goal has been to reduce the mass of the instrument to approximately one-half the value of the TIMED/SABER instrument (i.e., from ~ 75 kg to ~ 35 kg) with little or no loss in sensitivity from the current instrument. A corresponding reduction in instrument power from ~ 75 W to ~ 35 W is also feasible.

In Figure 4 below we show the optical system for the lightweight derivative of the SABER instrument, herein referred to as the Middle Atmosphere Sounder and Thermal Emission Radiometer (MASTER) instrument. It is an on-axis, re-imaging telescope preceded by a plane scan mirror. Its entrance aperture is 100 mm in diameter, and its F/number at the detectors is F/2. A picket-fence, tuning fork chopper is located at the focus of the fore-optics, which are re-imaged at

the detector. As demonstrated by SABER, this optical system configuration has excellent stray-light rejection, having been measured to be as good as an off-axis optical system. This unusual stray-light rejection capability for an on-axis optical system results from an unconventional cold Lyot stop design that blocks stray light from the central obscuration, the spider vanes, and from the aperture stop. This Lyot stop minimizes background photon noise by preventing the detectors from seeing any low emissivity surface except the 75K Lyot stop. The on-axis configuration is also mass efficient, when compared to an off-axis system, due to the inherent strength of its cylindrical shape. The radiometric sensitivity of this on-axis optical system has been shown by analysis to be significantly better than an off-axis optical system.

The largest contributor to mass reduction will result from changing the method of mounting and cooling the telescope. This change accounts for most of the mass reduction of the support structure, the scan and baffle, the re-imager optics, the fore optics, and MLI blanketing assemblies. This change will be made possible by shortening the baffle that precedes the scan mirror in the optical path. Making this baffle shorter in MASTER, when compared to SABER, results in a negligible decrease in stray light rejection in exchange for a very large decrease in mass.

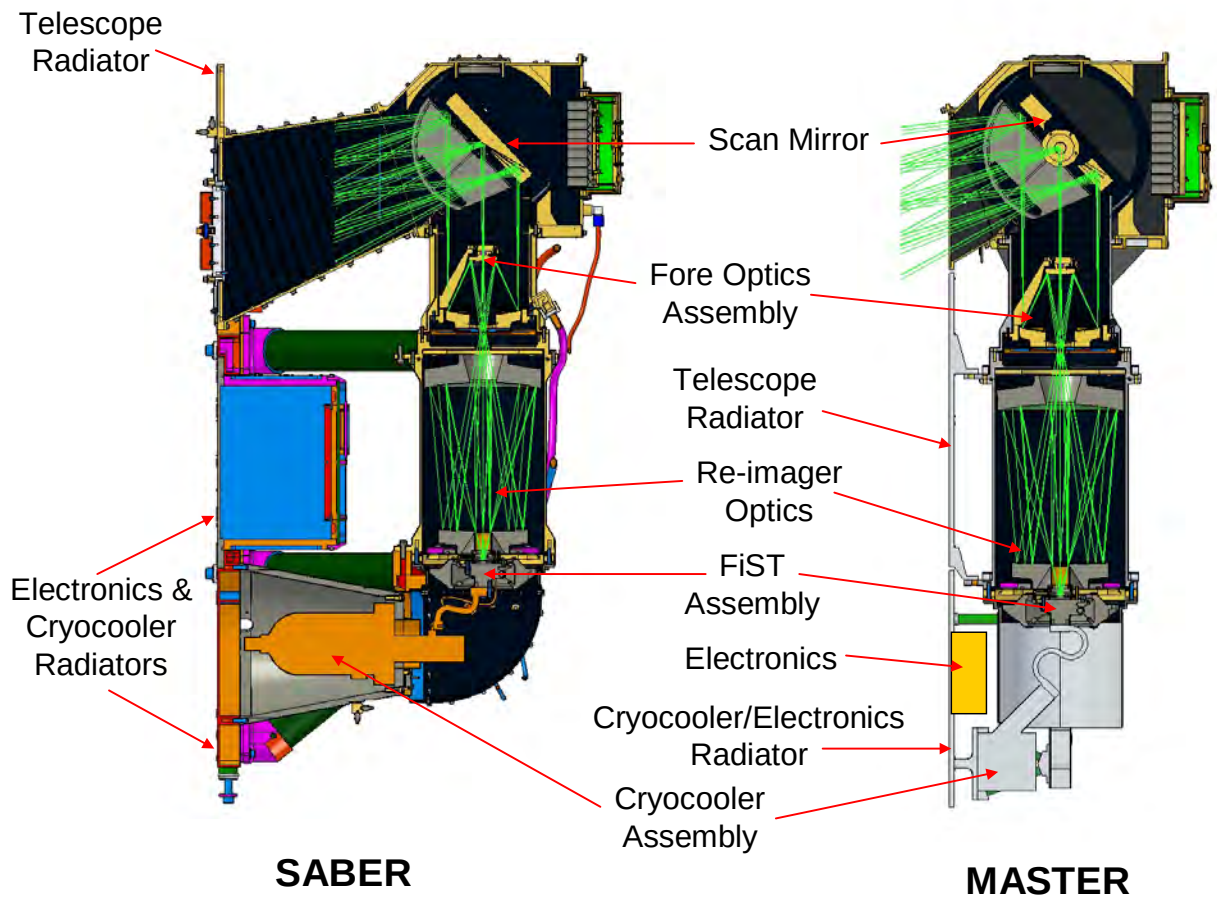


Figure 4. SABER and MASTER Optical System Comparison

The length of the SABER baffle was selected early in the SABER program to get every bit of stray light rejection possible; however, in hindsight, the SABER baffle length could have been shortened with only a negligible decrease in its stray light rejection capability. The shorter baffle in front of the scan mirror will enable the MASTER telescope to be directly attached to its radiator rather than to the isolator legs as was the case for SABER. The cooling path for the MASTER telescope will be radial along the

length of the instrument rather than longitudinal from the back of the instrument to the front as it is in SABER. This change in the cooling path reduces the wall thickness and hence mass required for thermal conduction in the optics. It also reduces the parasitic heat load that further reduces the mass by reducing needed radiator area. Additional mass reductions will be achieved by incorporating composite materials in the scan and baffle assembly and also by light-weighting the mirrors.

The next largest contributor to mass reduction results from updating and improving the packaging of the electronics and from advances in cryocooler technologies over the past decade. The electronics and cryocooler radiator is also reduced in size due to the reduced power needs to operate the next generation cryocooler and electronics. The smaller packages of these components also reduce the parasitic heat loading on the radiators, once again reducing the radiator size from SABER heritage. Finally, radiometric analysis of the MASTER design shows negligible change in predicted noise performance as compared with the achieved SABER flight performance. Anticipated mass and power specifications for MASTER and SABER are given in Tables 1 and 2.

Table 1. MASTER and SABER mass comparison.

Assembly Name	SABER Mass (kg)	MASTER Mass (kg)	Mass Savings (kg)
Support Structure	14.04	1.10	12.94
Instrument Electronics	11.51	2.18	9.33
Cryocooler	11.65	4.02	7.64
Scan and Baffle	15.98	10.94	5.04
Re-imager Optics	7.27	5.59	1.69
MLI Blanketing	2.20	1.50	0.70
Fore Optics	2.33	1.67	0.65
Hardware/Cables	7.71	5.77	1.94
TOTAL MASS	72.69	32.77	39.93

Table 2. MASTER and SABER power comparison.

ASSEMBLY NAME	SABER Power (W)	MASTER Power W	Power Savings (W)
Cryocooler	35.5	16.3	19.2
Heater	13.8	5.0	8.8
Analog Controllers	10.0	6.7	3.3
Signal Processing	5.5	2.3	3.2
C & DH	5.7	4.1	1.6
Housekeeping	1.8	0.6	1.2
TOTAL POWER	72.3	35.0	37.3

5. Summary and Additional Considerations

Geospace climate change is a frontier of scientific exploration. A unique, fundamental climate data record spanning 9 years from SABER now exists. It is proposed that this record be extended, especially the measurement of the NO and CO₂ emissions between 80 and 275 km. It is further proposed that the spectral signature techniques, which are now the foundation for space-based detection of tropospheric climate change, be applied to the infrared observations the geospace environment. These techniques, because they use the measured infrared radiances, offer the potential for detection and attribution of climate change in a much shorter time than conventional parameter retrievals. This is due to the inherently higher accuracy of the radiance measurements as compared to the accuracy of parameters retrieved from the radiance measurements. Lastly, the proposed MASTER instrument, a lightweight derivative of the SABER instrument, can be developed for a flight of opportunity. The MASTER instrument could also be part of a focused mission to study the ionospheric E-region, as discussed in the White Paper by *Mertens et al.* [2010] submitted to the Heliophysics Decadal Survey call.

References

- Anderson, J. G., et al., (2004), Absolute, spectrally resolved thermal radiance: a benchmark for climate monitoring from space, *J. Quant. Spectrosc. Rad. Trans.*, 85, 367-383, doi:10.1016/S0022-4073(03)00232-2
- Charlock, T. P., (1984), CO₂ induced climatic change and spectral variations in the outgoing terrestrial infrared radiation, *Tellus*, 36B, 139-148. 1984
- Dykema, J.A., and J.G. Anderson, (2006), A methodology for obtaining on-orbit SI-traceable spectral radiance measurements in the thermal infrared, *Metrologia* 43, 287-293.
- Goody, R. M., R. Haskins, W. Abdou, and L. Chen, (1995), Detection of climate forcing using emission spectra, *Remote Sensing and Earth Observation*, 5, 22-32.
- Harries, J. E., et al. (2001), Increases in greenhouse forcing inferred from the outgoing longwave radiation spectra of the Earth in 1970 and 1997, *Nature*, 410(6826), 355-357, doi:10.1038/35066553
- Huang, Yi, V. Ramaswamy, 2009: Evolution and Trend of the Outgoing Longwave Radiation Spectrum. *J. Climate*, **22**, 4637-4651. doi: 10.1175/2009JCLI2874.1
- Huang, Y., S. Leroy, P. J. Gero, J. Dykema, and J. Anderson (2010), Separation of longwave climate feedbacks from spectral observations, *J. Geophys. Res.*, 115, D07104, doi:10.1029/2009JD012766.
- Kiehl, J. T., (1983), Satellite detection of effects due to increased atmospheric carbon dioxide, *Science*, 222, 504-506.
- Leroy, S., et al. (2008), Testing climate models with thermal infrared spectra, *J. Clim.*, 21, 1863-1875, doi:10.1175/2007JCLI2061.1.
- Marcos, F. A., J. O. Wise, M. J. Kendra, N. J. Grossbard and B. R. Bowman, (2005), Detection of a long-term decrease in thermospheric neutral density, *Geophys. Res. Lett.*, 32, L04103, doi:10.1029/2004GL021269.
- Mertens C. J., et al., (2010) Ionospheric E-Region Chemistry and Energetics, White Paper submitted to the Heliophysics Decadal Survey
- Mlynczak, M. G., et al. (2008), Solar-terrestrial coupling evidenced by periodic behavior in geomagnetic indexes and the infrared energy budget of the thermosphere, *Geophys. Res. Lett.*, 35, L05808, doi:10.1029/2007GL032620.
- Mlynczak, M. G., et al. (2010a), Observations of infrared radiative cooling in the thermosphere on daily to multiyear timescales from the TIMED/SABER instrument, *J. Geophys. Res.*, 115, A03309, doi:10.1029/2009JA014713.
- Mlynczak, M. G., et al. (2010b), Evidence for Onset of Solar Cycle 24 from Observations of Increasing Radiative Cooling of the Thermosphere, *Geophys. Res. Lett.*, in preparation.
- National Research Council, 2007, Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, 456 pp., http://www.nap.edu/catalog.php?record_id=11820
- Roble, R. G., and R. E. Dickinson (1989), How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere?, *Geophys. Res. Lett.*, 16(12), 1441-1444.
- Slingo, A., and M. Webb, (1997), The spectral signature of global warming, *Q. J. R. Meteorol. Soc.*, 123, 293-307.

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.

Abstract

Active cause-and-effect experiments are an extremely leveraged way to develop scientific understanding of the behavior of space plasmas. The recent sophistication and innovative use of ionospheric heaters, along with the fact that we can repeat the experiments for varying space conditions makes them the mode of choice in using space as an open plasma and geophysics laboratory. The type of experiments that can be conducted include injection of whistler and Alfvén waves in the magnetosphere and the Radiation Belts (RB) to study wave-particle interactions under controlled conditions, creation of Artificial Plasma Layers and their interaction with HF and other waves, stimulated magnetospheric emissions, Alfvén masers and substorm triggering, Langmuir turbulence, parametric interactions, electron acceleration, artificial ionospheric turbulence, optical emissions and artificial ionospheric lenses. It is argued that a comprehensive program requires facilities at high latitudes such as HAARP and EISCAT, mid-latitude such as Arecibo and equatorial in order to access different magnetospheric regions, understand the effect of varying magnetic geometry and E-region conduction, including the role of the Cowling effect in the dip-equator. It is recommended that (i) AMISR be relocated from Poker Flat to HAARP; (ii) Arecibo heater be completed and upgraded; (iii) a heater be constructed in Jicamarca to cover the dip equator; (iv) the US contribute to the launch of the RESONANCE satellite under study by IKI, Russia. Completion of the above tasks requires a capital cost of approximately \$ 10 M. Maintenance of the facilities and funding of a vigorous active experiments research program will require an annual budget of \$ 10 M. This is an extremely leveraged low cost program in support of the US space sciences. It will provide excellent student training in physics and engineering in addition to its major science contribution. NSF, NASA, ONR, AFOSR and DARPA could provide coordinated funding.

1.0 Introduction:

Active experiments, where the environment can be perturbed with controlled injections of particles, waves, beams, ionized or neutral gases and then diagnosed to determine the response of the space plasma have for a long time provided the opportunity to perform laboratory-style, cause and effect experiments in space. Such experiments have been extremely valuable in aiding our understanding of basic plasma physics processes because they provide a great control on the injected perturbation and if diagnosed properly can address sharply focused questions. In the past a number of such experiments, such as electron beam injection (ARAKS, ZARNITSA, EXEDE, PRECEDE,...), neutral gas injection (CRRES, AMPTE,...), Electrodynamic Tethers (TSS1 and 1R), wave injection (AKTIVNY, WISP,...) and Shuttle/Spacelab provided stimulating results on key space physics issues. Unfortunately during the last two decades active experiments of such type have been eliminated, with the exception of some rocket based ones. The major drawback of such active experiments is that they were

essentially one-shot type experiments and therefore had a relatively low cost-to-benefit ratio.

The advent of and extreme sophistication of the “Ionospheric Heater” as evidenced by the Ionospheric Research Instrument (IRI) of the High Frequency Active Auroral Research Program (HAARP) and the EISCAT HEATER are ushering the active experiments and the use of space as an “open” plasma laboratory into a new era. Although ionospheric heaters operated in Arecibo, Platteville and HIPAS, it is the sophistication and power of modern phased-array type heaters that allows us to control the form of temporal structure of the perturbation injected in the ionospheric and magnetospheric plasma and in conjunction with modern ground and space based instrumentation to resolve critical problems in space plasma physics.

This white paper argues for

1. Completing the diagnostic instrumentation of the HAARP by adding Incoherent Scatter Radar (ISR) capability. This can be accomplished by moving the Advanced Modular ISR (AMISR) from Poker Flat to the HAARP site.
2. Completing the Arecibo heater
3. Constructing a heater in Jicamarca, Peru, a location in the dip equator.
4. Providing satellite diagnostic support to high latitude heaters from a satellite system in magneto-synchronous orbit that ensures that the satellite stays in the same flux tube over significant time periods. Alternatively the US might assist in the launch of the RESONANCE project that has been proposed by IKI in Russia.

With the first three items we will have available heaters and complete diagnostic instrumentation for three geomagnetic configurations, high, middle and dip equatorial latitudes while a RESONANCE type satellite in conjunction the Radiation Belt Probes and other satellites will provide space based diagnostics.

The remaining of the paper is organized as following: Section 2.0 briefly reviews recent investigations using ionospheric heaters and outlines the type of active experiments that can be performed with the suggested enhanced capability. The final Section 3.0 estimated budgetary requirements.

2.0 Results of Current Investigations and Implications to Future Studies:

2.1 ULF/ELF/VLF Wave Injection – Studies of Wave-Particle Interactions in the Radiation Belts under Controlled Wave Injection:

Both HAARP and EISCAT are located in ionospheric regions permeated by ionospheric currents, known as electrojet currents. ULF/ELF/VLF waves been generated in the ionospheric E-region during electrojet periods by injecting amplitude modulated HF waves from the heater. The resultant temperature modulation of the

heated electrons modulates the ionospheric conductivity resulting in the formation of a virtual antenna that injects waves at the modulation frequency in the earth-ionosphere waveguide as well as in the magnetosphere. Frequencies from .01 Hz to 20 kHz have been thus generated and detected as far as Guam 8000 km away. An important scientific study concerns the interaction of waves in the VLF range, 1-6 kHz with energetic electrons in the Radiation Belts. It has been found that if the injected amplitude of the VLF waves exceeds a threshold value an extremely strong non-linear interaction occurs that triggers emissions in other nearby frequencies and results in amplification of the wave power by 20-30 dB. An example showing amplification of more than 10 dB/sec obtained in a recent experiment using HAARP is shown in Fig.1. This phenomenon, known as Artificially Stimulated Emissions (ASE), in addition to being one of the most challenging issues in non-linear plasma physics, has enormous implications in understanding hiss and chorus in planetary systems and the dynamics of electron lifetimes in the Radiation Belts (RB). These phenomena can only be studied by active experiments using ionospheric heaters along with space based and conjugate region diagnostics.

Recent HAARP experiments, conducted under DARPA's BRIOCHE program, produced a new and surprising result. Namely, waves in the ULF/ELF range with frequencies below 50 Hz can be generated without the need for an electrojet. The physics of the process is illustrated in the simulation shown in Fig.2 while the observed waves are shown in Fig.3. Modulating the electron pressure of the F-region generated these waves. The process generates directly Magneto-Sonic (MS) waves that propagate laterally in the Alfvénic waveguide while Shear Alfvén waves (SA) are generated when the MS waves reach the E-region where the Hall conductivity dominates. The SA waves are injected upwards in the magnetosphere where they can interact with both the trapped electrons and protons. This discovery opens up a new research area by using the upcoming Arecibo heater to inject waves SA waves from 1-50 Hz into the inner RB and study their interaction with the protons under controlled conditions. The SA waves are generated when F-region driven MS waves reach the E-region and are then injected along the magnetic field tube in the inner belt (Fig.4). Finally the same F-region generation will allow MS wave injection upwards in the magnetosphere from a Jicamarca heater, allowing thus the first study of RB particle – MS wave interactions under controlled conditions.

Other relevant studies within the same context can address triggering of the Alfvén Maser by modulation of the F-region reflection coefficient (Fig.5) as well as study the impact of ionospheric conduction in triggering substorms.

2.2 Artificial Ionization:

Creation and control of artificial plasma layers in the ionosphere has been extremely challenging scientifically. Furthermore, the creation of such controlled layers will provide a new capability of testing propagation and scattering of HF waves, created by ground based HF transmitters, through the modified ionospheric plasma. Recent results at HAARP in experiments conducted by Todd Pedersen of AFRL created for the first time Artificial Plasma Layers (APL) as shown in Fig. 6. Potential experiments using such APL are illustrated in Fig.7. Notice that both the generation of APL as well as the

propagation of HF or other waves through those layers depends critically on the magnetic geometry. Having mid-latitude and equatorial heaters with appropriate diagnostics will allow us to study the role of the magnetic geometry in APL type experiments.

3.3 Langmuir Turbulence-Parametric Instabilities – Ionospheric Irregularities - Electron Acceleration-Optical Emissions:

Studies of the above topics using the old Arecibo heater and the EISCAT heater have been ongoing for some time. However, the first observation of Stimulated Brillouin Scattering (SBS) was accomplished only very recently at HAARP (Bernhardt et al., PRL 104, 165004, 2010) and is evidence of the importance of heater ERP and diagnostic instrumentation. Fig. 8 shows an example of the results. Many issues such as the type of instabilities leading to ionospheric irregularities and ionospheric turbulence, their saturation, the role of the magnetic field inclination, F vs. E region processes, the resultant electron acceleration and optical emissions, scintillations – generation and suppression are all frontier space plasma physics topics that can only be addressed using the ionosphere as an open laboratory (fig. 9).

3.4 Artificial Ionospheric Lenses

The possibility of creating artificial lenses by modifying the index of refraction of the ionosphere by HF heating has been discussed theoretically for a long time. Gennady Milikh of the University of Maryland obtained the first experimental evidence of the effect in recent HAARP experiments, under DARPA's BRIOCHE program. As seen in Fig. 10 more than 30 dB focusing was required to saturate the HF receiver located on DEMETER while overflying the HAARP site at a 650 km altitude. Again magnetic geometry is an important ingredient in the type, location and amplification capability of such artificial lenses.

3.0 Budgetary Estimates:

Two separate budget categories are required for the project. First the capital requirements for completion the facilities and the required diagnostics:

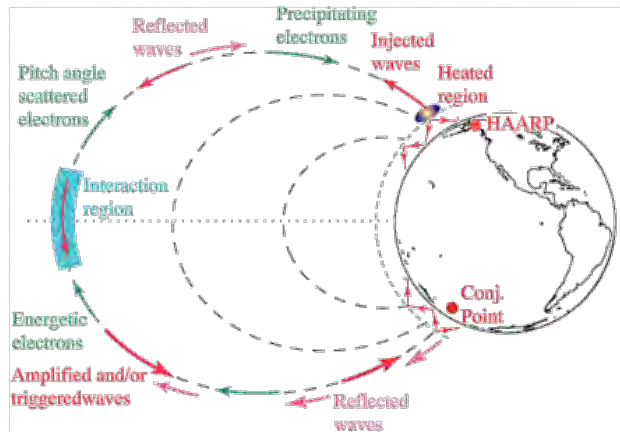
1. HAARP: The most important requirement for HAARP is an ISR. Moving the Poker Flat AMISR to HAARP rather than Argentina can be accomplished with less than \$ 2 M.
2. Completion and possible upgrade of the Arecibo heater approximately \$1-2 M
3. Construction of a heater in Jicamarca using similar construction and GFE equipment with Arecibo \$ 2-3 M
4. Contribution to IKI for launching of RESONANCE \$ 4M

This gives a total of approximately \$ 10 M

Second is the requirement of upkeep of the facilities and stable support of the research. A modest estimate is \$ 10M/annually.

Notice that this amount can be contributed by the key civilian and DoD basic research agencies. Namely NSF, NASA, AFOSR, ONR and DARPA.

FIGURES

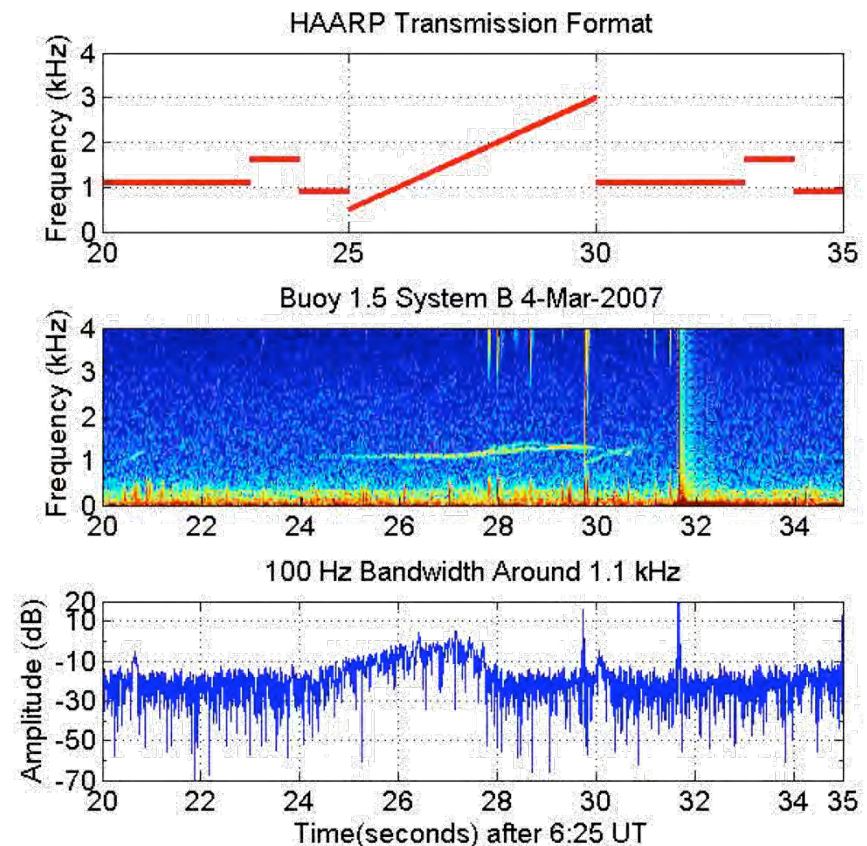


(a)

Fig.1: ASE observed during Stanford experiments using HAARP (Courtesy of U.Inan)

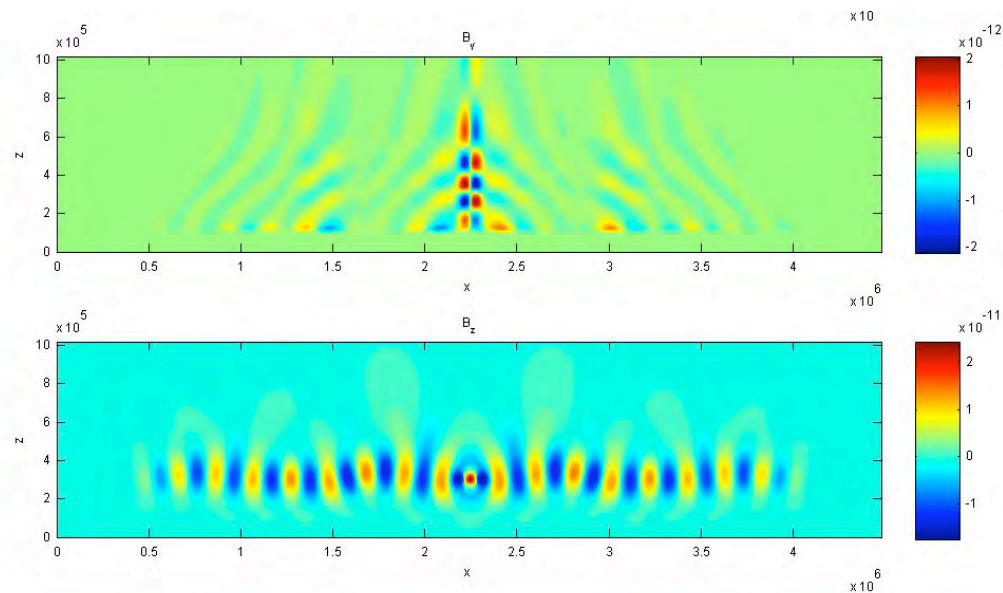
(a) Notional concept indicating the process of VLF injection by HAARP, amplification due to the interaction with trapped electrons and detection at the conjugates.

(b) Transmitted HAARP signals (upper panel); conjugate reception of triggered emission of 1.1 kHz pulse (middle panel); detected signal amplification (lower panel)

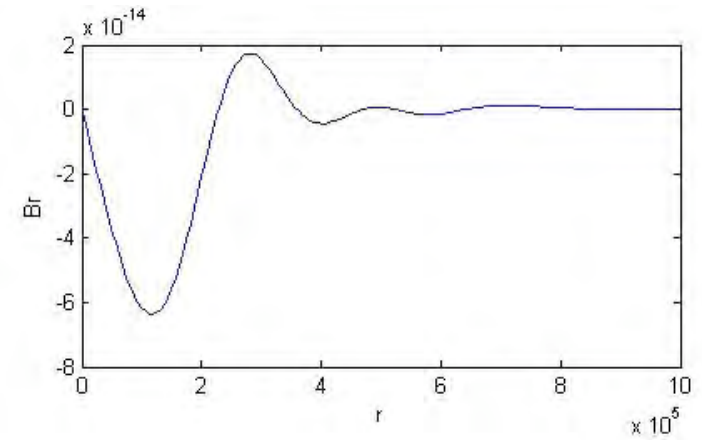


**Only the pulse at 1100 Hz is amplified
10 dB/sec**

(b)



(a)



(b)

Fig.2 Computer simulation of ionospheric current drive by F-region 2 Hz modulated heating at F-peak located at 300 km and for vertical magnetic field :
 (a) MS field (B_z component) propagating laterally 5 secs after source turn-on (lower panel); SA field (B_θ) generated when the MS wave reaches the E-region and confined on the original flux tube and injected in the magnetosphere (upper panel)
 (b) Magnetic field on the ground 2 sec after turn-on

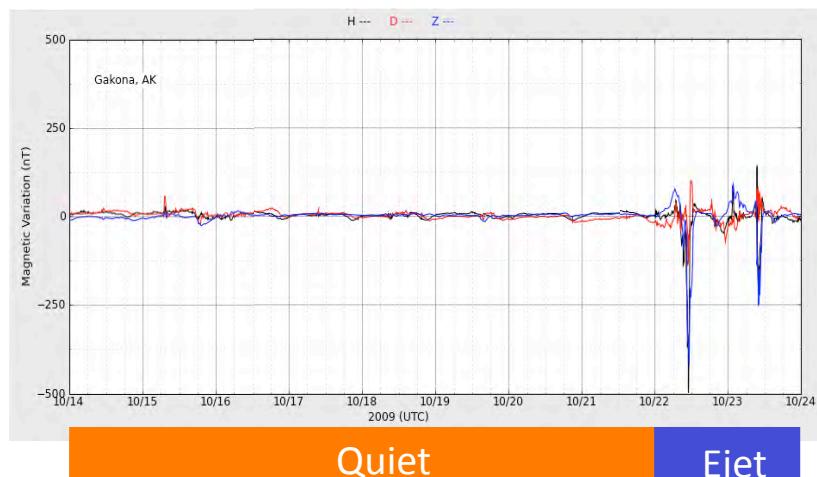
Simulations by University of Maryland (MURI project)
 Papadopoulos and Gunerov

Oct.14-21, *ICD* was the source

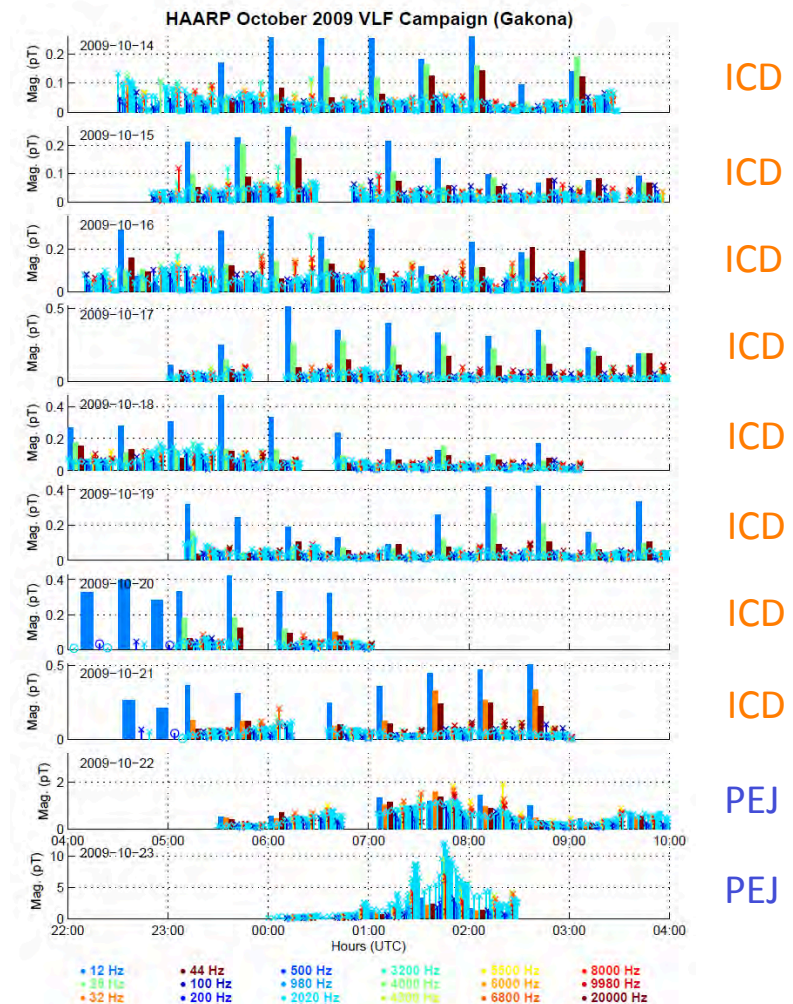
No electrojet, quiet ionosphere
Consistent daily ELF production, <
0.5 pT

Oct. 22-23, *PEJ* was the source

Active electrojet
Spur of ELF/VLF production



(a)



(b)

Fig.3: Experiments indicating Ionospheric Current Drive (ICD) vs. Polar Electrojet (PEJ) modulation – (C.L. Chang, BRIOCHE program): (a) Electrojet amplitude measured at Gakona. (b) Waves measured on the ground at various frequencies. ICD generates only frequencies below 50 Hz, while PEJ all frequencies up to 20 kHz.

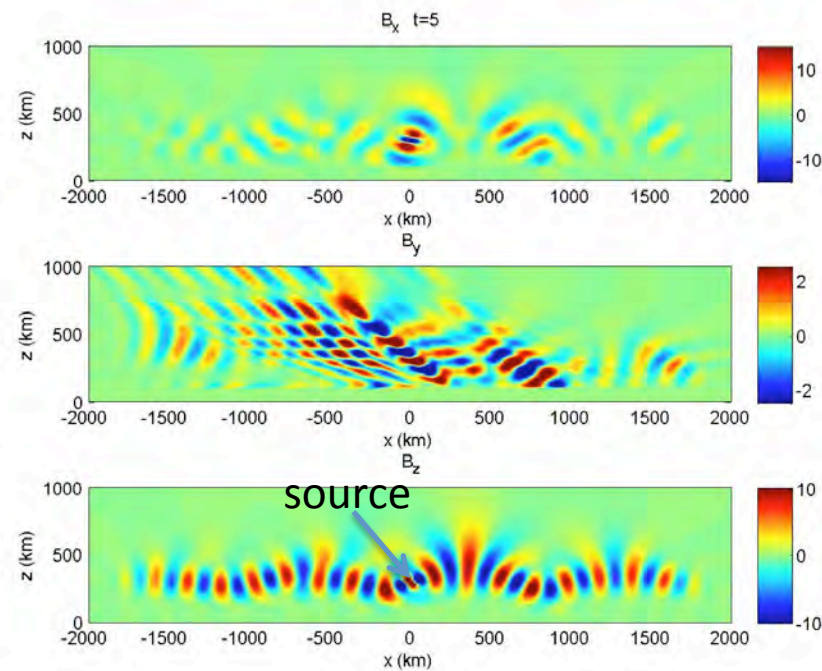
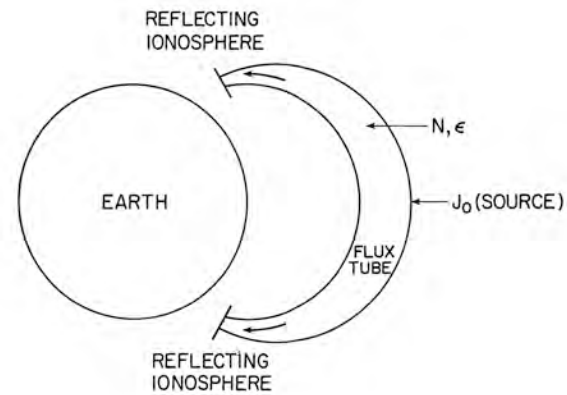
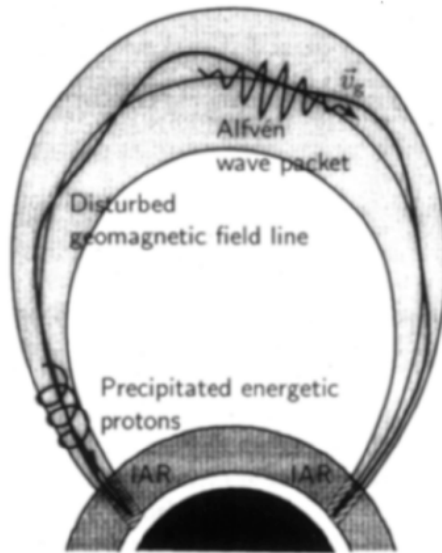
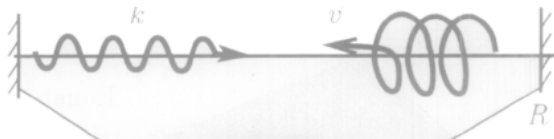


Fig. 4: Simulation of ICD at for oblique injection (B at 45 degree angle to vertical) (University of Maryland MURI program; Eliasson-Papadopoulos). Source similar to Fig.3. MS wave (B_z) propagates mainly laterally (lower panel) while SA (B_y) (middle panel) is injected mainly in the flux tube that goes through the closest bounce of the MS wave with ionospheric E-region.



- Inverted population $\rightarrow$ Proton anisotropy
- Reflection coefficient $\rightarrow R$ at F-region
- Gain depends on inversion extent and R
- Non-linearity due to effect of proton precipitation on reflection coefficient

Fig. 5: Schematic and elements of Alfvén Maser. Maser action is pumped by using HF heating to modulate the reflection coefficient at F-region heights $R(t) = R_0[1 + \cos(\omega t)]$ (Demekov et al., Adv. Space Res. 335, 2003)



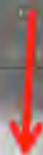
Strong Interaction in Bright Spot at ~160 km Altitude



Todd Pedersen

17 MAR 09
05:16:46 UT
5600A
13.000 S
-21.4 C

Central Spot



Ring

557.7 nm emissions
viewed up the magnetic
field at 5:16:46 UT

- Spot in center of ring but displaced by ~50 km downward in altitude
- Echoes at 160 km correspond to blob of artificially produced ionization

HAARPOON 2

Delta Junction, AK

Ring



Vertical displacement of ~50 km



Central Spot

557.7 nm emissions
viewed from the side at
5:16:46 UT

17 MAR 09
05:16:46 UT
5600A
13.000 S
-21.4 C

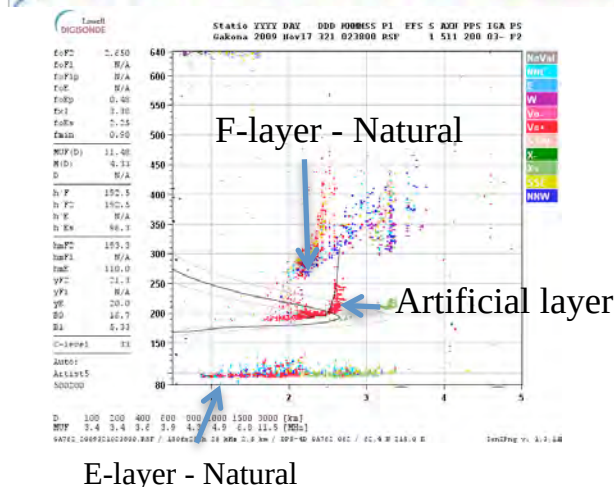


Fig. 6: First evidence of creation of Artificial Plasma Layers (APL) by Todd Pedersen of AFRL using HAARP

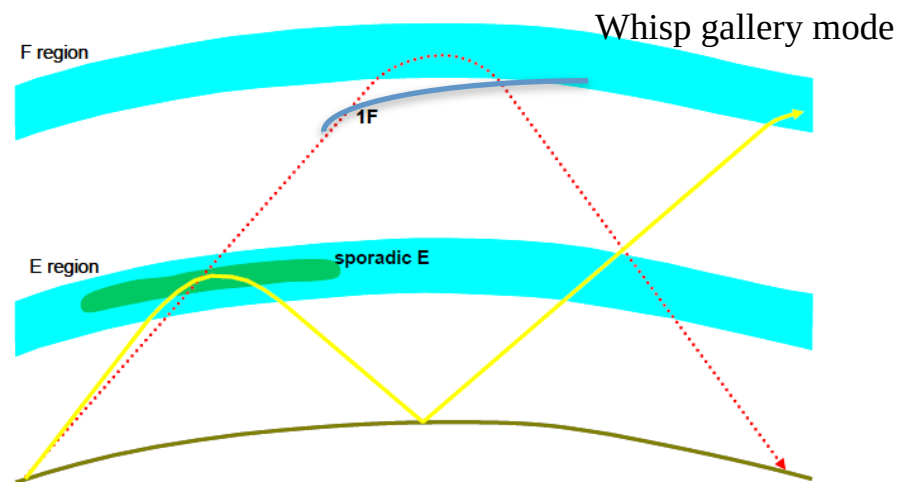


Fig. 7: Schematics of experiment concepts using APL to study their interaction and effects on HF propagation

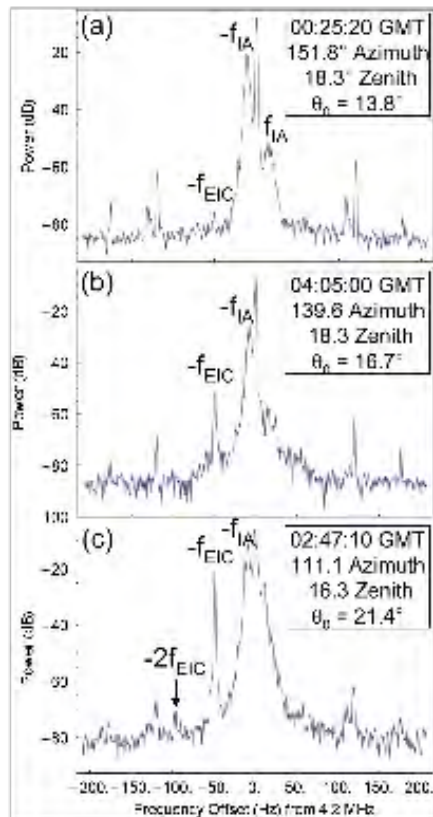


FIG. 1 (color online). Low-frequency SEE spectra showing ion-acoustic and electrostatic ion cyclotron spectral lines. When the beam is offset from the magnetic field by (a) 13.8 degrees, a strong IA and an extremely weak EIC line are produced. Increasing the offset angle to (b) 16.7 degrees increases the strength of the EIC line by 24 dB. With (c) 21.4 degree injection angle, the EM pump wave yields an additional 29 dB increase in the EIC emission.

Fig. 8: First detection of SBS in the ionosphere (Bernhardt et al., PRL 104, 165004, 2010)

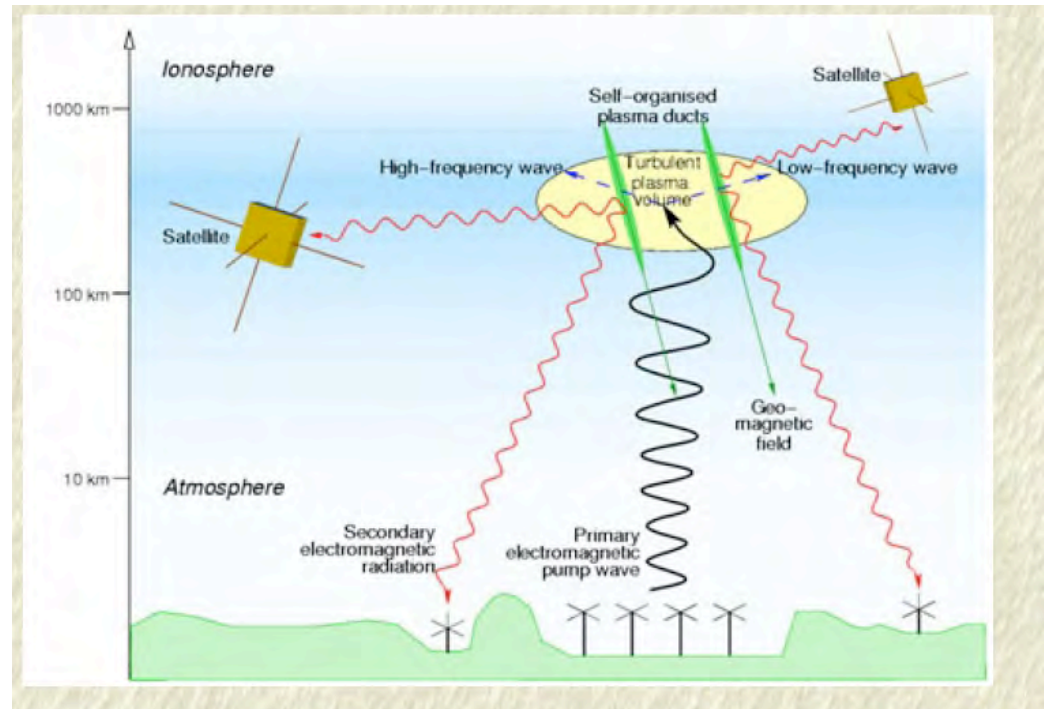
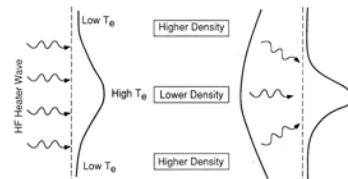


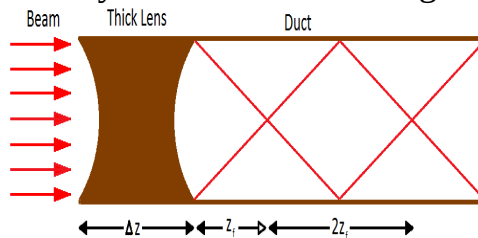
Fig.9: HF heater induced modifications to study key non-linear turbulent processes and induced electromagnetic emissions in space plasmas (after M. Rietveld)

Plasma Lens – Thin lens



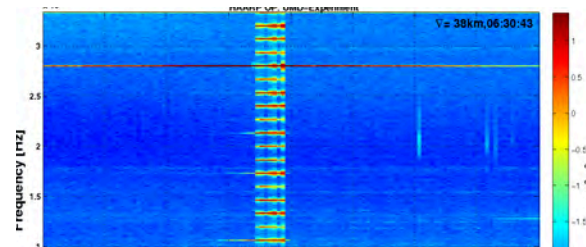
Never confirmed

Physics of Duct Focusing



(a)

Demeter Data



More than 30 dB focusing

(b)

Fig. 10: (a) Concepts of thin and thick (duct focusing) ionospheric lenses
 (b) Detection of 2.8 MHz HF by the DEMETER satellite overflying HAARP indicating saturation of the detector over the heated duct. More than 30 dB focusing is required to saturate the detector. (HAARP experiments during BRIOCHE campaign by G. Milikh, University of Maryland)

Mechanisms of Energetic Mass Ejection (MEME)

T. E. Moore¹, L. Andersson², C. R. Chappell³, G. I. Ganguli⁴, T. I. Gombosi⁵, G. V. Khazanov¹, L. M. Kistler⁶, D. J. Knudsen⁷, M. R. Lessard⁶, M. W. Liemohn⁵, J. P. McFadden⁸, A. F. Nagy⁵, W. K. Peterson², C. J. Pollock¹, A. J. Ridley⁵, D. E. Rowland¹, R. W. Schunk⁹, J. L. Semeter¹⁰, R. J. Strangeway¹¹, J. P. Thayer¹², R. Winglee¹³, A. Yau⁷

1. Executive Summary

MEME will achieve the overarching objective: “Origins of Near Earth Plasmas; to understand the transport of terrestrial gas and plasma from its atmospheric source into the Magnetosphere and downstream Solar Wind” [2009 Heliophysics Roadmap]. Plasma of atmospheric origin is widely recognized as a critical constituent of magnetospheric dynamics, providing the primary source of plasma for the ring current and plasma sheet during dynamic conditions. The key observables are the signatures of the process, involving several mechanisms, by which gas and plasma interact and are transported and heated or accelerated such that they escape Earth's gravity and are ejected into the Magnetosphere and downstream Solar Wind. Candidate mechanisms include convection into auroral regions, auroral kinetic and electromagnetic energy fluxes into the atmosphere, Joule dissipation through ion-neutral collisions, imposed convection and resulting ion pickup, plasma wave generation, and resultant plasma heating and acceleration. Plasma physical mechanisms are fundamental aspects of atmospheric energetic mass ejection from gravitational confinement, anywhere in the universe. MEME will provide comprehensive multi-point fields and plasma observations from the topside ionospheric ejection region, completing our knowledge of these mechanisms. MEME is essential to build predictive models of the energy transfer between drivers and loads, and the resultant mass ejection fluxes needed in support of predictive global circulation models of geospace weather.

2. Description

The MEME science problem comes down to understanding the varying degree to which terrestrial material mass loads the hot Solar Wind plasma and magnetospheric energetic particles [Szégo et al., 2000]. Known terrestrial plasma escape mechanisms include the polar wind [Barakat & Schunk, 2006]; the plasmasphere and its convection plumes, with their topside ionosphere “footprints”, consisting mainly of light ion outflows; and auroral ionospheric heating mechanisms, which eject highly variable heavy ion outflows including O^+ and molecular ions into the steady and pervasive polar wind light ion outflows [Moore & Horwitz, 2007; Lotko, 2007]. Reconnaissance opens the Magnetosphere to exchange of solar and terrestrial plasmas and energy, to a degree that increases as the Solar Wind magnetic field turns southward and magnetospheric convection activity increases. Yet, paradoxically, the more open the Magnetosphere, the more it becomes mass loaded with terrestrial material, mainly the heavy O^+ component [Kistler et al. 2006]. This basic fact argues strongly that reconexion of the Solar Wind to the Magnetosphere and Ionosphere produces more ablative ejection of terrestrial plasma than entry of solar plasma. This escape process has been extensively observed and hypothesized, but its component mechanisms remain poorly understood because of the challenges of comprehensively observing low energy plasmas and fields, combined with the space-time ambiguities of single point observation.

Known auroral mass ejection mechanisms include ionization and heating by soft electron precipitation, with resultant enhancement of the ambipolar electric field; Alfvén waves propagating into the Ionosphere from turbulent magnetospheric boundary layers; current-driven instabilities leading to lower hybrid waves in the auroral acceleration region; and

Mechanisms of Energetic Mass Ejection

convection and shear driven instabilities along auroral flux tubes, including pick-up ions in the topside [Wilson, 1994; Moore & Khazanov, 2010; Paterson and Frank, 1989], and Joule-frictional heating in the collision dominated F region. Much of the energy is dissipated in the neutral gas, leading to observable atmospheric upwelling features above the auroras, including the ejection of molecular ions. Auroral heavy ion outflows are observed to be dominated by superthermal (eV to 10's eV) ions that are transversely heated and whose flux rises in a power law relationship to the incident DC and-or AC electromagnetic energy fluxes and the auroral production of soft electrons [Strangeway et al., 2005; Zheng et al., 2005; Zeng & Horwitz, 2007]. As summarized in *Figure 1*, these facts may be consistent with lifting of ions by the ambipolar electric field, enhanced by soft electron precipitation [Burchill et al., 2009], combined with heating and-or ponderomotive acceleration of the ions by broadband waves in the cyclotron frequency range. The ambipolar electric field can be derived from hot plasma precipitation rates, but the source of ion resonant waves is indeterminate, so neither their amplitudes nor heating rates are derivable from macroscopic disturbance conditions such as magnetohydrodynamic (MHD) flow, current, and thermal conditions.

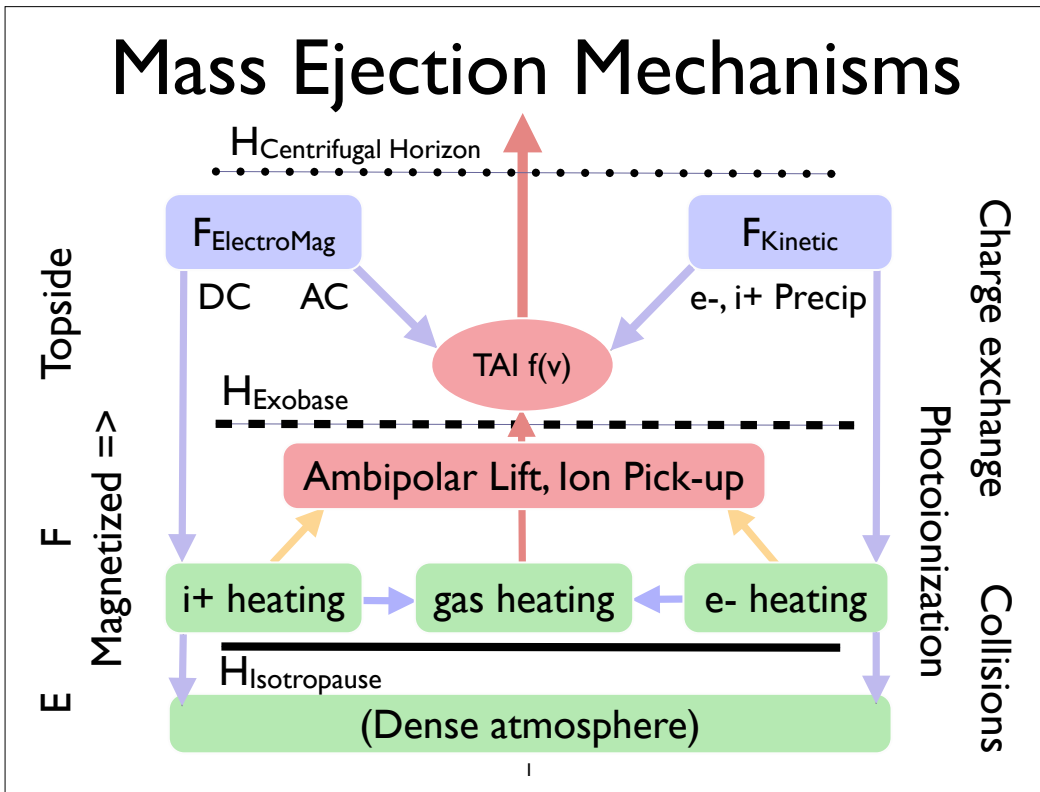


Figure 1. Schematic diagram of mass ejection mechanisms after Moore and Khazanov, [2010 JGR] and Strangeway et al., [2005 JGR]

The Outstanding Question is then: "What are the mechanisms by which Solar Wind energy flux is dissipated in planetary gas and plasma to produce ejected mass flux?" To answer this question, observations are needed to provide a comprehensive picture of ionospheric mass ejection, including: i) detailed observations of the 3D energy-angle distribution of transversely accelerated ions and electrons down to zero energy (≤ 0.1 eV); ii) control of sensor potential at or near the plasma potential; iii) true velocity measurements based on composition analysis; iv) observation of plasma wavelengths and frequencies for definitive mode identification; v) direct observations of auroral neutral gas upwelling to complement recent CHAMP and other

Mechanisms of Energetic Mass Ejection

spacecraft accelerometer observations. The successful mission design will deliver simultaneous conjugate diagnostics from the F region (200-500 km), from the topside exobase region (500-1500 km), and from the auroral acceleration region (1500-3000 km). Advanced incoherent scatter radar (AISR) provides needed diagnostics of the F region. As illustrated in *Figure 2*, two identical spacecraft, placed in high inclination 500 km x 3000 km orbits, with oppositely oriented semi-major axes through the radar sampled zone, can be maintained out of phase so they pass over the auroral zone in conjunction. In this configuration, one would be near apogee and the other near perigee, with alternating placements in the northern and southern auroral zones, with fixed orientation in inertial space so that a scan of all local times is performed each year.

Measurement Requirements

We seek the operational capability to predict ionospheric mass ejection fluxes from imposed Solar Wind conditions, via global simulation models capable of computing the energetic excitation of the Ionosphere. The key element now missing from our understanding of this process is the ability to compute ionospheric response to those excitations, including the generation of instabilities that turn a portion of the available energy into ionization, heating, and acceleration of ionospheric gas and plasmas. MEME will achieve this overarching goal, advancing Heliophysics toward improved operational capabilities, as follows:

1. Determine the excitation of the global topside Ionosphere by solar and magnetospheric energy inputs.
 - a. Measure the DC and AC Poynting flux into the Ionosphere from high altitude energy sources, in the range $0.01\text{-}100\text{ mW m}^{-2}$ with a resolution of 0.1 sec (1 km).
 - b. Measure the frequency and wave vector spectrum associated with significant AC Poynting flux to determine the wave modes, propagation direction, and identify source regions.
 - c. Measure the kinetic momentum and heat flux into the Ionosphere from high altitude particle sources, from thermal to auroral energies (0 eV - 20 keV).
2. Determine the response of the global topside Ionosphere to solar and magnetospheric forcing.
 - a. Measure the local plasma wave environment generated by solar and magnetospheric forcing and the relaxation of auroral electron beams and pick-up ion ring beams.
 - b. Determine the capability of such waves to produce superthermal power law tails in relevant velocity distributions.
 - c. Measure the plasma upflow and outflow in the magnetized collisional zone (125-500 km), the exobase transition zone (~500-1500 km), and the centrifugal forcing zone of the auroral acceleration region (~1500-3000 km).
 - d. Measure thermospheric upwelling and fast atom emission in response to auroral energy inputs throughout, and determine its effect on collisionality and exobase height.
3. Use physical models of the Ionosphere to predict the responses from the excitations observed, and refine these models to the point where they are demonstrably successful.

3. Cost

MEME requires a multi-point *in situ* observation program allowing simultaneous observations in the magnetized collisional F region, in the topside exobase transition zone, and in the region of highly structured field-aligned currents and shear flows of the auroral acceleration region. As such its cost will be comparable to the inflation-adjusted cost of previous multi-spacecraft

Mechanisms of Energetic Mass Ejection

missions such as the International Sun Earth Explorer (ISEE), Cluster, and Time History of Events and Microscale Interactions during Substorms (THEMIS), albeit with fewer spacecraft, and therefore lower costs. The closest cost comparison may be with the Dynamics Explorer-1 and -2 mission (DE-1,-2), although MEME's emphasis is on lower altitude phenomena associated with the ionosphere, in contrast with DE-1's focus on of the plasmasphere or polar wind. Achieving the synchronized opposing major axis orbits over the appropriate auroral radar facilities and scanning all local times may consume the apparent propulsion savings over higher altitude orbits, so this orbital strategy requires study to determine its feasibility and cost.

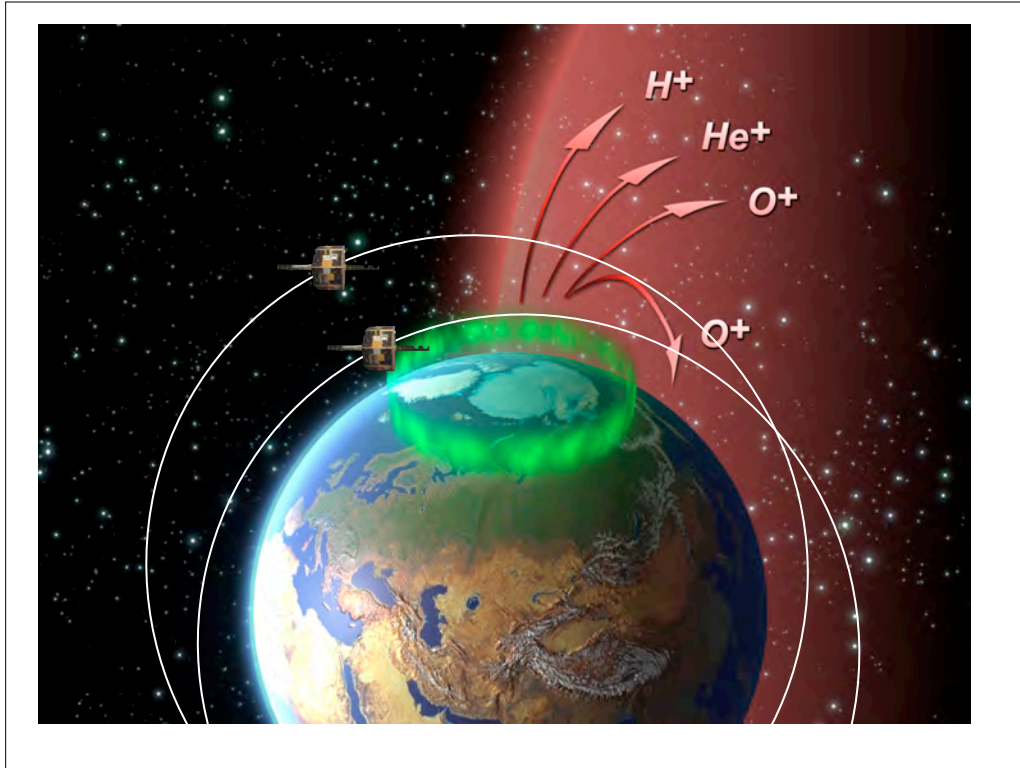


Figure 2. Schematic diagram of an effective MEME orbital strategy that places both spacecraft over auroral forms in approximate conjunction.

New instrument capabilities not seen on DE-1 and -2 are required and available for MEME. These will enhance the spatiotemporal resolution of plasma measurements in support of diagnostics of thin structures (km) sampled at orbital velocity (10 km/s). It is a requirement to observe the lowest energy ionospheric plasmas without assuming that their velocity distributions are in local thermodynamic equilibrium. This requires fast 3D plasma instruments that image multiple dimensions simultaneously, with the sensors maintained at or near the plasma potential to prevent exclusion of the core of ionospheric ions and electrons, which again, may not be "cold". Such instruments require boom deployment to assure that they view normal to the residual spacecraft sheath. Wave measurements must be interferometric to permit wave vector measurement and definitive mode identification. Finally, there must at minimum be background gas composition measurements comparable to laboratory Residual Gas Analyzers (RGA), with gas transport analyzers preferred for completeness.

A trade must be made between onboard or ground based auroral imaging. Onboard imaging would give the tightest integration with the *in situ* observations. But the ground elements of MEME may include a network of all sky cameras and scanning photometers in addition to the required Advanced Incoherent Scatter Radar (AISR) diagnostics. Together these may provide

Mechanisms of Energetic Mass Ejection

adequate diagnostics of auroral ionospheric context in support of interpreting the spacecraft data, providing definitive discrimination between ionospheric upwelling in response to electron heating and upwelling in response to ion heating.

Significant cost elements for the MEME mission include:

1. Collisional and Magnetized Zone Probe (125-500 km)
 - a. Advanced Incoherent Scatter Radar (AISR) plasma diagnostics
 - b. Ground optical observations and auroral imaging (baseline)
2. Exobase-Auroral Transition Zone Probe 1 (500-3000 km)
 - a. Superthermal plasma (0 - 20 eV ions and electrons, resolving 0.1 eV, with aperture bias)
 - b. Precipitating plasma (20 eV - 20k eV ions and electrons)
 - c. E field and wave interferometer (DC - f_{cO2+} - f_{pe})
 - d. Magnetometer (DC - f_{cO+})
 - e. Neutral gas composition analyzer (RGA)
3. Exobase-Auroral Transition Zone probe 2 (500-3000 km)
 - a. Same subelements as for Probe 1
4. Mission & Science Operations Center
 - a. Maintains spacecraft systems and orbits
 - b. Operates instruments
 - c. Merges ground and space data sets into mission database
 - d. Supports community research collaboration

MEME is comparable in scope to the NASA Dynamics Explorer program, which also consisted of two spacecraft, one in circular ~1000 km orbit, the other in a 1000 km x 4.5 R_E (geocentric) orbit. The reported cost of DE1-2, including two unique observatories and launch vehicle, was about \$65M (FY81\$), which inflates to approximately \$160M (FY10\$). However, DE was undertaken before full cost accounting of NASA missions. The value of additional civil service support to the DE mission is roughly estimated to be FY10\$100M (~400 wk-yr). To the above must be added the cost of dedicated supporting ground observations. Thus the full cost of MEME is thought to be in the range \$275-300M (FY10\$) range. This is a generous figure, since the two MEME spacecraft are identical rather than unique. Ideally, the supporting ground observations would be provided by an agency that is responsible for the relevant facilities, and as such, could be contributed from outside the NASA budget, sharing the overall program cost.

4. Relevance

MEME is designed in response to the STP#5 (ONEP) Science Target based on the fundamental process question: "How are planetary parent gases and plasmas heated, accelerated and transported?" It also aims to answer the societal relevance investigation: "How do the Solar Wind and the Magnetosphere-Ionosphere-Thermosphere systems interact with each other?" Understanding planetary plasma production, energization and transport mechanisms is fundamental in determining how planetary atmospheres lose material and contribute to magnetospheric plasma populations and their dynamics. Some of the matter ultimately escapes into the Solar Wind. Loss of ionospheric plasma is effectively atmospheric loss, since the plasma is created through ionization of the neutral atmosphere. Moreover, upwelling gas and plasma often consists of heavier species, such as atomic oxygen, and during intense activity, molecular atoms and ions. Increased mass density changes the atmospheric drag on low orbiting spacecraft, and also the characteristics of the magnetospheric plasma, by loading down flux tubes with enhanced density, thereby slowing wave propagation and rate of reconnection.

Assessment of ionospheric plasma distribution throughout the Magnetosphere requires knowledge of how the ions are energized and ejected from Earth's gravitational potential well. For example, oxygen ions must be heated or accelerated to energies higher than typical

Mechanisms of Energetic Mass Ejection

ionospheric temperatures to escape; however, protons can escape relatively easily at ionospheric temperatures. At the same time, since the heavier ions tend to have lower velocities, they are more likely to be affected by convection. Distinctly non-thermal ion velocity distributions are known to be produced in the magnetized but still collisional F region. Prediction of the ejection of these ions requires detailed knowledge of their energy distribution as they leave the collisional Ionosphere.

The fundamental process of atmospheric mass ejection has especially important implications for life on Earth and other planets, not only in the impact of atmospheric loss but also in the understanding of geomagnetic storms, their magnitude, and duration. We must better understand the nonlinear interactions and feedbacks that enhance or moderate electrical currents in the Magnetosphere to successfully predict storm strength and subsequent societal effects. Ionospheric plasma is a major contributor to the plasma pressure and currents throughout the Magnetosphere during these storms. Because the plasma ejection derives from storm strength but also contributes to storm strength, the exchange of mass and energy between the Ionosphere and Magnetosphere constitutes the largest nonlinear feedback system among the terrestrial planets.

MEME satisfies the relevance criteria cited by the Decadal Survey panels, as follows:

- a. It is a high priority science target of the 2009 NASA Heliophysics Roadmap/Science Plan.
- b. It cuts across all NRC panel themes. Relevance to Solar and Heliospheric Physics lies in the physics of mass loaded plasmas, a common thread applying to the Solar Wind interaction with the solar photospheric gas, dust in the solar system, interstellar gas, planetary atmospheres, and satellite and bolide outgassing.
- c. It answers top scientific questions facing solar and space physics today, as identified in the 2003 Decadal survey and 2009 Heliophysics Roadmap, notably:
 - i. What are the sources of energy that drive ionospheric outflows?
 - ii. How are these energy sources partitioned as a function of altitude and location (e.g., the auroral oval versus the dayside cusp)?
 - iii. Where do the different energy sources deposit their energy?
 - iv. How does the neutral atmosphere respond to the energy deposition?
 - v. Does neutral atmosphere variability affect ionospheric response to energy inputs?
 - vi. To what degree do feedback and saturation processes control the outflow of plasma?
 - vii. What feedback is there on the magnetospheric drivers of the outflows?
- d. It contributes to applications and-or policy making because space weather prediction cannot be accurate without a more realistic model of Atmosphere-Ionosphere-Magnetosphere-Solar Wind interactions. Contemporary models of the storm time ring current cannot produce a realistic ring current without being “spiked” with additional plasma at the boundary between the outer Magnetosphere and the inner Magnetosphere. This additional plasma is believed to originate in the Ionosphere but is not provided for by existing models, owing to the lack of an integral physical model of the outflows.
- e. It complements and indeed is dependent upon ground observational systems including optical spectral imaging and IS radar observations of the auroral Ionosphere.
- f. The cost of MEME to resolve these problems is in the mainstream of typical missions.
- g. It is at a readiness level (technical, resources, staffing) that merits implementation.

Mechanisms of Energetic Mass Ejection

- h. MEME is synergistic with other national and international plans and activities, such as the National Space Weather Program, and other missions such as Cluster, STEREO, THEMIS, SDO, Heliospheric Sentinels, Magnetospheric Multiscale, and RBSP.

5. Discussion

MEME will solve long-standing problems of the origin and loss of terrestrial plasma, from its source in the Atmosphere to the Magnetosphere and Solar Wind. It will use multipoint observations to relate simultaneous conjugate phenomena across key altitude ranges. It will evaluate the importance of electron precipitation, via the creation of superthermal electron distributions or heating of the thermal electrons. And it will determine the source and generation mechanism of the waves that are observed and known to heat ionospheric ions. When these mechanisms are fully understood, it will become possible to derive the heating rates, upward mass fluxes, and composition of mass loss produced by imposed Solar Wind conditions acting upon any upper atmosphere system. This in turn will make possible more accurate and assumption-free predictions of space weather in geospace, as driven by Solar Wind energy inputs and atmospheric ionization, ablation and circulation. MEME is a carefully considered response to the challenge of organizing an interdisciplinary study of the coupled Solar Wind-Magnetosphere-Ionosphere-Atmosphere system. There is an immediate need to advance the predictive capabilities of our increasingly competent models of solar variability and its interactions with planetary bodies, including magneto-plasma coupling leading to terrestrial space weather disturbances. But there is also a long term need to advance our basic understanding of the Heliophysics of planetary system evolution, including the important atmospheric effects of planetary magnetization.

6. Conclusion

Solar variability and planetary magnetization control the flows of matter and energy in our upper atmosphere system, but we cannot begin to predict their nonlinear consequences without a more general understanding of the mechanisms through which stellar winds interact with planetary magnetospheres and atmospheres. Our own home planet offers the best possible opportunity to directly observe magnetized plasma media interacting.

7. Affiliations

1. Heliophysics Sci. Div. Code 670, NASA's Goddard SFC, Greenbelt, MD 20771 USA
2. Laboratory for Atmos. & Space Phys., Univ. Of Colorado, Boulder, CO 80309 US
3. Dept. Of Physics and Astronomy, Vanderbilt Univ., Nashville, TN, 37235 USA
4. Plasma Physics Div., Naval Research Laboratory, Washington DC, 20007 USA
5. Dept. Atmos. Oceanic, Space Sci., Univ. Of Michigan, Ann Arbor, MI 48109 USA
6. University of NH, Space Science Center, Morse Hall, Durham NH 03824 USA
7. Dept. Of Physics and Astronomy, Univ. Of Calgary, Calgary, AB T2N1N4 Canada
8. Space Sciences Laboratory, Univ. of Calif. Berkeley, Berkeley, CA 94720 USA
9. Center for Atmos. & Space Sci., Utah State University, Logan UT 84322 USA
10. Dept. Electrical and Computer Engineering, Boston Univ., Boston MA 02215 USA
11. Inst. Geophys. Planet. Phys., Univ. of Calif. in LA, Los Angeles, CA 90024 USA
12. Aerospace Engineering Dept., Univ. Of Colorado, Boulder, CO 80309 US
13. Geophysics Program, Univ. Of Washington, Seattle, WA 98195 USA

References

- Barakat, A. R., and R. W. Schunk (2006), A three-dimensional model of the generalized polar wind, *J. Geophys. Res.*, 111, A12314, doi:10.1029/2006JA011662.
- Burchill, J. K., D. J. Knudsen, et al. (2010), Thermal ion upflow in the cusp ionosphere and its dependence on soft electron energy flux, *J. Geophys. Res.*, 115, A05206, doi: 10.1029/2009JA015006.
- Heliophysics Roadmap Team Report to the NASA Advisory Council, Heliophysics Subcommittee, (2009), THE SOLAR AND SPACE PHYSICS OF A NEW ERA, Recommended Roadmap for Science and Technology 2009–2030, May 2009, NP-2009-08-76-MSFC, 8-402847.
- Kistler, L. M., et al. (2006), Ion composition and pressure changes in storm time and nonstorm substorms in the vicinity of the near-Earth neutral line, *J. Geophys. Res.*, 111, A11222, doi: 10.1029/2006JA011939.
- Lotko, W., (2007), The magnetosphere–ionosphere system from the perspective of plasma circulation: A tutorial, *J. Atmos. Solar-Terrest. Phys.* 69, p. 191–211.
- Moore, T. E., and J. L. Horwitz, (2007), Stellar ablation of planetary atmospheres, *Rev. Geophys.*, 45, RG3002, doi:10.1029/ 2005RG000194.
- Moore, T. E., and G. V. Khazanov, (2010), Mechanisms of Ionospheric Mass Escape, 2010, *J. Geophys. Res.*, DOI: 10.1029/JA014905.
- Paterson, W. R., and L. A. Frank (1989), Hot ion plasmas from the cloud of neutral gases surrounding the space shuttle, 94(A4), p. 3721.
- Strangeway, R. J., R. E. Ergun, Y.-J. Su, C. W. Carlson, and R. C. Elphic (2005), Factors controlling ionospheric outflows as observed at intermediate altitudes, *J. Geophys. Res.*, 110, A03221, doi:10.1029/2004JA010829.
- Szegő, K. et al. (2000), Physics of mass loaded plasmas, *Space Science Reviews* 94: 429–671, 2000.
- Wilson, G. R. (1994), Kinetic modeling of O⁺ upflows resulting from ExB convection heating in the high-altitude F region ionosphere, *J. Geophys. Res.*, 99, A9, pp. 17453.
- Zeng, W, and J. L. Horwitz (2007), Formula representation of auroral ionospheric O⁺ outflows..., *Geophys. Res. Lett.* 34, L06103, doi:10.1029/2006GL028632.
- Zheng Y., et al. (2005), Polar study of ionospheric ion outflow versus energy input, *J. Geophys. Res.*, 110, A07210, doi:10.129/2004JA010995.

Long Term Observations for Trend Studies

A White Paper for the Decadal Survey of Solar and Space Physics

Susan M. Nossal
Physics Department
University of Wisconsin-Madison
nossal@physics.wisc.edu
608-262-9107

1. The importance of long-term observations

Long-term observations have been essential to understanding the Earth's climate system and changes resulting from human activities. It is necessary to accumulate observations over many years in order to truly understand how a complex system such as the atmosphere behaves and responds to natural and human caused forcings. Perhaps the most famous such data set is the *Keeling* Mauna Loa carbon dioxide record that indicated increasing atmospheric concentrations of this radiatively important molecule [*Intergovernmental Panel on Climate Change*, 2007]. There are multiple other long-term lower atmosphere observations, including atmospheric temperature, precipitation, Arctic sea ice extent, and lake ice records [*Intergovernmental Panel on Climate Change*, 2007]. Taken together and with other long term data sets, they provide knowledge of geographic patterns of change beyond that which would be possible with a single data set.

Remote sensing, satellite drag, and in situ observations have now enabled the building of multi-decade data sets in the middle and upper atmosphere, facilitating the investigation of vertical as well as geographic patterns of variability due to natural and human factors. The middle and upper atmosphere are very sensitive to changes in the concentrations of gases such as carbon dioxide and methane that have increased due to anthropogenic sources. Sensitivity studies indicate that the upper atmospheric response to a doubling of these gases would be *cooling* by tens of degrees Celsius [*Roble and Dickinson*, 1989]. In the lower atmosphere the projected *warming* is on the order of a few degrees Celsius, accompanied by major impacts to weather and ecosystems [*IPCC, Solomon et al.*, 2007]. Understanding natural variability, such as that due to the solar cycle, is important for characterizing the upper atmosphere and for isolating potential signs of anthropogenic change. Continued model development of the whole atmosphere system can assist in better understanding mechanisms for these responses.

Observations of the middle and upper atmosphere can be used as a diagnostic together with those in the lower atmosphere for model validation to test understanding of the physical, chemical, and biological processes affecting the whole climate system. Long-term data sets of middle and upper atmospheric parameters can potentially be used for attribution. For example, changes in the vertical atmospheric temperature and density profiles may assist in distinguishing between changes due to anthropogenic increases in greenhouse gases and

those due to changes in the solar output. While the magnitude of projected temperature trends in the upper atmosphere is greater than that projected in the lower atmosphere, long term observational records in the middle and upper atmosphere have been difficult to acquire due to the faint signals and sensitivity requirements for remote sensing.

Long-term observations of the middle and upper atmosphere are now being used to study natural and anthropogenic factors influencing the Earth's climate. A few examples are as follows. Satellite and balloon observations of stratospheric water vapor have been used to investigate the influence of stratospheric water vapor concentrations on the global surface temperature [Solomon *et al.*, 2010]. This study concludes that stratospheric water vapor is an important driver of decadal global surface climate change [Solomon *et al.*, 2010].

She et al. have obtained a record of mesospheric temperature measurements derived from sodium lidar observations spanning 1990-2007. These measurements indicate a temperature decrease of ~ 1.5 K/decade at 91 km, when they account for the solar cycle and the warming following Pinatubo when fitting the temperature observations [She *et al.*, 2009; Krueger *et al.*, 2009]. The observations are consistent with predictions from both the Spectral Mesosphere/Lower Thermosphere Model and the Hamburg Model of the Neutral and Ionized Atmosphere as shown in Fig. 4(b) of She *et al.*, 2009, and reproduced here.

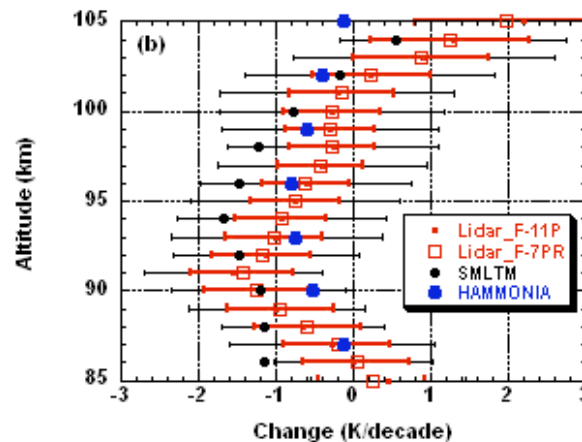


Figure 1. Observed nocturnal temperature trend in degrees K/decade between 85 and 105 km based on 1990-2007 observations from Fort Collins, Colorado. From She *et al.*, 2009.

The thermospheric density is predicted to decrease as the atmosphere contracts in response to carbon dioxide cooling of the upper atmosphere [Roble, 1995]. Emmert *et al.*, [2008] and have compiled a 40-year record of the globally averaged thermospheric mass density from analysis of the drag on the orbits of thousands of objects. The Emmert *et al.* [2008] analysis indicates an overall decline in the thermospheric density at 400 km of $-2.68 \pm 0.49\%$ per decade, with larger decreases during solar minimum conditions. Modeling studies by Qian *et al.* [2006] also predict that the thermospheric density decrease due to CO₂ radiative cooling is larger during solar minimum conditions. The thermospheric density trends derived from satellite drag analysis are in fair quantitative agreement with modeling predictions [Emmert *et al.*, 2008]. Anomalously low densities were observed during the period from 2007-2009 compared with climatologically expected levels [Emmert *et al.*, 2010]

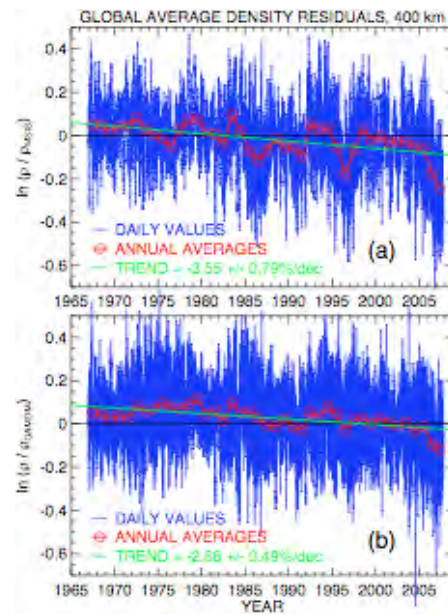


Figure 2a,b. Global average log-density residuals at 400 km plotted as a function of time. In figure 2a the residuals are relative to NRLMSISE-00 and in figure 2b the residuals are relative to the empirical Global Average Mass Density Model. From Emmert *et al.*, 2008.

Atomic hydrogen becomes an increasingly dominant component of the upper boundary of the Earth's atmosphere, the geocorona. Hydrogen in this region is a byproduct of radiatively important water vapor and methane below. The solar cycle is a major source of natural variability in the upper atmosphere and its impact must be understood to characterize this region and to isolate potential signs of anthropogenic change. The University of Wisconsin has compiled a long-term set of ground-based geocoronal hydrogen Balmer α observations spanning three solar minima [Nossal *et al.*, 2008]. The geocoronal Balmer α emission is primarily excited by solar Lyman β ; thus the geocoronal

Balmer α column emission is a function of the hydrogen density profile and the strength of the line center solar Lyman β flux, both of which change over the course of the solar cycle. The Wisconsin observations have indicated a statistically significant solar cycle variation in the geocoronal Balmer α emission with emission intensities higher by a factor of 1.5 ± 0.15 during solar maximum conditions [Nossal *et al.*, 2008]. Work to retrieve hydrogen densities from multi-line hydrogen emission observations is important for separating the portion of trends due to changes in density from that due to changes in the solar flux [Bishop *et al.*, 2004; Mierkiewicz *et al.*, 2006].

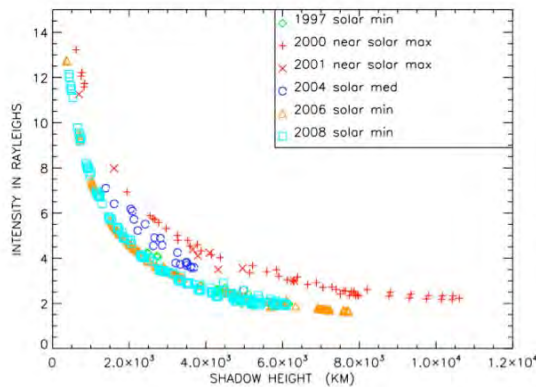


Figure 3a

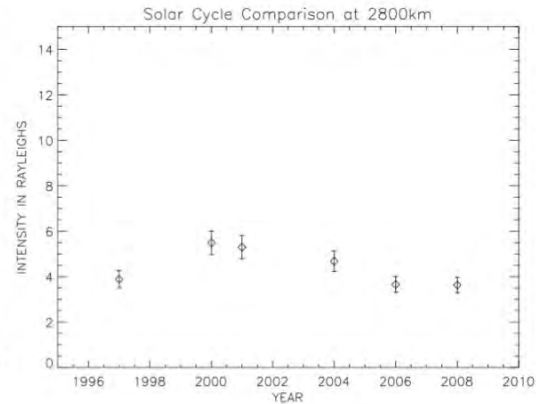


Figure 3b

Figure 3a,b: Solar Cycle 23 observations by the Wisconsin H-alpha Mapper Fabry-Perot (WHAM) of the geocoronal hydrogen Balmer-alpha (656.3 nm) column emission are plotted versus shadow altitude in Figure 3a. The data of figure 3a were fit with a multi-order polynomial and the intensity at a mid-range shadow altitude of 2800 km is plotted versus year in figure 3b. Error bars represent the uncertainty in the *relative* intensity determination. The variation in the Solar Cycle 23 observations is consistent with that from Solar Cycles 21-22 [Nossal *et al.*, 2008].

2. Recommendations for obtaining observations to enable long-term trends studies into the future.

- Achieving measurement continuity is essential for assessing long-term trends in the Earth's atmosphere. Requirements for multiyear comparisons include well-characterized instrumentation, a stable calibration source, cross-calibration between instruments, reproducible observing conditions, and consistent data analysis. Sources of uncertainty must be well characterized to understand the limits of the observations.
- The acquisition of high quality, well-calibrated benchmark observations is required for comparisons with observations taken decades into the future. These observations must be well documented to facilitate such comparisons.
- Overlap between observations taken by older and newer instruments is important for corroboration of measurements and linking of data sets.
- Development of improved relative, absolute, and inter-calibration techniques between different instruments will facilitate connecting observations from different instruments to provide global pictures of atmospheric parameters.
- Continuous measurements of solar fluxes is essential for understanding middle and upper atmospheric processes and natural variability in these regions.
- Joint ground-based and satellite observations provide complementary information about local variability and global scale processes. The observations can also be used to corroborate one another.
- Further collaboration between researchers studying the lower atmosphere and upper atmosphere could facilitate potential new insights into the coupling between atmospheric regions and whole atmosphere processes. Techniques commonly used when studying one region may provide new approaches to probing another region.

References

- Bishop, J., E.J. Mierkiewicz, F.L. Roesler, J.F. Gomez, and C. Morales (2004), Data- model comparison search analysis of coincident PBO Balmer α , EURD Lyman β geocoronal measurements from March 2000, *J. Geophys. Res.*, doi:10.1029/2003JA010165.
- Emmert, J.T., J.M. Picone, and R.R. Meier (2008), Thermospheric global average density trends, 1967-2007, derived from orbits of 5000 near-Earth objects, *Geophys. Res. Lett.*, 35, L05101, doi:10.1029/2007GL032809.
- Emmert, J.T., J.L. Lean, and J.M. Picone (2010), Record-low thermospheric density during the 2008 solar minimum, *Geophys. Res. Lett.*, 37, L12102, doi:10.1029/2010GL043671.
- Krueger, D.A., She, CY (2009), Seasonal variability in mesopause region temperatures over Fort Collins, CO (41 °N, 105 °W) based on lidar observations from 1990 through 2007, *JASTP*, 71, 1565-1570, doi:10.1016/j.jastp.2009.06.003.
- Mierkiewicz, E.J, Roesler, F.L., Nossal, S.M., Bishop, J., Reynolds, R.J., Haffner, L.M. (2006), Geocoronal hydrogen studies using Fabry-Perot Interferometers, Part 1: instrumentation, observations, and analysis, *JASTP*, 68, 1520-1552, doi:10.1016/j.jastp.2005.08.024.
- Nossal, S., R.J. Reynolds, F.L. Roesler, and F. Scherb (1993), Solar cycle variations of geocoronal Balmer α emission, *J. Geophys. Res.*, 98, 3669-3676.
- Nossal, S.M., E.J. Mierkiewicz, F.L. Roesler, R.J. Reynolds, and L.M. Haffner (2006), Geocoronal Hydrogen Studies Using Fabry-Perot Interferometers, Part 2: Long Term Observations, *JASTP*, 68, 1553-1575, doi:10.1016/j.jastp.2005.08.025.
- Nossal, S.M., E.J. Mierkiewicz, F.L. Roesler, L.M. Haffner, R.J. Reynolds, and R.C. Woodward (2008), Geocoronal hydrogen observations spanning three solar minima, *J. Geophys. Res.*, 113, A11307, doi:10.1029/2008JA013380.
- Qian, L., R.G. Roble, S.C. Solomon, and T.J. Kane, Calculated and observed climate change in the thermosphere, and a prediction for solar cycle 24, *Geophys. Res. Lett.*, 33, L23705 doi:10.1029/2006GL027185, 2006.
- Roble, R.G. (1995), Major Greenhouse cooling (yes cooling) – the upper atmosphere response to increased CO₂. *Rev. of Geophys.*, 33, 539-546.
- Roble, R.G., Dickinson, R.E. (1989), How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? *Geophys. Res. Lett.*, 16, 1441.
- She, CY, D.A. Krueger, R. Akmaev, H. Schmidt, E. Talaat, S. Yee (2009), Long-term variability in mesopause region temperatures over Fort Collins, Colorado (41 °N, 105 °W) based on lidar observations from 1990 through 2007, *JASTP*, 71, 1558-1564, doi:10.1016/j.jastp.2009.05.007.
- Solomon, S., K.H. Rosenlof, R.W. Portmann, J.S. Daniel, S.M. Davis, T.J. Sanford, G.K. Plattner (2010), Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, *Science*, 327, no. 5970, pp. 1210-1223, doi:10.1126/science.1182488.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds.) (2007), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Short-term variability of the IT system

J. Oberheide, M. F. Larsen, G. A. Lehmacher, and J. W. Meriwether
Dept. of Physics and Astronomy, Clemson University

Summary

We propose that the A-I-M panel recommends a comprehensive observation program to untangle the complex web of interacting processes and wave coupling that causes day-to-day variability in the ionosphere-thermosphere (IT) system. The program should comprise a minimum of four identical small spacecrafts capable of measuring in-situ plasma density and neutral winds throughout the thermosphere, and a network of ground-based facilities for high cadence wind and temperature observations. Our overarching science objective is: How does the “space weather” of the IT system vary on short time scales (≤ 1 day) and what drives this variability?

Short-term IT variability is poorly understood but critically important for answering a number of fundamental science questions identified in the new CEDAR strategic plan and in the 2009 NASA Heliophysics Roadmap, for instance: “*What governs the coupling of neutral and ionized species?*”, “*How do coupled middle and upper atmospheres respond to external drivers and each other?*”, and “*What is responsible for the dramatic variability in many of the state variables describing the IT region?*”. High-cadence global IT data and understanding the underlying physical processes are also necessary prerequisites for developing data assimilation systems to attain the predictive space weather capabilities that are important for space-related technologies and humans in space. This white paper emphasizes the wave coupling from below aspect of short-term IT variability but the proposed satellite constellation would also provide an unprecedented view of the magnetic and solar driving including geomagnetic storms.

Introduction

Numerous observations and model studies made during the past decade have unequivocally revealed that the “space weather” of the IT system owes a considerable amount of its longitudinal, local time, seasonal-latitudinal, and day-to-day variability to atmospheric waves that originate in the lower part of the atmosphere.

What we now know is that gravity waves generated by hurricanes or convective plumes may propagate into the thermosphere and seed or drive ionospheric instabilities such as plasma bubbles, sporadic-E patches and traveling ionospheric disturbances. Atmospheric tides due to nonlinear interactions and persistent tropical rainstorms produce large longitudinal and local time variations in bulk IT properties, e.g., temperature, wind, composition, airglow and plasma density, to name a few (Figure 1). Oscillations of F-region plasma density and other IT properties at planetary wave periods at least partly reflect the planetary wave activity and variability in the stratosphere. Thermospheric waves generated by high-latitude auroral energy input change the ionospheric and thermospheric environments even at low latitudes. All these waves may further interact with each other and the background atmosphere, i.e., when propagating upward through the stratosphere to the ionosphere, hence producing additional variability and/or secondary waves. The important processes involved include E-region dynamo modulation by nonmigrating tides to produce F-region plasma variability, tidal-ion drag interactions to produce the midnight temperature maximum and midnight pressure bulge in the thermosphere, and the initiation of various ionospheric plasma instabilities by gravity waves.

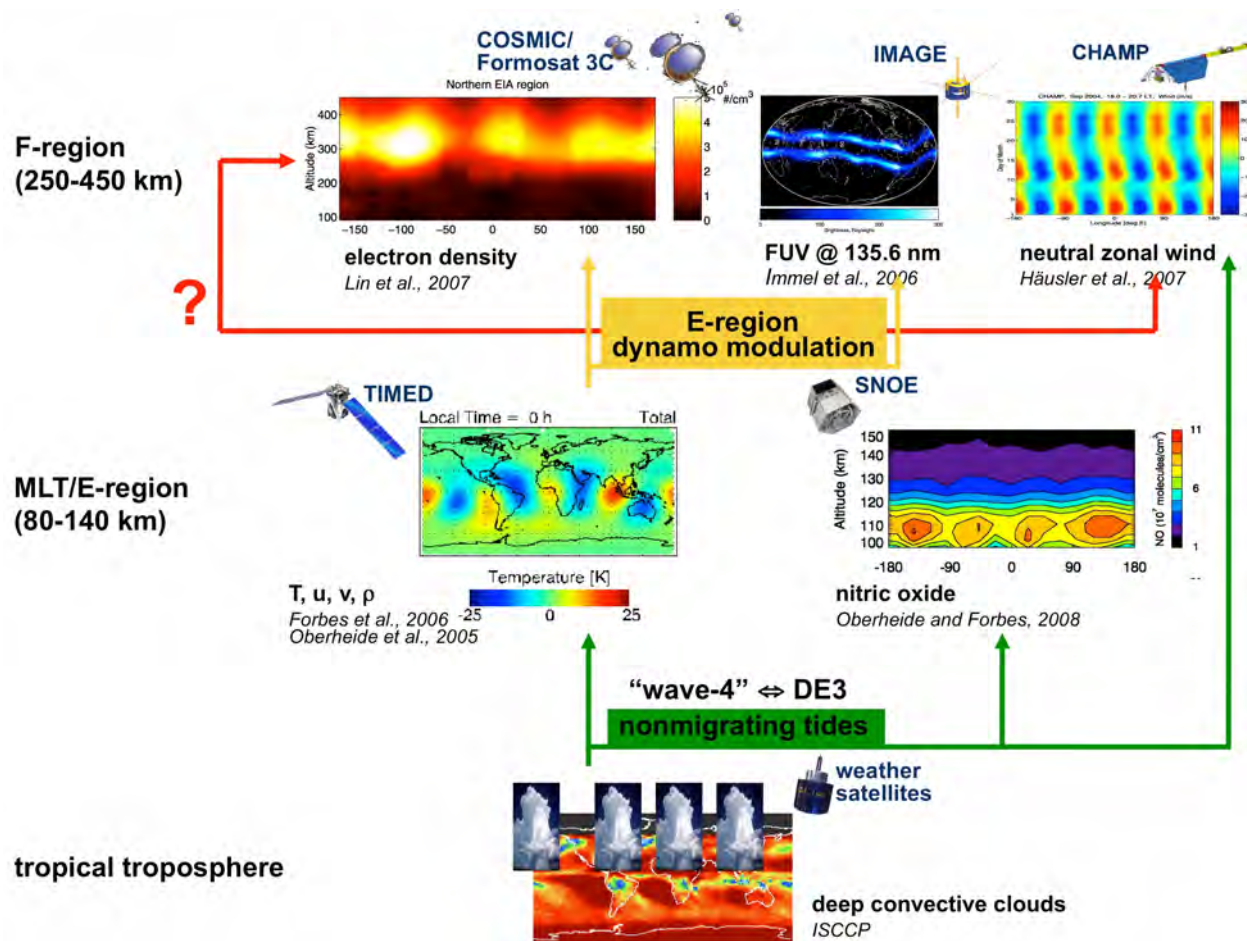


Figure 1. Graphic summarizing some IT impacts of atmospheric tides due to persistent tropical weather systems. Other tides than the depicted DE3 can be equally important. The red arrow probably includes F-region meridional and vertical tidal winds including their effect on ionization, i.e., by $[O]/[N_2]$ advection.

What we do not know is the short-term, that is, the day-to-day variability of the waves in the IT system, their resulting effects and how this compares to the solar and magnetic forcing in the ionosphere. Significant wave variability on all spatial scales must be expected because of source variations (i.e., weather), secondary wave generation through either interactions with the background atmosphere or other waves, or changes in wave propagation and dissipation. It must also be emphasized that we do not know much about the global patterns of tides and gravity waves in the IT at all. Current ground-based observations, primarily by optical instruments or radar systems, do not have the required longitudinal and latitudinal coverage. Current satellite observations are too limited in their local time coverage and/or do not provide the necessary data in the thermosphere at all. Most of the global wave patterns thus come from either empirical or first-principles modeling and not from observations. This is clearly unsatisfactory and severely hampers further progress in understanding the processes that couple the lower with the upper atmosphere, how important they are compared to solar and magnetic driving within the ionosphere, and the development of an IT assimilation system to improve space weather forecast.

Our overarching science question is: How does the “space weather” of the IT system vary on short time scales (≤ 1 day) and what drives this variability?

Tides map polar dynamics (SSW) into low-latitudes, causing MLT and F-region variability on a **day-to-day** basis; **tidal-SPW interaction**

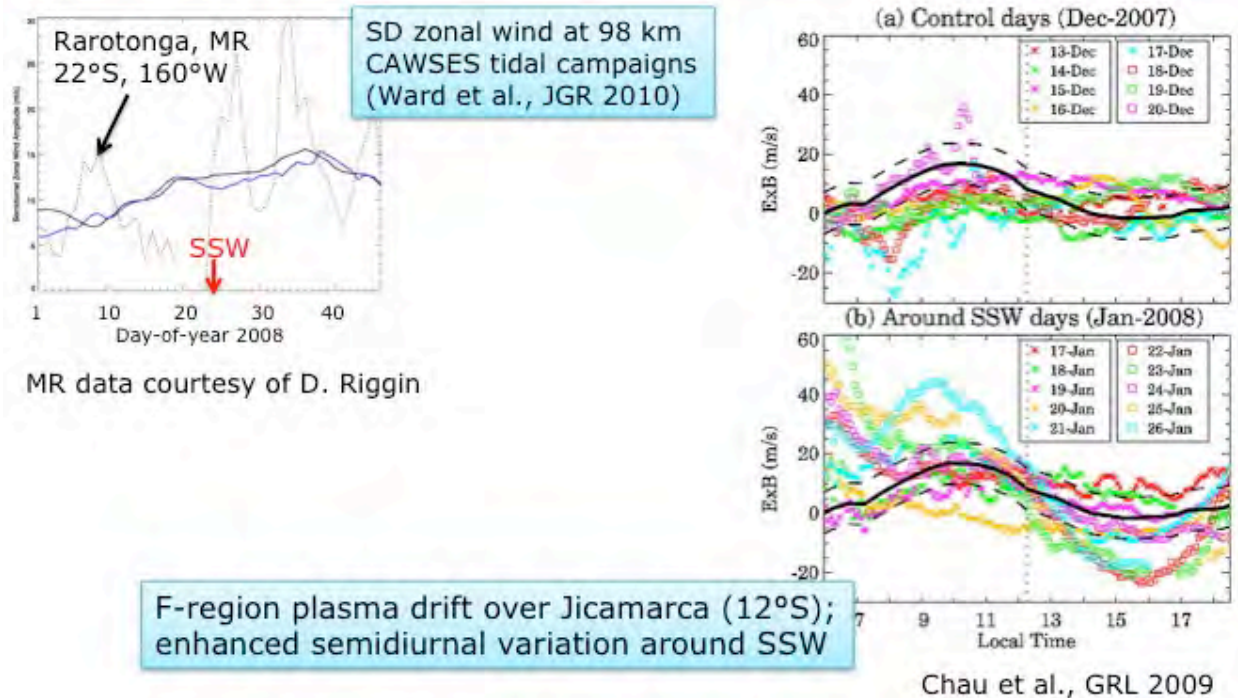


Figure 2. Ionospheric day-to-day variability at the equator due to tidal/planetary wave interactions in the polar stratosphere. The thick black line in the left panel is a 60-day running mean average of the meteor radar tides, thus mimicking the time resolution of TIDI/TIMED tidal diagnostics (blue line).

Science Objectives

Tides

Understanding tidal impacts requires the untangling of the primarily diurnal, semidiurnal and terdiurnal oscillations into zonal wavenumber/frequency pairs since the different wave components have different propagation characteristics and magnitudes that lead to very different impacts on the IT. A tide with a short vertical wavelength, for instance, is unlikely to significantly alter the ionospheric dynamo because of sign cancellation over the E-region. Long wavelength components on the other hand may have vanishing zonal wind amplitudes at the equator, thus having minimal effect on the E-region dynamo, but may be associated with large meridional winds across the equator at F-region heights, thus transmitting the tidal signal into the ionospheric plasma through a different process. Separating the tidal contributions from other sources of variability, as required, relies on global wind observations with full 24-hours local solar time coverage. Local ground-based observations by themselves are not suitable for this task. The only satellite currently in orbit that provides suitable data is TIMED, but because of the 12 min local time precession per day, 60-day averages are required. The time resolution of the derived tides is thus insufficient to delineate tidal variability that occurs on time scales shorter than about one month. A similar argument applies to in-situ tidal observations from satellite-borne accelerometers such as GRACE. The upper boundary of the TIMED tidal diagnostics is 120 km for temperature and 105 km for winds, the latter due to the well-known instrument

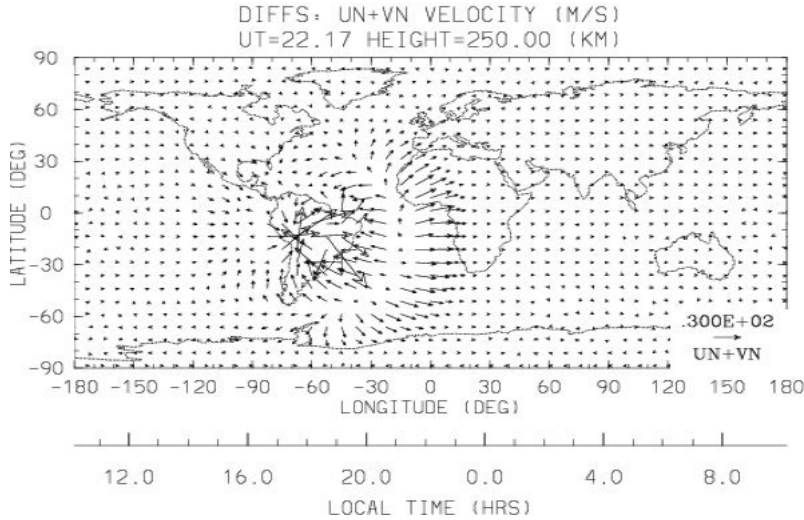


Figure 3. GW @ 250 km due to secondary wave generation in the lower thermosphere. The original source is a convective plume in Brazil. Note the large wave amplitudes over Africa. Model simulation courtesy of Sharon Vadas.

problems of the wind instrument (TIDI). We thus have no tidal observations between the TIMED and the accelerometer altitudes, leaving a thermospheric gap between 120-400 km! Current thermospheric tides come from fitting a tidal model to TIMED tides around 100 km and extrapolating them into the upper atmosphere.

The current database is clearly inadequate to study short-term variability and tidal impacts on the IT system. Tropospheric weather is an important tidal source and varies significantly from day-to-day. We must therefore expect that the IT response (Figure 1) will vary accordingly. Another dramatic IT effect is shown in Figure 2. Polar stratospheric warmings that occur frequently in winter cause low latitude tidal variability, presumably due to tidal/planetary wave interactions. This obviously changes vertical drift velocities at the equator. It is not understood if the tidal variability is due to migrating or nonmigrating tides because TIDI cannot resolve it. The physical processes that couple polar stratospheric warmings with the equatorial electrodynamics are thus unclear and significant amounts of IT variability remain unexplained.

Gravity Waves

Models (Figure 3) predict significant GW activity in the thermosphere. A single convective plume may produce significant thermospheric perturbations over adjacent continents with several thousand kilometer horizontal wavelength and a period in the order of one hour. The cause-and-effect chain is complicated and involves wave breaking and secondary wave generation in the MLT. If the model simulation is accurate, the effect will have severe consequences for the interpretation of ground-based and space-borne observations since a single tropospheric point source can have global effects. The resulting ionospheric effects will be similar since GWs may seed or drive instabilities. Such GWs can only be resolved and tracked to their sources by a combination of multiple space-borne instruments (spatial resolution) and appropriately distributed ground-based instruments (cadence). It is also necessary to remove the tides from the observations because they may mask the horizontal and vertical structure of the GWs.

There are numerous other sources of GWs in the IT system, including auroral forcing, either through Lorentz forcing or Joule heating. Delineating their morphology and resulting ionospheric

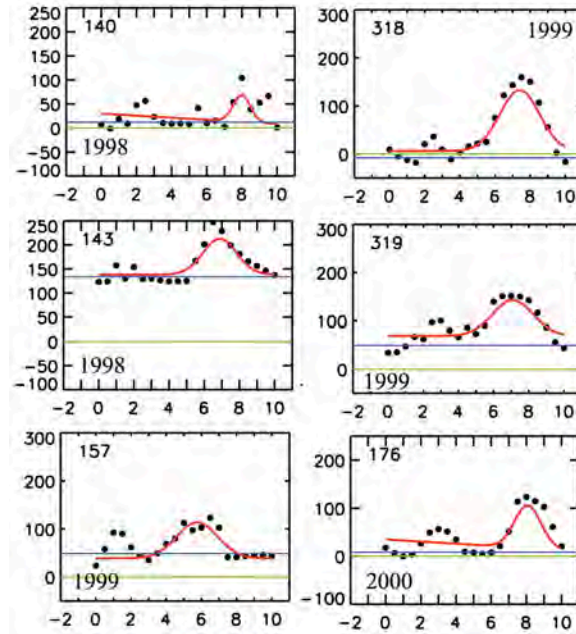


Figure 4. Temperature residuals obtained after the removal of the underlying thermal variation with a MSIS00 model fit. In addition to the primary temperature peak (attributed to the MTM phenomenon) is also seen a secondary peak that appears just after evening twilight.

effects including MSTIDs and plasma bubbles on a global scale is challenging and clearly relies on combined satellite and ground-based observations, such as the proposed satellite constellation. We also expect that it will shed light on the formation of the unusual vertical winds formed in the vicinity of the auroral oval and the role of aurorally-generated waves in the formation of the equatorial ionization anomaly.

Midnight Temperature Maximum – MTM

A prominent example for short-term IT variability due to wave-background interaction is the Midnight Temperature Maximum (MTM). Ground-based Fabry-Perot measurements show that the MTM phenomenon occurs repeatedly throughout the equatorial region on a consistent day-to-day basis but with a MTM peak amplitude that varies over the range of 35 to 200 K (Figure 4). What is not understood is how the progressive stages of the development of the associated midnight pressure bulge throughout the equatorial region relate to the upward propagating tidal wave structure. Also not understood is how the varying mix of higher order harmonics of the equatorial tidal wave structure that reach the thermosphere contributes to the propagation of the MTM from the geographic equator to polar latitudes. These questions can be addressed by the proposed ideal satellite constellation. Also of interest is the secondary temperature peak that is revealed by Fabry-Perot temperature measurements that occur near evening twilight. This feature is currently interpreted to be a result of the heating that occurs when upward tidal harmonics reach the nightglow region near 250 km.

Importance of problem

The 2003 decadal survey already identified the volatile weather in the upper atmosphere as one of the key A-I-M issues and recognized an urgent need for comprehensive observation programs

to untangle the complex web of interacting processes and global wave coupling, including the upward transport of energy and momentum from the lower atmosphere. Discoveries made in the past five years, namely the abovementioned tidal, planetary and gravity wave effects on the ionospheric plasma and upper thermosphere, lend further impetus to this requirement. A fundamental science question that remains is the geographical and seasonal distribution of short-term (1 day or less) variability in the IT system and its driving mechanisms. Ground-based observations show broad evidence of dramatic variations in neutral and plasma parameters from day to day, or even within one day, but neither the global morphology nor its causes are known. Tides due to tropospheric weather and/or nonlinear interaction processes may inherently vary from day to day and the range of wavenumber/frequency components may react quite differently to forcing variability. Gravity waves due to convection may have the most profound effects in the thermosphere several thousand kilometers away from their source region. The extent to which the lower atmospheric variability is a driver for the short-term variability of the IT system needs to be determined, as well as the relative importance of the solar and magnetic driving within the ionosphere. This is also important to space weather applications as it will likely have consequences for telecommunications, re-entry of space vehicles, and satellite operations. The implications for ionosphere-magnetosphere coupling still need to be determined.

Short-term IT variability was already identified as a high priority in the 2003 decadal survey. Referring to the upward coupling of energy due to tides, it reads: *“Any satellite provides a global picture of a tide at only two local times on a given day. Continuous ground-based observations are needed, but the existing facilities are too few to provide a complete geographic perspective on these inherently global-scale waves. A complete characterization of atmospheric tides and their important effects in the upper atmosphere can only be accomplished by correlative analysis and assimilation of both ground-based and satellite-borne measurements. Further, tidal diagnostics from a satellite constellation are needed to fully resolve the spatial and temporal variability of atmospheric tides.”*

We note that this task has not been accomplished since the earlier survey was written! The status of the Geospace Electrodynamics Connections (GEC) mission is uncertain and the recommended *“aggressive program to field hundreds of small automated instrument clusters”* did not take place. Such space-borne and ground-based data, however, are not only needed to resolve the short-term IT variability from the data side alone but also to develop working assimilation systems. The latter concept was extremely successful in meteorological weather forecasting and similar progress can be expected if implemented for the upper atmosphere.

The goals outlined here are consistent with the new CEDAR strategic plan and the priority investigations outlined in the 2009 NASA Heliophysics Roadmap, specifically: *“What governs the coupling of neutral and ionized species?”*, *“How do coupled middle and upper atmospheres respond to external drivers and each other?”*, and *“What is responsible for the dramatic variability in many of the state variables describing the IT region?”* Delineating and understanding short-term IT variability on a global scale is crucial for achieving closure. Data from the proposed multiple spacecraft constellation, combined with ground-based instruments, would also provide additional knowledge concerning the properties of upper atmosphere winds and would enhance the design and implementation of NASA’s STP#7 mission (Ion-Neutral Coupling in the Atmosphere, INCA) that is currently scheduled as a Medium size mission, but not until 2025.

Proposed implementation

A minimum requirement is to have global observations of thermospheric winds every 3 hours local time. This is needed to resolve and separate migrating and nonmigrating diurnal, semidiurnal and terdiurnal tidal components. A three-hour local time coverage can only be achieved by using four or more spacecrafts in appropriate orbits. Every spacecraft should have an in-situ plasma instrument on board, to provide ionospheric data with high cadence. The global data must be complemented by longitudinal networks of ground-based stations to provide a higher time resolution where needed, i.e., at the equator, at mid-latitudes and in the polar regions, to resolve higher frequency waves, such as thermospheric GW waves, with periods that are on the order of 1 hour. Ground-based Fabry-Perot interferometers that provide both temperature and wind data would be as useful complement to the spacecraft measurements.

An ideal mission would comprise 12 satellites that all measure thermospheric winds and temperatures. The latter is required due to the wave effects on reaction and collision rates that govern the neutral and ion chemistry, and for determining the energy budget. Temperature measurements would almost certainly come from nitric oxide observations and would thus provide a detailed view on Earth's natural thermostat. The 12 satellite constellation must be in four different orbital planes to achieve the 3-hour local time coverage. Having three satellites per orbit would then provide 30-min time resolution along-track, which is sufficient to resolve the GWs from Figure 3 from space-based platforms. Again, this needs to be complemented by a network of ground-based instruments to facilitate observations with higher time resolution.

Feasibility of implementation

Both the required space-borne and ground-based measurements are feasible. In-situ electron density observations are a standard technique. The space-borne wind observations will likely come from green and red line Fabry-Perot observations, as a heritage of HRDI and TIDI type instruments. Temperature data (for the ideal mission) would come from the same instruments by measuring rotational temperatures in the upper thermosphere and/or from 100-170 km nitric oxide observations. The latter technique is essentially a line-width absorption measurement at 5.3 μm using a pressure modulation cell. It is independent of knowing the atomic oxygen concentrations and would also provide infrared cooling rate observations with an unprecedented cadence. The ground-based optical Fabry-Perot instruments are also well developed and of proven quality.

It is not our purpose here to provide a specific mission budget but the cost for a multi-spacecraft mission is obviously significant. Some economy of scale can undoubtedly be achieved by using identical instrument packages for each spacecraft, but it seems clear that a different approach to spacecraft component hardening and reliability will be required if large constellations with tens of vehicles are to be feasible. Recent experience with some of the smaller spacecraft form factors have shown that significant cost reductions can be achieved by establishing acceptable failure rates and using spacecraft arrays that meet the minimum science requirements statistically rather than individually. The development of innovative quality control management approaches will be an important challenge and may be more demanding than the instrument techniques themselves.

**Understanding Geospace on a Grand Scale:
The Global Ionosphere/Thermosphere Constellation**

A White Paper for Consideration by the
Heliophysics Decadal Study Committee
National Research Council
Washington, D.C.

By

Robert Pfaff, Dean Pesnell, Doug Rowland
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

12 November 2010

I. Introduction and Scientific Motivation

The initial exploration of the near-Earth regions of Geospace -- essentially, the ionosphere and thermosphere -- has revealed a space environment exceedingly rich in physical processes, dynamics, and chemistry that has far transcended even the most studied contemplations of scientists who sought to explain how planetary upper atmospheres might actually work. Among the transformative revelations of modern experimental space physics is the recognition that the behavior of the Earth's ionized and neutral gases depends both on non-linear coupling between the different species within the gaseous envelope, as well as on complex, highly varying external sources. These external sources include the fluctuating, unpredictable EUV radiation from the sun as well as intricate energy and momentum exchanges with the troposphere/stratosphere/mesosphere below and the magnetosphere and solar wind above. The dependency of the Earth's upper atmosphere on both internal and external drivers is so fundamental that any effort to understand the ionosphere/thermosphere region without incorporating such a "systems view" would be futile.

Seventy-five years of pioneering observational research -- initially with ground-based instruments and later with space based *in situ* probes and imagers -- in concert with advances in theory and modeling, has provided the fundamental underpinnings upon which we base our knowledge and predictions of how the ionosphere/thermosphere system might be characterized. Although much progress has been made from the standpoint of the limited surveys to date and the physical explanations currently on hand, our understanding of the dynamics and fundamental non-linear physics of this region is still rudimentary at best. In fact, it is precisely these initial observations, intriguing though they may be, that have revealed the overwhelming gaps in our knowledge of the Earth's ionosphere/thermosphere, primarily resulting from the lack of basic, systematic, global observations. As will be demonstrated here, the next logical step to expand our knowledge -- one that would enable scores of immediate, fundamental advances in our understanding -- is to simultaneously observe the ionosphere/thermosphere system both globally -- at all local times and latitudes -- and at the appropriate altitudes where the

Although the knowledge to be acquired here addresses fundamental physics inherent to all planetary upper atmospheres and their evolution and hence is important in its own right, there is another motivation for the proposed initiative. The Geospace environment increasingly affects society and technology in numerous areas, many of which were not appreciated even in the recent past. These areas include communication and navigation systems, over-the-horizon radars, satellite drag and debris management, all rocket/missile launches including those used by the emerging industry of space tourism, and the vitality and safety of the world's power grid systems. Thus, from an applications standpoint, the data to be gathered as part of the grand scheme described herein will also be used to significantly advance space weather model characterization and forecast accuracy.

The Earth's upper atmosphere is both intrinsically inhomogeneous and highly dynamic, varying markedly with latitude, longitude, local time, and altitude. A representation of some of the diverse processes that help characterize this region of Geospace is shown in Figure 1. From a general standpoint, this depiction illustrates that many different physical processes are at work in the near-earth region of Geospace, and that these processes are considerably different between the high latitude/polar caps and the low latitude regions and between night and day. Furthermore, the fact that Geospace is threaded by a dipole-like magnetic field enables efficient coupling of electrical energy between the different altitudes within Geospace, between the northern and southern hemispheres, and between the near-earth region of Geospace and that of the more distant magnetosphere and solar wind.

1. High latitude electrodynamics and atmosphere forcing.

The diagram illustrates the complex interactions between the solar wind, the Earth's magnetic field, and the ionosphere. Key components and processes shown include:

- Solar Wind:** Incoming particles and magnetic field lines from the Sun.
- Magnetopause and Bow Shock:** The boundary where the solar wind is deflected by the Earth's magnetic field.
- Magnetosphere:** The region dominated by the Earth's magnetic field, containing the ring current and auroral oval.
- Ionosphere:** The layer of the atmosphere containing free ions and electrons, with distinct E and F layers.
- Particle Populations:** Energetic particles, O₂⁺ ions, and neutral atoms are shown moving between the magnetosphere and ionosphere.
- Neutral Wind:** Neutral atoms flowing from the ionosphere into the magnetosphere.
- Plasma Wind:** Plasma flowing from the ionosphere into the magnetosphere.
- Field Lines:** Magnetic field lines connecting the magnetosphere to the ionosphere, showing the flow of particles and energy.

Arrows indicate the direction of flow and interaction between these various components.

Figure 1

to the magnetosphere are not characterized (or understood) at all, even in a static case. In order to understand the dynamic plasma and atmospheric high latitude environment and its response to magnetospheric forcing, particularly on time scales commensurate with storm-driven events, a variety of measurements including plasma, neutral gas, electric and magnetic field, and energetic particle data must be gathered simultaneously at all local times and at low altitudes (< 350 km).

Furthermore, high latitude imagers will provide the large scale context of the auroral measurements in both the north and the south. (See Figure 2b).

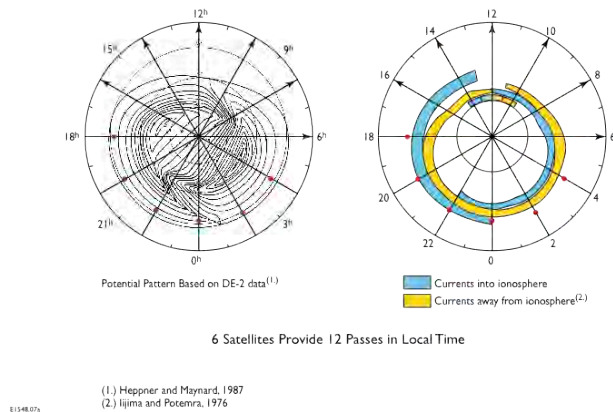


Figure 2a

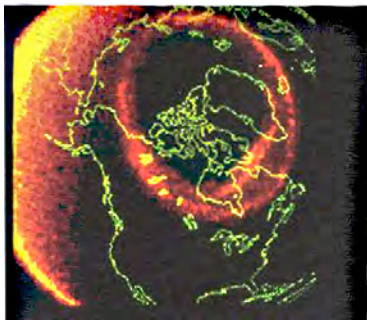


Figure 2b

Understanding the high latitude electrodynamics and atmospheric input is critical as it is this region of the upper atmosphere that is the main depository for magnetospheric energy via Joule heating and other processes. Furthermore, the atmosphere is believed to play a major role in regulating the aurora and magnetospheric energy deposition, via braking of energetic particles, creation of thermal plasma that subsequently shorts out electric fields, and sustaining Hall and Pedersen currents via the enhanced conductivities at lower altitudes. Simultaneous measurements of electric fields, currents, plasma,

neutral winds and density, and energetic particles, at all local times in both the northern and southern hemispheres will enable our knowledge and understanding of these critical processes to advance by a marked degree.

2. Response of the Upper Atmosphere to Magnetic Storms

A long standing goal of the space physics community is to determine how the ionosphere/thermosphere responds globally to magnetic storms, including electric fields, plasma density, neutral winds, and composition. From “spot” measurements and satellite passes at given local times, we know that the ionosphere and thermosphere are radically changed, and in some sense, the mid and low latitude ionosphere is obliterated during magnetic storms. However, we do not understand how this happens. We have very little knowledge of how the global system responds to such changes in magnetic activity, including small magnetic storms that may occur frequently when solar activity is moderate.

Understanding the response function of the upper atmosphere and how energy and matter are redistributed during storms is of paramount importance to knowing how the upper atmosphere works. Besides uncovering the basic processes responsible for the deformation of the upper atmosphere during storms, we need to understand the time constants over which the ionosphere and thermosphere “remember” the effects of the magnetic storm. Over what period (hours or days?) does the upper atmosphere need to recover? What if the upper atmosphere is battered by a second magnetic storm while “recovering” from an earlier one, as so often happens with large scale magnetic activity?

One of the main requirements to understand the effect of magnetic storms on the I/T system is the need to look at all longitudes and local times simultaneously. For example, it is essential to identify both the longitude/local time that is at the epicenter of the storm effects as well as to ascertain which local times/longitudes are not affected. (see Figures 3, 4a, 4b, 4c). Are there longitudes on the Earth for which, when at a prescribed local time during a large magnetic storm, the effects will be most severe? These questions are important not only for understanding the physical processes that affect the global upper atmosphere during magnetic storms but also for predicting the space weather effects at the time when the initial solar disturbance is identified.

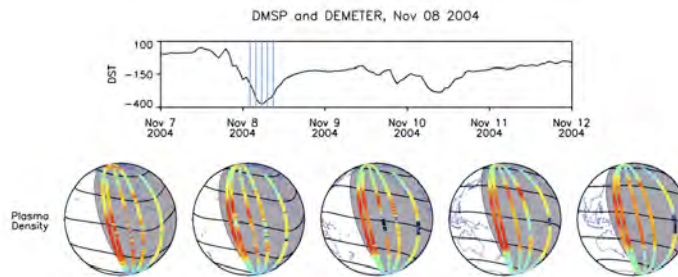
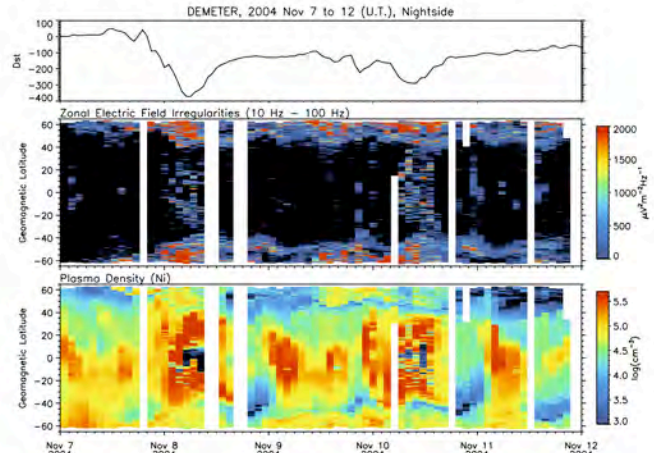


Figure 4a

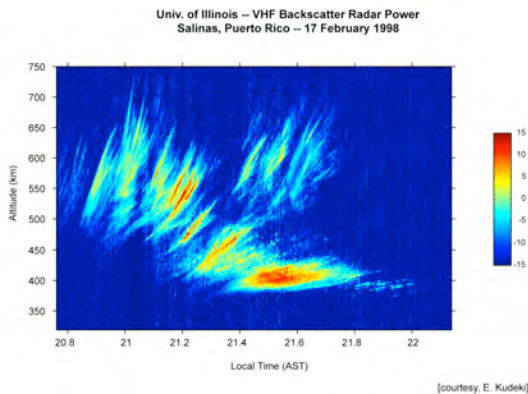


Figure 4c

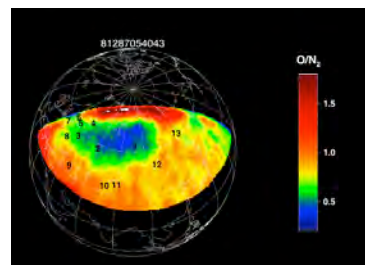


Figure 4b

3. Global neutral wind circulation patterns, neutral density structure, and planetary and gravity waves

The Earth's global neutral wind circulation patterns are very poorly known. Sounding rocket data has shown that the E-region winds are much larger than those predicted by tides and this underscores not only our uncertainty regarding the wind profiles, but also the drivers responsible for the winds. In concert with this uncertainty, the structuring of the neutral density is not known. Recent CHAMP (Figure 6) and STREAK (Figure 6) measurements have shown large scale neutral structuring that does not agree at all with the MSIS models, but rather, are organized by the magnetic field and other parameters that normally control the ionized gas population only. When gravity waves and planetary scale waves are included, it becomes obvious that a large scale sampling of the upper atmosphere simultaneously at all latitudes and local times is needed to uncover the dynamics of the neutral atmosphere in addition to the main physical processes responsible for this behavior.

The neutral atmosphere specification is vital for our understanding with respect to the evolution of the atmosphere (i.e., understanding loss and escape of the neutral gas) as well as its coupling to the plasma environment. Evidence of structured electric fields appear ubiquitous at low latitudes at night without a large counterpart in the plasma density and this may be due to coupling with the neutral atmosphere. In this manner, a global network of measurement platforms will enable the upper atmosphere's response to hurricanes and other tropospheric forcing functions to be known and understood for the first time.

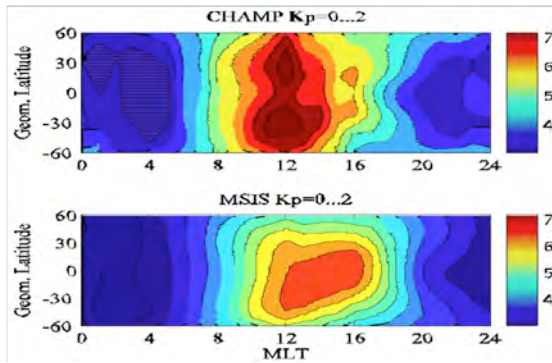


Figure 5

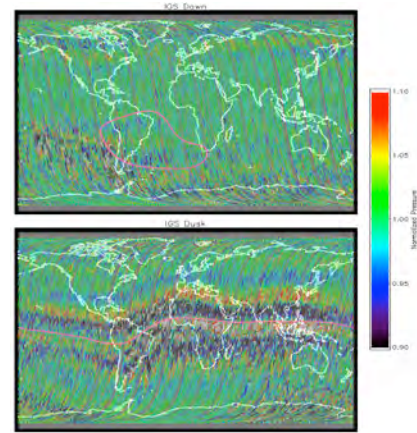


Figure 6

III. The Ionosphere/Thermosphere Constellation Concept

Observing the upper atmosphere system shown in Figure 1 presents an immediate challenge: after determining what are the key data parameters to observe, deciding how best to acquire these data and their temporal and spatial variations is not straightforward. Without reviewing all of the techniques here, it is sufficient to say that a robust, systematic observation approach would, in our view, require simultaneous measurements using both *in situ* probes and imagers in space, as well as arrays of ground-based instruments located around the globe. First, different techniques measure different quantities with different accuracies and integration times. Further, space-based observations in near-earth orbit acquire measurements about prescribed trajectories

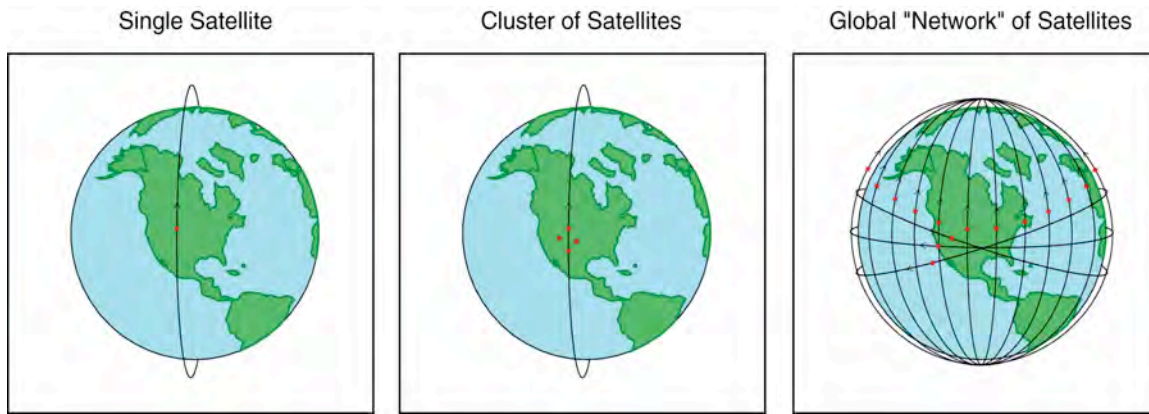


Figure 7

defined by orbital mechanics whereas ground-based measurements gather high time resolution measurements frequently at all altitudes but essentially at fixed locations. We focus here on new strategies to deploy multiple platforms to acquire the *in situ* data, although space-based imagers and ground-based data are important parts of this measurement strategy.

The concept of a global network of measurements is shown in Figure 7. The panel on the left shows a single satellite in a polar orbit. Measurements gathered on such a platform would be excellent for event studies and initial exploration. Such data would be quite useful for providing average global conditions, as shown by the Dynamics Explorer-2 satellite. Next, consider a cluster of satellites traveling in tandem or perhaps strung out along an orbit, such as in a string of pearls configuration, as shown in the middle panel. In this fashion, the data from the joint measurement platforms would be exceedingly useful for studies of events in one location that need to be resolved in space and time. The data could also be used to provide refined average global conditions. An example is NASA's 3-4 satellite suite that forms the Geospace Electrodynamics Connections mission. Finally, in the panel on the right, we depict a global network of satellites. These platforms would provide global, simultaneous observations covering all latitudes and local times, and promise to reveal global-scale processes, coupling to other regions, and large-scale event studies particularly within the "big picture" of Geospace. An example would be the Ionospheric-Thermospheric Constellation mission described herein.

As stated above, the problem is that global dynamics are not captured by single satellite. Average conditions characterized from data from a single satellite are static, masking large-scale spatial and temporal changes. Ionospheric and thermospheric models need to be constrained at more than one point (or one trajectory) and furthermore, the dynamic variations must be uncovered to make progress with our understanding.

How many satellites are needed? Modeling will provide the basis of the response to this question. Here, we consider 24, all with station keeping so that they will fly in formation. One can well imagine 12 satellites in the northern hemisphere and 12 in the southern hemisphere, trading places every 45 minutes (since the orbital period is essentially 90 minutes.)

The basic approach is thus very straightforward: The constellation will gather simultaneous, global data of key ionospheric parameters using

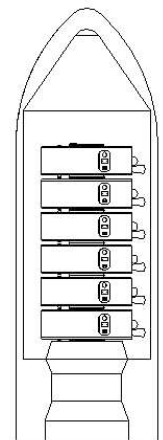


Figure 8

identical instruments on multiple platforms distributed in local time. The data will be used, among other things, to create space-time maps of all of the basic electrodynamic, gas property, and particle measurements in the ionosphere/thermosphere system. The satellites will use propulsion to provide continuous measurements at (say) 300km, 400km, 500 km, as well as station keeping to enable formation flying. The data will be used to develop an unparalleled physical description of the I/T system and the sources of energy and momentum that drive this system. Coupled with data from simultaneous imagers (one above each pole and one at geosynchronous altitudes monitoring the day and night sides of the earth), our understanding of Geospace will be dramatically advanced to levels never before imaginable. Finally, coupled with the measurements will be an augmented theory and modeling program that will dramatically improve both empirical and predictive ionospheric models.

IV. An Initial Hardware Approach Consisting of *In situ* and Imaging Satellites

We consider satellites with identical suites of “notional” instruments divided among these categories:

- Gas properties package
(Neutral and plasma densities, drifts, temperatures, composition)
- Fields properties package
(Electric and Magnetic fields, plasma wave interferometers, Langmuir probe)
- Energetic Particles package
(High and low energy electrons, ions)

Other candidate instruments for each platform include: an ionospheric sounder that will provide, along the satellite track, density profiles from spacecraft to hmF2 altitude and irregularities regions, regardless of whether the satellite is above or below the F-peak; GPS receivers for TEC and scintillations, tomography and beacon experiments, and a Fabry Perot neutral wind instrument, among others.

Included in this array of satellites would be higher altitude satellites with FUV imagers with filters, stationed over the poles to gather the auroral dynamics picture and at geosynchronous altitudes to view the upper atmosphere variations at noon and midnight.

The concept mission design would notionally include six polar orbiting satellites to be launched on a single vehicle, evenly spaced in local time after 9 months, in a manner similar to how the COSMIC satellites were launched and separated. (See Figures 8, 9).

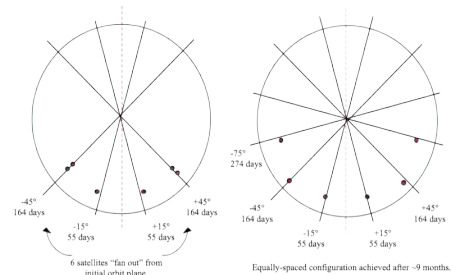


Figure 9

VI. Summary

The IT-Constellation proposed here would provide a major leap forward, addressing fundamental, global physical processes and providing fundamental, new knowledge of Geospace. In particular, it addresses “system science”, revealing how the ionosphere-thermosphere connects, globally, to the magnetosphere above and the troposphere below.

As a NASA “Flagship” mission, the Global IT-Constellation is the next logical step in NASA’s observing program to understand the Earth’s space environment.

Next Gen IT Modeling Infrastructure for Space Weather Forecasting

John Retterer
Air Force Research Laboratory

The phenomena and processes that help shape the structure of the ionosphere and thermosphere (IT) are understood in a general way, but the overall agreement of our models of the IT system with observations of thermospheric and ionospheric parameters is still unsatisfyingly poor. In large part this is due to lack of knowledge of the instantaneous values of parameters which vary from day to day or faster, the 'weather' of this region of the atmosphere. This has consequences not only for the forecasting of potential impacts of phenomena of the IT system on human activities, but is also limiting our progress in further understanding the system, because this variability produces 'noise' which obscures the additional 'signals' in our data due to other processes which have not yet emerged in our understanding of the system and prevents their discrimination.

Three general thrusts will be important for making more than incremental improvements in our modeling of the IT system: 1) greater use of data assimilation, 2) stronger coupling of regions of the atmosphere which are often now treated separately, and 3) higher spatial and temporal resolution in our modeling of structures in the atmosphere. The best use of all of these ideas will require planning for infrastructure development that is appropriate to consider in this decadal survey.

Demonstrations of the effectiveness of data assimilation for improving the accuracy of IT modeling have been made, most notably with the GAIM model of the ionosphere implemented in the Air Force Weather Agency. This model uses measurements of total electron content (TEC) of the ionosphere and other measurements to improve the specification of the electron density in the ionosphere; a newer version promises to use other measurements to help specify the 'drivers' of the ionosphere, *e.g.*, electric fields and winds, which will allow other quantities besides the electron density to be specified, *e.g.*, conditions for plasma instability within the ionosphere, as well as potentially to improve the forecasts, the predictions of structure in the future, of the ionosphere. Now we need to develop the infrastructure of data collectors to provide the optimum information for the data assimilation process. One of the key drivers of the plasma density in the ionosphere is photoproduction of plasma by solar EUV and X-ray emissions. We need to guarantee that the solar spectrum is monitored continuously by satellites outside the Earth's atmosphere. Another important source of information concerns the conditions in the solar wind, which is presently monitored by the ACE satellite; we need to guarantee that our flow of data from ACE and its follow-on do not falter. In-situ measurements of plasma density, drift, and temperature, as well as measurements of neutral winds and densities, are probably the most useful data for assimilation into an IT model, but the current satellites which can make these measurements are woefully sparse in their coverage of the globe. We need to launch a constellation of satellites which will monitor these parameters around the globe and relay that information to the ground in near real-time.

To take advantage of this information requires the development of models which can use it, and a computing infrastructure which can run the models quickly enough so that forecasts of IT events can be made before they happen. A popular mechanism for implementing data assimilation is the ensemble Kalman filter. This requires multiple simultaneous model runs to cover a multi-dimensional parameter space so that the set of parameters that yield the best fit to the data can be chosen. More accurate fitting requires higher dimension parameter spaces, which require more ensemble members to sample accurately. To run the model for all the necessary ensemble members quickly enough to produce a forecast of phenomena requires an investment in computation facilities.

Another thrust that is required to make our IT models more realistic and accurate, is the coupling to models of other regions of the atmosphere. Tropospheric and stratospheric phenomena affect the thermosphere and ionosphere through winds, waves and tides, and energy deposition. Solar-wind, magnetosphere, polar cap, and inner magnetosphere phenomena have great effect, too, but these couplings have already been recognized, *e.g.*, in the development of CISM, the NSF effort to couple solar-wind, magnetosphere, and IT models. It has come to be recognized since then that lower portions of the atmosphere are significant, too, and the most satisfying IT model will be a whole-atmosphere model which includes phenomena from the ground up. The data collection for the assimilation process for tropospheric weather forecasting performed by weather agencies is so extensive and expensive that it would be foolish to attempt to duplicate the effort; instead use of the output of models like those at NCEP, National Center for Environmental Prediction, should be used to provide our lower boundary condition. One step toward maintaining our coupled-model capability is to find a home for the CISM suite of models when the NSF program expires.

The third thrust identified here is a focus on increasing the resolution of models so that coupling between different spatial-scale regimes can be made more accurate. A case in point is the development of radio scintillation through plasma turbulence in the ionosphere. Observations point to atmospheric gravity waves as one of the likely phenomena that produce the initial irregularities in density that develop into the turbulence when the plasma is unstable. Yet no global thermospheric model has the spatial resolution to resolve gravity waves and describe their occurrence let alone their spectrum. At the other end of the chain of processes that lead to scintillation, the wavelengths of the plasma density irregularities that couple with radio waves to perturb their propagation are in the sub-kilometer range. Not even the models that describe the development of the turbulent plumes can describe the production of wavelengths this short without some extrapolation of the spectrum. Extension of the dynamical range of a simulation model is often a matter of greater computational resources, which can be provided with infrastructure investment. Note that the computational requirements often grow more quickly than linearly in the resolution, because the execution steps of the algorithms often scale as a higher power of the number of elements.

To summarize, with what we know now, to make significant progress in modeling the ionosphere/thermosphere system for both discovery of new phenomena and processes as well as for space weather forecasting to minimize the impact of IT phenomena on human activities, we need to invest in the following IT modeling infrastructure:

1. Guaranteed solar-spectra data in near real-time

2. Guaranteed continuation of solar-wind data in near real-time
3. A constellation of low-altitude earth orbit satellites in both polar and equatorial orbits, with in-situ sensors for neutral densities, composition, and winds, plasma (electron and ion) composition, density, drift, and temperature, delivered in near real-time
4. Continuation of the CISM coupled-model system, with expansion to include coupling with lower-atmosphere models (mesosphere, stratosphere, and troposphere), with data assimilation
5. Investment in a high-performance computing facility where new algorithms for breakthrough increases in resolution and dynamic range, along with data assimilation techniques, could be pursued.

A Constellation Mission to Understand the Thermospheric Reaction to Energy Input Across Scales

Aaron Ridley (University of Michigan)

Andrew Nicholas, Douglas Drob (Naval Research Laboratory)

Jeffrey Thayer, Jeffrey Forbes (University of Colorado)

Timothy Fuller-Rowell, Tomoko Matsuo (NOAA)

Mark Conde, Bill Bristow (University of Alaska, Fairbanks)

Larry Paxton (JHU/APL)

Fredrico Herrero (NASA, GSFC)

1 Abstract

The thermosphere and ionosphere are tightly coupled, overlapping regions of the atmosphere that are forced from above and below. The system has no horizontal boundaries and therefore waves and winds can advect mass, momentum and energy easily from one place to another. Therefore the system is extremely dynamic. Utilizing single satellites to examine the dynamics of such a complex environment is extremely difficult because of the space/time ambiguities that exist with all satellites. Therefore, it is suggested that the ionosphere/thermosphere system instead be studied with a constellation of low mass Nano-Satellites that are complimented with miniaturized instruments that are capable of measuring the in situ state of the system (i.e., density, temperature and winds of the neutrals and ions). It is suggested that having four planes that contain six spacecraft is roughly the minimum that will allow the dynamics of the system to be adequately studied.

2 Description

The thermosphere is a crucial boundary between the Earth's atmosphere and space. Orbiting satellites and debris are strongly affected by the lifting of the atmosphere that results from thermospheric weather [Bruinsma *et al.*, 2006; Fuller-Rowell *et al.*, 1994; Prölss, 2004]. On March 12, 2009, International Space Station astronauts had to take shelter because a piece of debris passed too close to the ISS. On February 10, 2009, an Iridium cell phone satellite was destroyed in a collision with a defunct Russian satellite. These types of events can be mitigated with better understanding of the upper atmospheric conditions, which leads to more accurate orbit determination [e.g., Bowman *et al.*, 2008]. While current empirical models of the thermosphere can describe the general trends, they can not capture the true dynamics of the system [e.g., Roble *et al.*, 1988; Fuller-Rowell and Rees, 1980]. For an accurate specification of the density, an understanding of the thermospheric response to energy input is required.

There have been few satellites that have measured the parameters needed for a detailed understanding of what controls the thermospheric density structure. The CHAMP and GRACE satellites have accelerometers, inferring the mass density by assuming that the in-track wind speed is zero and that the exact orientation and drag coefficients of the satellites are known [e.g., Lühr *et al.*, 2004]. While this has provided a wealth of information on the global thermospheric distribution of density, there are still many unanswered questions on the energization of the thermosphere that can not be addressed by this type of measurement alone.

Further, single satellites assist us in our understanding of the climate of the system, but they are greatly hampered when an understanding of the dynamics is desired. For example, Figure 1 shows thermospheric density measurements from five consecutive ascending nodes. There are many interesting things occurring during this time including: (1) the density in the Southern hemisphere increases dramatically at the highest latitudes; (2) the density in the midlatitudes in the southern hemisphere barely increases; (3) the density at the highest latitudes in the northern hemisphere actually decreases; (4) The density in the equatorial region has the strongest reaction

to the energy input; and (5) there is evidence of large-amplitude wave structure in the northern hemisphere midlatitudes.

From this figure it is clear that the thermospheric density can change by a factor of two in less than 90 minutes, which is the approximate orbital period of the CHAMP satellite. In between these orbits, there is no clear indication of what is occurring. While there is an indication of wave structure, it is almost impossible to determine whether the waves are travelling north or south. The waves also gain in amplitude from the high latitudes to the equatorial region, which is opposite from simple geometrical considerations, given that the equatorial region encompasses a larger area than the mid- or high-latitudes.

The example is illustrative of the fundamental problem of single-satellite in situ measurement missions - they can not unravel the dynamics of the system due to the fact that the orbital period is typically longer than the inherent time-scales in the system and that they can not determine whether something is propagating into the field-of-view or is created where it is measured.

Typical satellites that are launched are very large for a few reasons, one of which has been that the size of the electronics has been quite large and heavy. Advances in commercial industry has pushed towards miniaturization of electronics, as clearly indicated by the progression from large desktop computers to laptops to iPads and smart phones. This technology has been embraced by many researchers in Heliophysics, and has resulted in many miniature instruments that, ten years ago, would have had a mass of over 10 *kg*, but now are under one. This miniaturization of instruments fundamentally changes how satellites can be designed, due to the instrument's decreased mass, volume and energy needs. Satellite busses take advantage of miniaturization also - on board computers and memory banks become smaller, radios become smaller, power systems become smaller, etc.

Practically this miniaturization means that satellites can become smaller and smaller or can be complemented with more and more instrumentation. In many ways, both paths are good for science in general. More instruments on one satellite provides more measurements of the relevant quantities that may be important for a physical understanding of the local processes. Smaller spacecraft may allow for more of them to be launched, providing multipoint measurements of the system. This is the path that we are advocating.

There have been two NASA Heliophysics missions that have utilized more than two satellites: Cluster and Themis. Both of these satellite missions have been magnetospheric oriented, studying very different processes in different ways. Cluster is a mission in which the satellites are flown relatively close together in order to understand gradients and meso-scale processes in the magnetosphere. The five Themis satellites are widely separated, allowing for large-scale processes to be examined. Each mission has been wildly successful.

We are advocating the use of constellation missions in the study of the ionosphere and thermosphere in order to understand the dynamics of the system after energy has been deposited at high latitudes. We suggest the use of relatively simple, miniature instrumentation mounted on Micro-Satellites or Nano-Satellites for such a constellation. Such a constellation of satellites deployed in the thermosphere/ionosphere could address many science questions, such as:

- **What are the scale sizes of thermospheric gradients in the auroral zone?** It has often been thought that the thermosphere has large scales and long time constants [e.g., *Innis and Conde*, 2001; *Guo and McEwen*, 2003]. While this may be true in some regions of the thermosphere when there is little activity, there is strong evidence that the thermosphere in the auroral zone is quite dynamic with strong gradients that can evolve rapidly

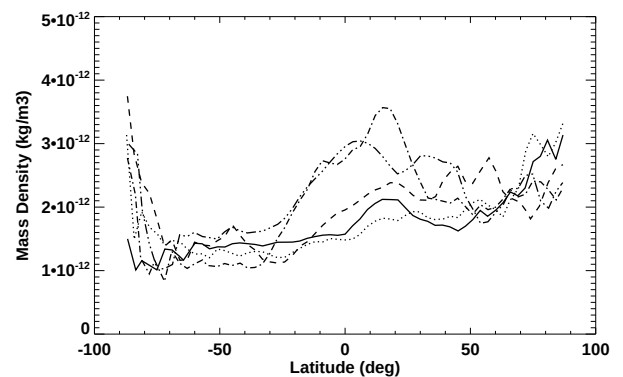


Figure 1: The thermospheric mass density measured by the CHAMP satellite during five consecutive ascending nodes. The order of the orbits in terms of line-style is: solid, dotted, dashed, dash-dot and dash-dot-dot-dot.

[e.g., *Conde and Smith*, 1998; *Lühr et al.*, 2004; *Liu et al.*, 2010]. One of the primary science drivers of such a constellation mission could be to examine the spatial scales within the auroral zone thermosphere as energy is deposited into the system on a wide variety of scales.

- **How correlated are the neutral winds and ion flows on different temporal/spatial scales, and what controls that correlation?** A constellation mission that measures both the ion drifts and neutral winds could easily correlate them across many different spatial scales within the high latitudes.

- **On what scales do upwellings occur due to heating?** Results from Fabry Perot Interferometer data have shown that strong vertical winds can result in the auroral zone [e.g., *Spencer et al.*, 1982; *Crickmore et al.*, 1991; *Conde and Dyson*, 1995; *Price et al.*, 1995; *Guo and McEwen*, 2003]. Very little information is known about these vertical winds. It is unclear whether they are highly localized, resulting in significant horizontal shears in the vertical flow, or whether they occur over much larger scales. A constellation mission could provide direct measurements of these winds, allowing a quantification of the scale sizes, magnitudes and correlations with activity level.

- **What are the scale-sizes of neutrals away from the auroral zone after a heating event?** It can be speculated that the thermosphere may have small scales in the auroral zone, where the Joule and auroral heating may be quite localized [e.g., *Innis et al.*, 2001]. As the increased densities and temperatures propagate away from the source location of the heating, it is unclear how much structure will remain. One of the primary goals of Armada is to quantify the scale-sizes of the density, temperature and neutral winds as a function of distance away from the auroral zone, activity level, and amount of time after a heating event (e.g., time after a sudden increase in the auroral electrojet index, hemispheric power or cross polar cap potential).

- **What is the influence of the background temperature, density and wind structure on the propagation away from the heating location?** In order to quantify how the thermosphere reacts to energy input, the dissipation of the energy needs to be examined. One of the primary dissipation mechanisms within the thermosphere is the advection of the energy to different locations [e.g. *Sutton et al.*, 2005]. Modeling results have shown that after a solar flare, the increased dayside temperature and density cause a wave to propagate to the nightside, where the reconvergence of the wave causes a significant peak in density near the equatorial region around midnight [*Pawlowski and Ridley*, 2008]. The wave then begins to propagate towards the dayside again, but it is propagating against the bulk day-to-night neutral wind flow, which causes the wave to propagate more slowly and to dissipate almost completely before reconverging on the dayside. It is clear that the sound speed will be different on the dayside and nightside, which should result in an asymmetric propagation of gradients away from the high latitudes, possibly allowing the waves to dissipate at different rates. Further, nitric oxide, which is a significant radiative cooler in the thermosphere [*Kockarts*, 1980; *Barth*, 1992], has an asymmetric day to night distribution [e.g., *Cravens and Stewart*, 1978; *Gerard and Roble*, 1988; *Siskind et al.*, 1997; *Ridley et al.*, 1999], which may strongly influence the dissipation rate of the enhanced temperature wave.

- **How do different sized fronts lose coherence as they propagate away from the source?** The CHAMP satellite has shown that high latitude heating events cause perturbations that propagate towards the equator [e.g., *Sutton et al.*, 2005]. Modeling shows that these perturbations should pass through the equatorial region, into the other hemisphere and reconverge at the other pole. This has not been clearly observed in satellite data. Therefore, there is a clear lack of understanding in why disturbances lose coherence as they propagate through the equatorial region and into the opposite hemisphere. Having both ion and neutral measurements on a constellation mission, researchers could investigate whether the coupling between the two influences the dissipation mechanisms.

- **What are the interhemispheric asymmetries in the response of the thermosphere to energy input?** Studies have examined how the magnetic pole offset, the interplanetary magnetic field B_y component and the seasons affect the response of the thermosphere and ionosphere to energy input [e.g., *Thayer et al.*, 1987; *Newell and Meng*, 1988; *Crowley et al.*, 1999; *Lu et al.*, 1994; *Rodger et al.*, 1994; *Gasda and Richmond*, 1998]. Each of the studies that we have described above will include an interhemispheric component to it, where the northern and southern hemispheres are considered separately and together as a system. This will allow the examination of any hemispheric, seasonal, or magnetic differences in the scale-sizes and response of the thermosphere to

energy input.

3 Objectives

This white paper advocates the use of a constellation of satellites to radically improve our understanding of the **dynamics** of the coupled thermosphere/ionosphere system. In order to achieve this goal the following top level requirements must be met:

1. Each satellite needs to be complemented with instrumentation that is capable of measuring both the neutral and ion state variables, such as density, temperature and winds. There are a few different instrument packages that are capable of achieving this requirement, with many of them small enough to fit on very small spacecraft. We specifically advocate the use of the Wind Ion Neutral Composition Suite (WINCS) [Herrero *et al.*, 2009] that is composed of four electrostatic analyzers (two for ions and two for neutrals, each mounted perpendicularly to each other) and two mass spectrometers (one for ions and one for neutrals). This suite of instruments contained in a single enclosure that is extremely small (less than 10cm x 10cm x 10cm) and is less than 1 kg in mass.
2. Each satellite needs to pass through the auroral zone, since this is the primary science zone.
3. There must be enough satellites in a single orbital plane that the return time to a location is less than dynamical time-scales within the system. Because our understanding of the dynamics of the system is quite weak at this point, it would be best to have to return times be variable. This means that the “pearls on a string” satellites within a single plane should have slightly different orbital periods, allowing them to catch up to and pass each other, giving return times between 0 minutes and the orbital period of the fastest satellite. Because of ground-based instrumentation, we understand that the inherent time-scale in the large-scale ionospheric convection is roughly 15 minutes [Ridley *et al.*, 1998], we should have more than six satellites within each plane (i.e., six satellites provides an average of 15 minutes return times to the same place).
4. Multiple planes of satellites should be utilized to provide information on the longitudinal structure and the dynamics of system. The scale-sizes of the thermosphere/ionosphere in longitude are poorly understood, and so it would, once again, be good to have variable spacing in longitude, allowing systematic studies of the longitudinal and local time variability in the the system. The number of orbital planes is quite difficult to specify, because our understanding is so limited. Four planes at equal spacing would allow for the comparison of many different local time sectors, but may not allow for the dynamics of longitudinally propagating waves to be examined. Therefore, it is suggested that four is the minimum that could be deployed in order to begin to understand the global structure of the thermosphere/ionosphere system.
5. A side-benefit from having different orbital periods for each satellite in an orbital plane is that in order to actually accomplish the different periods, the satellites must be deployed at slightly different altitudes or with slightly different ellipticities. Either way, the satellites will pass over or under each other, providing a view of the vertical structure in the thermosphere and ionosphere on small-scales.

4 Cost

It is suggested above that the minimum number of satellites that should be launched is 24 (four orbital planes containing six satellites). If one were to simply assume that the 24 satellites would cost roughly the same as a Themis satellite (assuming that the Themis mission cost approximately \$180M, so each satellite cost 1/5 of this,

or \$36M), such a mission would cost approximately \$864M. This is obviously a simplification, but it may put an upper cap cost on the mission.

There are a few reasons why this cost may be a significant over estimate:

1. The size of the spacecraft could actually be significantly smaller than those within the Themis mission. It is now possible to incorporate sensors such WINCS into NanoSats under 10 kg. Building a smaller spacecraft typically costs less money than a larger spacecraft, due to power, thermal and structural downsizing.
2. Companies exist that build multiple satellites in relatively large quantities. For these types of builds, the companies validate the process of building, and conduct less testing on the individual satellites being produced. A switch to this type of building requires more non reoccurring expenses, but, overall, a huge savings is gained in the production of many systems.
3. At the level of four or five spacecraft, it is still possible to treat the satellites as individual entities and runs communication systems and such the typical way. As the number of satellites grow beyond a small number, the method for communication and operations needs to change. For example, automated scheduling of downloads needs to be done. While the operations cost for a constellation mission may be larger than a single or double satellite mission, it will be significantly less than the cost for a single satellite multiplied by the number of satellites.

It is suggested that building 24 satellites at 10kg each would cost on the order of \$10M each, for a total of \$240M for satellites. Considering that operations costs will be significantly larger than typical satellite missions, this could run as large as \$24M, or 10% of the cost of the satellites. With costs of management, systems engineering, science, etc., the cost of such a mission is estimated to be on the order of \$350M, excluding launch costs.

There are two alternatives to launch 24 10kg spacecraft to relatively high inclinations (above 80°) into four different orbital planes: (1) launch all the satellites on a single large rocket with a bus that is capable of making small orbital changes needed to deploy the spacecraft into the slightly different inclinations; (2) launch two smaller rockets in which the upper stage can perform a single orbital maneuver that can change inclinations by a small amount, and disperse two orbital planes. Each method has different costs and benefits. With a single launch vehicle and a bus to deliver them into orbits, they can be placed more exactly into the orbital planes, and a second rocket is not needed. For two rockets, the orbital insertion may not be as precise, but there is redundancy in the system if one of the rockets were to fail. If the single rocket were a Taurus and the two other rockets were Pegasus, then the cost of building a deployment bus along with the cost of the Taurus may be comparable to the cost of two Pegasus with simple deployers attached to the upper stage. Either way would cost approximately \$100M, bringing the total mission cost up to about \$450M.

5 Relevance

National Academy of Sciences Decadal Survey for Solar and Space Physics has the following highlights that are specifically addressed by such a thermosphere/ionosphere constellation mission:

- 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.

- Advanced science instrumentation. Highly miniaturized sensors of charged and neutral particles and photons will be essential elements of instruments for new solar and space physics missions. **Recommendation: NASA should continue to assign high priority, through its recently established new instrument development programs, to supporting the development of advanced instrumentation for solar and space physics missions and programs.**

- Gathering and assimilating data from multiple platforms. Future flight missions include multipoint measurements to resolve spatial and temporal scales that dominate the physical processes that operate in solar system plasmas. **Recommendation: NASA should accelerate the development of command-and-control and data acquisition technologies for constellation missions.**

The 2010 Heliophysics Roadmap for Space and Technology has the following highlights that could be specifically addressed by such a constellation mission:

- Origins of Near Earth Plasma (ONEP): Understand the origin and transport of terrestrial plasma from its source to the magnetosphere and solar wind.
- Ion-Neutral Coupling in the Atmosphere (INCA): Understand how neutral winds control ionospheric variability.
- Climate Impacts of Space Radiation (CISR): Understand our atmospheres response to auroral, radiation belt, and solar energetic particles, and the associated effects on nitric oxide (NO) and ozone.
- Dynamic Geospace Coupling (DGC): Understand how magnetospheric dynamics provide energy into the coupled ionosphere-magnetosphere system.

In addition, such a constellation could address the following research focus areas: **F2** Particle acceleration and transport; **F3** Ion-neutral coupling; **F4** Creation and variability of magnetic dynamos; **H2** Earths magnetosphere, ionosphere and upper atmosphere; **H3** Role of the Sun in driving change in the Earths atmosphere; **H4** Apply our knowledge to understand other regions; **J1** Variability, extremes and boundary conditions; and **J4** Effects on and within planetary environments.

6 Technical Challenges

In order for NASA to accomplish a constellation mission with 24 or more satellites, three areas should be addressed specifically:

1. NASA should recognize that single satellites launched on single rockets are single points of failure, while a satellite mission with many satellites can lose a certain percentage of them without a serious degradation of the science that can be accomplished. This is not to say that the quality of the satellites should be lowered in any way, but there should be a fundamental shift in validating the build/integrating/testing process instead of a focus of the individual satellites when considering constellations. This will significantly reduce the per spacecraft cost which may enable more satellites to be deployed at the possibility that a slightly higher percentage may fail. For example, if 48 satellites could be built under this scheme instead of 24 under the current scheme, and the failure rate only increases by a few percentage points, then the mission would have significantly more satellites in orbit, even though a small percentage may fail sooner than expected.
2. NASA needs to fund programs focused on the miniaturization of both instruments and subsystems such as attitude determination (e.g., star trackers, horizon sensors, rate gyros, etc.) and communication (e.g., UHF, S-band and X-band radios). By shrinking these critical subsystems, more small instrumentation may have the capability of being put on Micro- and Nano-Sats.
3. NASA needs to upgrade its communication network to enable the possibility of many simultaneous contacts, very fast acquisition of new contacts, and automated handling uplinks and downlinks. With the

current network, it is very hard to envision a large constellation mission that is taking significant amounts of data to succeed using only the NASA near-Earth network.

7 Summary

This white paper advocates a Heliosphysics mission to study the dynamics of the ionosphere/thermosphere system that would be composed of a constellation of 24 Nano-Satellites with a mass of less than 10kg. These satellites would be dispersed with six satellites in four different orbital planes. Giving each satellite a slightly different periodicity would allow the satellites within the plane to spread out over time, and giving each of the planes a slightly different inclination would allow the planes to spread out over time. The scientific aim that is suggested in this white paper is to examine the response of the thermosphere/ionosphere system to high latitude heating, so each of the planes would need to be at approximately 80° - 85° inclination angle.

References

- Barth, C. (1992), Nitric Oxide in the lower thermosphere, *Planet. Space Sci.*, 40, 315.
- Bowman, B., W. K. Tobiska, F. Marcos, and C. Valladares (2008), The j2006 empirical thermospheric density model, *J. Atmos. Sol.-Terr. Phys.*, 70, 774, doi:10.1016/j.jastp.2007.10.002.
- Bruinsma, S., J. M. Forbes, R. S. Nerem, and X. Zhang (2006), Thermosphere density response to the 20-21 November 2003 solar and geomagnetic storm from CHAMP and GRACE accelerometer data, *J. Geophys. Res.*, 111, 6303, doi:10.1029/2005JA011284.
- Conde, M., and P. L. Dyson (1995), Thermospheric vertical winds above mawson, antarctica, *Journal of Atmospheric and Terrestrial Physics (ISSN 0021-9169)*, 57, 589.
- Conde, M., and R. W. Smith (1998), Spatial structure in the thermospheric horizontal wind above poker flat, alaska, during solar minimum, *Journal of Geophysical Research*, 103, 9449, doi:10.1029/97JA03331.
- Cravens, E., and A. I. Stewart (1978), Global morphology of nitric oxide in the lower *E*-region, *J. Geophys. Res.*, 83, 2446.
- Crickmore, R. I., J. R. Dudeney, and A. S. Rodger (1991), Vertical thermospheric winds at the equatorward edge of the auroral oval, (*Thermospheric and ionospheric dynamics; Symposium*, 53, 485.
- Crowley, G., A. Ridley, D. Winningham, R. Frahm, J. Sharber, J. R. III, and R. Roble (1999), On the hemispheric symmetry in thermospheric nitric oxide, *Geophys. Res. Lett.*, 26, 1545.
- Fuller-Rowell, T., and D. Rees (1980), A three-dimensional, time-dependent, global model of the thermosphere, *J. Atmos. Sci.*, 37, 2545.
- Fuller-Rowell, T. J., M. V. Codrescu, R. J. Moffett, and S. Quegan (1994), Response of the thermosphere and ionosphere to geomagnetic storms, *J. Geophys. Res.*, 99, 3893.
- Gasda, S., and A. Richmond (1998), Longitudinal and interhemispheric variations of auroral ionospheric electrodynamics in a realistic geomagnetic field, *J. Geophys. Res.*, 103, 4011.
- Gerard, J., and R. Roble (1988), The role of nitric oxide on the zonally averaged structure of the thermosphere: Solstice conditions for solar cycle maximum, *Planet. Space Sci.*, 36, 271–279.
- Guo, W., and D. J. McEwen (2003), Vertical winds in the central polar cap, *Geophysical Research Letters*, 30, doi:10.1029/2003GL017124.
- Herrero, F., L. Beasley, R. Comes, and J. Marbourg (2009), The gas kinetic method for measurements of the neutral wind vector, temperature and densities in the thermosphere, www.thermosphere.com, website:.
- Innis, J. L., and M. Conde (2001), Thermospheric vertical wind activity maps derived from dynamics explorer-2 wats observations, *Geophysical Research Letters*, 28, 3847, doi:10.1029/2001GL013704.
- Innis, J. L., P. A. Greet, and P. L. Dyson (2001), Thermospheric gravity waves in the southern polar cap from 5 years of photometric observations at davis, antarctica, *Journal of Geophysical Research*, 106, 15,489, doi:10.1029/2001JA000032.
- Kockarts, G. (1980), Nitric oxide cooling in the terrestrial thermosphere, *Geophys. Res. Lett.*, 7, 137.

- Liu, R., H. Lühr, and S.-Y. Ma (2010), Storm-time related mass density anomalies in the polar cap as observed by champ, *Annales Geophysicae*, 28, 165.
- Lu, G., et al. (1994), Interhemispheric asymmetry of the high-latitude ionospheric convection pattern, *J. Geophys. Res.*, 99, 6491.
- Lühr, H., M. Rother, W. Köhler, P. Ritter, and L. Grunwaldt (2004), Thermospheric up-welling in the cusp region: Evidence from champ observations, *Geophys. Res. Lett.*, 31, 06,805, doi:10.1029/2003GL019314.
- Newell, P., and C.-I. Meng (1988), Hemispherical asymmetry in cusp precipitation near solstice, *J. Geophys. Res.*, 93, 2643.
- Pawlowski, D., and A. Ridley (2008), Modeling the thermospheric response to solar flares, *J. Geophys. Res.*, 113, A10,309, doi:10.1029/2008JA013182.
- Price, G. D., R. W. Smith, and G. Hernandez (1995), Simultaneous measurements of large vertical winds in the upper and lower thermosphere, *J. Atmos. Terr. Phys.*, 57, 631.
- Prölss, G. W. (2004), Space Weather Effects in the Upper Atmosphere: Low and Middle Latitudes, in *Lecture Notes in Physics, Berlin Springer Verlag, Lecture Notes in Physics, Berlin Springer Verlag*, vol. 656, edited by K. Scherer, H. Fichter, and B. Herber, p. 193, doi:10.1007/b100037.
- Ridley, A., C. Clauer, G. Lu, and V. Papitashvili (1998), A statistical study of the ionospheric convection response to changing interplanetary magnetic field conditions using the assimilative mapping of ionospheric electrodynamics technique, *J. Geophys. Res.*, 103, 4023.
- Ridley, A., G. Crowley, R. Frahm, D. Winningham, J. Sharber, J. R. III, and R. Roble (1999), Variations of the nitric oxide mass mixing ratio in the thermosphere as a function of kp , *Geophys. Res. Lett.*, 26, 1541.
- Roble, R., E. Ridley, A. Richmond, and R. Dickinson (1988), A coupled thermosphere/ionosphere general circulation model, *Geophys. Res. Lett.*, 15, 1325.
- Rodger, A., M. Pinnock, J. Dudeney, J. Watermann, O. de la Beaujardiere, and K. Baker (1994), Simultaneous two-hemisphere observations of the presence of polar patches in the nightside ionosphere, *Ann. Geophys.*, 12, 642.
- Siskind, D., J. Bacmeister, M. Summers, and J. R. III (1997), Two-dimensional model calculations of nitric oxide transport in the middle atmosphere and comparison with Halogen Occultation Experiment data, *J. Geophys. Res.*, 102, 3527.
- Spencer, N. W., L. E. Wharton, G. R. Carignan, and J. C. Maurer (1982), Thermosphere zonal winds - vertical motions and temperature as measured from dynamics explorer, *Geophys. Res. Lett.*, 9, 953, doi:10.1029/GL009i009p00953.
- Sutton, E. K., J. M. Forbes, and R. S. Nerem (2005), Global thermospheric neutral density and wind response to the severe 2003 geomagnetic storms from champ accelerometer data, *Journal of Geophysical Research*, 110, doi:10.1029/2004JA010985.
- Thayer, J., T. Killeen, F. McCormac, C. Tschan, J.-J. Ponthieu, and N. Spencer (1987), Thermospheric neutral wind signatures dependent on the east-west component of the interplanetary magnetic field for northern and southern hemispheres as measured by Dynamics Explorer-2, *Ann. Geophys.*, 5A, 363.

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.

Acronyms and Glossary

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 Heliophysics 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 Ejections and other solar activity, launched 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.

The Atmospheric Density Specification Experiment

Patrick A. Roddy and John O. Ballenthin

Summary

Current atmospheric models essential to accurate orbit prediction of satellites and space debris are inadequate now and will remain so in the future without new measurements of neutral density and composition. Our current understanding of the neutral thermosphere is due in large part to neutral mass spectrometers flown on a number of satellite mission in the 1970's and 80's. The methods and techniques used to assimilate that neutral mass spectrometer data are well established and need only the addition of new, high-accuracy measurements to greatly increase our understanding of how the neutral thermosphere responds to inputs from the sun. The Air Force Research Laboratory's Atmospheric Density Specification mission (ADS) is a mass spectrometer and neutral wind sensor package that if flown is capable of supplying the quality of measurements required to reduce errors in the modeled thermospheric neutral density from 20% to 5%.

Mission Objectives: Acquire comprehensive, global high accuracy atmospheric dataset across all local times, seasons and levels of geomagnetic activity for use in empirical and physics based thermospheric neutral density models.

Instrumentation

The ADS mission will achieve these objectives by flying AFRL's Absolute Density Mass Spectrometer (ADMS) and the University of Texas at Dallas's Anemometer Cross-Track Measurement Experiment (ACME) on a three axis stabilized satellite in an elliptical near-polar orbit. Figure 1 shows the two instruments. The ADMS instrument is neutral quadrapole mass spectrometer capable of operating in either closed or open-source modes. In closed-source mode it capable of measuring absolute neutral density to an accuracy of 5% with the use of an onboard calibration source. In open-source mode it capable of measuring the ram energy of incoming neutrals and thus determine the ram component of the neutral wind vector. The orthogonal components of the neutral wind are then determined by the ACME instrument. ACME determines the arrival angle of the neutral wind by measuring a pressure differential in series of chambers exposed to the ram. The ACME instrument is identical to the Cross-Track neutral wind instrument flying on the Air Force C/NOFS satellite.

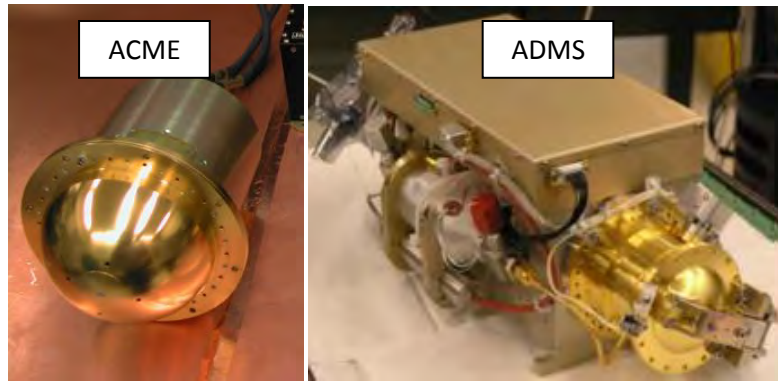


Figure 1. The ACME (left) and ADMS (right) instruments.

Measurements

- Absolute neutral density accurate to 5%
- Neutral composition
- Neutral temperature
- 3D neutral wind vector

The ADS instrument will directly measure absolute neutral density, neutral composition and the ram component of the neutral wind. ADS will be able to determine neutral temperatures from density and composition scale heights for non-circular orbits and possibly from energy analysis of the ram wind. The ACME instrument will provide cross track winds as well as some measure of neutral density.

Summary

Our models of the neutral thermosphere are inadequate for our current and future needs; this is due primarily to them being data starved. We as a community desperately need to field new sensors, whether they be ADS or not, to address this problem.

The Tropical Coupler Mission

D.E. Rowland¹, R. F. Pfaff¹, W. Farrell¹, R.L. Bishop²

¹NASA / Goddard Space Flight Center

²The Aerospace Corporation

Executive Summary

Tropical Coupler will, for the first time, provide a complete understanding of the forcing of the ionosphere and lower thermosphere by stratospheric and tropospheric effects at low latitudes, and the critical role of ion-neutral coupling in driving both large-scale dynamics and internal instabilities of the low-latitude I-T system. This science objective is both extremely “local” and very universally applicable. The same types of phenomena drive planetary atmospheres throughout the solar system, to a lesser or greater extent depending on their solar / magnetospheric forcing and the presence or absence of a magnetic field. Tropical Coupler will be the ultimate “local” mission in that it provides critical measurements to help us understand our home planet – the effects of the sensible atmosphere, in which we live, on the upper atmosphere, upon which we depend so much for technological applications (radio communications, GPS, satellite drag, etc.). Tropical Coupler is essential to build on recent advances made by the C/NOFS mission, and to complete the picture of energy transfer and coupling in the Sun-Earth system. By launching in the near future, while SDO and other magnetospheric assets are returning data about those forcing terms, Tropical Coupler will be able to disentangle the effects of the lower atmosphere driving from these other sources of energy.

Introduction and Motivation

The upper atmosphere of the Earth (and other planets) is energetically driven by three sources: the solar irradiance (in the EUV, which deposits its energy at high altitudes), solar wind / magnetospheric forcing (predominantly at high latitudes for planets with an internal magnetic field, at all latitudes for non-magnetic planets), and from below (tropospheric and stratospheric forcing). The critical altitude range from 80 km (base of the ionosphere) to 600 km (base of the exosphere) is where much of this energy is ultimately deposited, either directly, or at the end of a chain of processes (see **Fig 1.**).

While many of these processes take place at all latitudes on the Earth, the low-latitude region is unique in several respects:

- 1) Understanding the ion-neutral coupling and lower-atmospheric forcing is made simpler by the reduced effects of magnetospheric and solar wind coupling at low latitudes.
- 2) Large-scale weather systems that produce strong convective gravity waves are much more common than at high latitudes

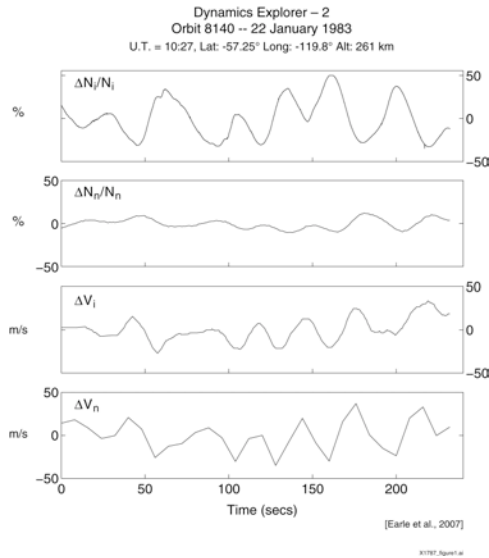


Figure 3: DE-2 observations of gravity waves in the thermosphere, at 261 km

coupling (wind-shear driven, in the 200 km region) is required to “seed” the instabilities and provide the initial growth, before the RT instability can take over.

Another example of how ion-neutral coupling drives energy and momentum transfer from the lower atmosphere into the ionosphere-thermosphere system is shown in Figure 3. This figure shows DE-2 observations of gravity waves in neutral and ion measurements at 261 km altitude at night. These gravity waves, whether propagating directly from tropospheric and stratospheric sources, or secondarily generated in the lower thermosphere, can transport energy easily across the magnetic field, resulting in

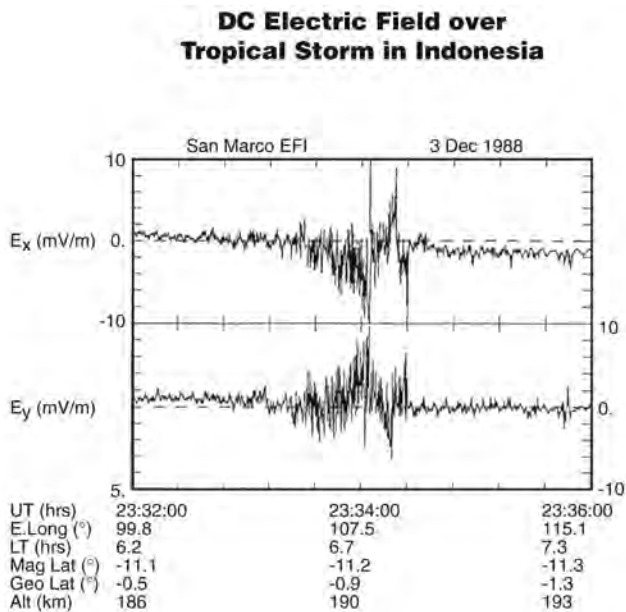


Figure 4: San Marco observations of electric fields driven by tropospheric activity [Farrell et al., 1994]

velocity, showing a characteristic “vortex” near the terminator, followed by the development of the radar “plume” indicating plasma uplift and the formation of strong irregularities in the plasma density. These types of internal ionospheric instabilities are strongly modulated by tidal and gravity wave forcing, and are strongly dependent on neutral wind shear and ionospheric conductivity profile.

This ion-neutral coupling is a critical feature of the partially ionized gas in the Earth’s upper atmosphere, and is required to drive much of the observed “electrodynamics” features that are observed. In particular, the generalized Rayleigh-Taylor instability, long believed to be the cause of the small scale structure, has too slow a growth rate to produce the observed dynamics on the observed timescales. Thus, lower-altitude ion-neutral

coupling (wind-shear driven, in the 200 km region) is required to “seed” the instabilities and provide the initial growth, before the RT instability can take over.

A final example of the as-yet not understood coupling of the lower atmosphere to the I-T system is given in Figure 4 [Farrell et al., 1994], which shows San Marco DC electric fields in the ionosphere, at 190 km, above a tropical storm. While San Marco was below the F-peak in this example, the expectation is that the ionosphere should short out these DC electric field structures. The fact that they are still observable at this altitude must mean that some secondary process

involving ion-neutral coupling is

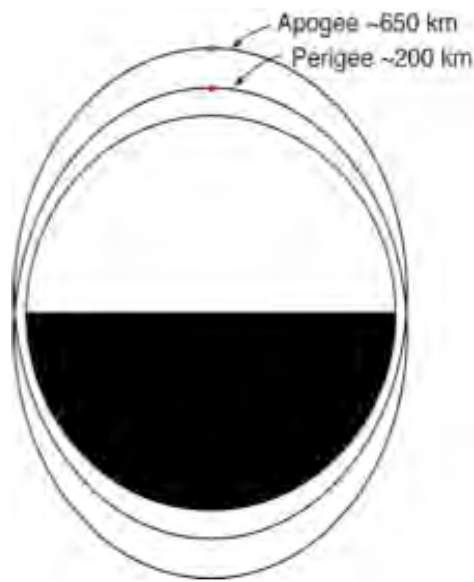


Figure 5: Tropical Coupler orbits

sustaining these structures and modifying the I-T dynamics at high altitudes. Note that the electric field associated with the thunderstorm is an order of magnitude larger than the background convection at these altitudes, implying that these events can be strong local drivers of ionospheric dynamics.

The next step to bring more complete understanding about the roles of ion-neutral coupling in mediating tropospheric/stratospheric forcing of the I-T system is to fly a new mission, called “Tropical Coupler” (TC). TC will consist of two identically instrumented spacecraft, with a suite of in situ electrodynamics and gas dynamics instruments, launched into a single orbit plane, with opposing arguments of perigee so that they are 180 degrees in orbital phase. This

configuration will result in repeated conjunctions between the two spacecraft, and allow for measurement of low latitude phenomena at all local times and longitudes, and over a range of altitudes (with two different altitudes at the same footpoint measured simultaneously). **Figure 5** shows the orbit concept. The spacecraft will have apogees of 650 km and perigees near 200 km, and contain onboard propulsion systems to counteract drag and achieve the initial orbit from a single launch vehicle.

Tropical Coupler Science Objectives

- Determine the extent and significance that tropical storms in the troposphere play in producing localized electric fields and neutral gas motions in the ionosphere.
 - o Tropical Coupler will measure the electric fields at lower altitudes directly above large tropical storms – by measuring the altitude profile of these electric fields over the range 200 to 600 km, we will be able to compare to models of gravity wave propagation and dissipation, and ionospheric conductivity, to see what the causes and consequences of these large fields are.
- Understand how large-scale neutral winds, gravity waves, and ion drifts are coupled in the low latitude upper atmosphere and how they respond to solar variations and to variable influences from the lower and middle atmosphere.
 - o Tropical Coupler will perform low altitude observations of neutral winds and gravity waves with simultaneous in situ measurements of ion drift, plasma conductivity, and other drivers. Do this over a range of altitudes
- Discover the cause of planetary-scale plasma upwellings at low latitudes and investigate the effect of unstable flux tubes on inner plasmasphere dynamics.
 - o Tropical Coupler will perform low altitude, low-latitude observations of plasma density and ion drift, from the large scale down to meter scale, to

determine what causes the formation of plasma depletions, what the drivers and instability mechanisms are, and what determines the formation of small-scale structure that can interfere with radio signals used for communication and navigation.

- Determine the effect on lower thermosphere chemistry due to gravity wave breaking and turbulence, as well as plasma uplift.
 - o Tropical Coupler will measure local composition, ions and neutrals, over the altitude range 200-600 km. Supplement with remote sensing measurements of the composition profiles with altitude over the range 100-300 km.
- Explain the chaotic and explosive transfer of energy between irregularity scale lengths ranging from 1000 km to 1 cm that occur in the low latitude ionosphere at night.
 - o Tropical Coupler will use multiple platforms to look at both the “seed” region (200-300 km) and the “fully developed plume” region (>400 km) to determine both the instabilities responsible, and to dramatically improve models of irregularity development and evolution.

Measurement Requirements

Tropical Coupler will contain a suite of high heritage instruments, in order to measure both the electrodynamics and gasdynamic properties of the lower thermosphere-ionosphere system. Critical measurements include neutral wind, ion drift, plasma density, ion and neutral composition, and remote sensing to return the plasma density and neutral density profiles with altitude.

in situ neutral wind	accurate to +/- 5 m/s, at least 50 km spatial resolution
in situ ion drift	accurate to +/- 5 m/s, at least 5 km spatial resolution
in situ electric field	accurate to +/- 0.3 mV/m, 10m scale spatial resolution
or finer	
in situ plasma density (down to 1 meter scale)	10% accurate, 100-10 ⁷ /cc
in situ ion composition from 1-46 amu,	50 km spatial resolution
in situ neutral composition from 1-46 amu,	50 km spatial resolution
in situ neutral density,	10% accurate, 50 km spatial resolution
vertical electron density profile to F peak (topside and bottomside)	
remote detection of irregularities / scintillation	
neutral density and composition vertical profiles (limb scan)	100-300 km regime

Mission Concept

Three (or two) identical three-axis stabilized spacecraft, with onboard propulsion, They have low-inclination orbits that are in the same plane but nearly 180 degrees out of phase. Apogee near 650 km and perigee near or just below 200 km results in the best measurements of density structure formation, neutral wind and electric field drivers, and complete tracking of the near-F-peak density profile using onboard sounding. The

two spacecraft will often be in near-conjunction, remaining within a few hundred km horizontally of each other, while spanning this altitude range. This orbit also provides for some separation of spatial variations from temporal variations, similar to “pearls on a string” orbit. An optional third identical spacecraft stays in the initial 400 km circular orbit, and provides a stable altitude reference for imaging and a third altitude reference point. Because the three spacecraft are identical, the system has flexibility and robustness.

Launch Vehicle and Orbit

We assume a single Taurus II or similar launch vehicle, launching the three (or two) TC spacecraft into a low inclination (below 30 degrees) 400 km circular orbit. Then the TC spacecraft (or two of three) execute propulsion maneuvers to go into the science orbit of 200 x 650 km, with opposite arguments of perigee, but “in phase” so that they are constantly near conjunction, both with each other and with the optional third “reference spacecraft” at 400 km. Propulsion is then used only for constellation maintenance and drag makeup.

Cost Estimate: **\$572M** (assumes three identical spacecraft, launched on a single Taurus II launch vehicle). Assumes risk class C.

1.0 Mission Management:	6% of WBS 5+6 = \$12M
2.0 Systems Engineering:	6% of WBS 5+6 = \$12M
3.0 Safety and Mission Assurance:	6% of WBS 5+6 = \$12M
4.0 Science:	\$35M for 2 years, including GI program, SOCs, support for some ground instrumentation
5.0 Instrumentation:	\$50M for first copy, \$25M for each subsequent set = \$100M
6.0 Spacecraft bus:	\$40M for first copy, \$30M for each subsequent set = \$100M
7.0 Mission Ops:	\$10M initial (MOC plus ground station setup), plus \$5M per year per spacecraft = \$10M + \$5M*6 = \$40M
8.0 Launch Vehicle and Services	\$100M
9.0 Ground Systems	\$9M
10.0 Systems Integration and Test	\$15M (\$5M per spacecraft)
11.0 Education and Public Outreach	\$5M (1%)
Total before reserves:	\$440M
Project-Level Reserves (30%):	\$132M
Total Mission Cost with reserves:	\$572M

If we delete the 400 km circular orbit spacecraft, and fly only two identical spacecraft, the total mission cost with 30% reserves goes to **\$467M** (launch vehicle cost remains the same).

Schedule

2 year prime mission

Co-ordinated Campaigns:

- Utilize ionosonde, ISR, and Global Lightning Network with appropriate space-based observations.
-

Modeling Efforts:

- Continue efforts to extend models ionospheric/thermospheric models such as TIEGCM to include lower altitude forcing.
- Extend low altitude tropospheric models into the upper atmosphere.
- Include low altitude forcing in other assimilative models such as SAMI2, GAIM, IDA4D by using low-altitude model output as lower boundary conditions.

Magnetic Meridian Ring of Incoherent Scatter Radars supporting science discovery and tracing AIM weather, climate, and global change

Ennio Sanchez, Anja Strømme, Anthony van Eyken, Craig Heinselman, John Kelly, and Michael Nicolls, SRI International, 333 Ravenswood Avenue, Menlo Park CA 94025.

Contact: Ennio Sanchez: ennio.sanchez@sri.com



Figure 1: Conceptual Magnetic Meridian Ring: Like early map makers, we know much about geospace, but lack the big picture. The Magnetic Meridian Ring concept exploits existing and planned facilities and partnerships to provide comprehensive measurements along a complete magnetic meridian for the first time.

Aeronomers have made huge strides in understanding the detailed physics and chemistry of the upper atmosphere and in designing, constructing and operating advanced facilities to monitor, measure, and in some cases, perturb that medium. The big picture, with its accompanying ability to predict the behavior of the geospace system both in response to natural, solar, and anthropogenic factors, remains elusive. Neither existing ground-based facilities, nor current or planned space missions, can provide the detailed, reliable, global coverage of geospace required to support both new and insightful scientific understanding

The Magnetic Meridian Ring addresses an overarching scientific and technical need for continuous, global, high quality, and reliable data determined under all conditions and covering all magnetic latitudes in both hemispheres. By exploiting existing and planned facilities, the plans and aspirations of international partners, our proven and cost-effective technology, and the favorable distribution of available landmasses around the Pacific Rim, the Magnetic Meridian Ring at last provides global system scientists, modelers, and data consumers with the high quality observational and monitoring capabilities they require.

and future operational capabilities to trace AIM weather, climate, and global change.

Current incoherent scatter radars (ISRs), developed and deployed following the recommendations of the first NRC Decadal Survey [2003], have the ability to operate reliably, remotely, and largely autonomously for extended periods and the systems to build, deploy, operate and maintain them are well developed. For the first time, it is now practical to propose a global ISR deployment capable of providing the precision measurements required.

We propose the installation of a Magnetic Meridian Ring of ISRs and related instrumentation which, taking advantage of existing installations, planned installations, and potential international collaborations, will provide continuous detailed data along a full geomagnetic meridian encircling the globe and affording scientists, modelers, and data consumers uninterrupted coverage including inter-hemispheric, day-night, and seasonal measurements under all conditions and at all times.

Based on current ISR technology, as already deployed at Poker Flat, Alaska, and Resolute Bay, Canada, approximately \$40M will deploy and operate a single facility for ten years. Seven to ten new sites, each employing two phased array radars, together with potential international partnerships, will drive major scientific progress in (for example): geospace science, system science, solar terrestrial interactions, and planetary energy budget, geospace system modeling, now, and fore-casting as well as providing the reliable and detailed inputs needed to support effective geospace management, threat evaluation, and mitigation procedures, as recommended and elucidated in the NSF GeoVision Report [2009].

*'Just as a thunderstorm cannot be properly analyzed with a single weather station, [geospace] cannot be properly studied with a single satellite in the vastness of space'.
Kissinger and McPherron (2010).*

1: Science Concept

Extending observational capabilities over larger areas of geospace provides both an essential foundation for continuing cutting-edge scientific research and an essential input to forecasting activities, including space weather and climate change/disruption. The most effective way to maximize observational coverage is to locate the most capable instruments at the accessible edges of geospace.

The previous Decadal Survey [2003] recognized the importance of global networks of ground-based ionospheric and geomagnetic measurement stations as integral components of the arrays needed to solve the complex, nonlinear, coupled atmosphere-ionosphere-magnetosphere systems.

Incoherent Scatter Radars (ISRs) are ideal for studying energy and mass transport within the Atmosphere-Ionosphere-Magnetosphere system (A-I-M) and the effects of Solar Wind – Magnetosphere interactions. ISRs measure both the predominantly horizontal redistribution of energy about the planet and vertically between atmospheric regions, and between the lower atmosphere and the ionosphere and magnetosphere.

Facilities can be arranged in many different configurations to sample various regions of the A-I-M system. However, there is one arrangement that will perform all the measurements that are needed to answer each of the fundamental questions outlined below. This consists of a ring of facilities located along a complete magnetic meridian encircling the globe. The proposed configuration provides simultaneous day and night measurements (when the ring is at or near the noon-midnight meridian), measurements along the day-night terminator (when the ring is in the dawn-dusk meridian), and inter-hemispherical asymmetries and summer-winter asymmetries at all latitudes. It can follow the latitudinal propagation of ionosphere-thermosphere disturbances such as neutral winds and gravity waves with time resolution of a few minutes to hours, as well as long term circulation effects including the redistribution of ice and aerosol particles in the mesosphere.

Figure 1 shows the conceptual Magnetic Meridian Ring constructed around the magnetic meridian passing through existing and planned ISRs in the American sector, passing close to the South Atlantic Anomaly and including radars and instrumentation in Southeast Asia, China, and Russia. Some facilities that would be part of the Magnetic Meridian Ring are already in operation, as discussed in Section 4, or in

detailed planning, such as the Antarctic Radars proposal, Figure 2, and the AIREs radar in Argentina. Other Magnetic Meridian choices are possible, but the Figure 1 concept maximizes the use of existing installed infrastructure, capitalizes on the availability of land masses to maximize coverage and simplify support, and provides excellent opportunities for meaningful and effective international collaboration.

The step from local to global measurements mandated by system science objectives in the next ten years requires three types of measurements.

(1) Tracking Space Weather effects from high- to low-latitudes and between high and low altitude.

Satellite-borne imagers and field/particle instruments, as well as ground-based instruments such as coherent scatter radars, optical imagers, and magnetometers, have already started to expose and document fundamental clues about the onset and development of storms and substorms, the principal mechanisms of transient energy transfer from the solar wind to the A-I-M system. However, an ISR is the only instrument that can simultaneously measure convection, aurora, and particle/dynamical boundaries, independently of conditions and without temporal/spatial ambiguities. Any individual ISR can make simultaneous measurements of plasma density (particle energy), temperature (electron energy dissipation and ion heating), and flow (elucidating reconnection dynamics and magnetic tension) that are crucial to discriminate candidate substorm triggering mechanisms and to differentiate among competing processes affecting precipitation and loss of radiation particles. However, during storm and substorm conditions,

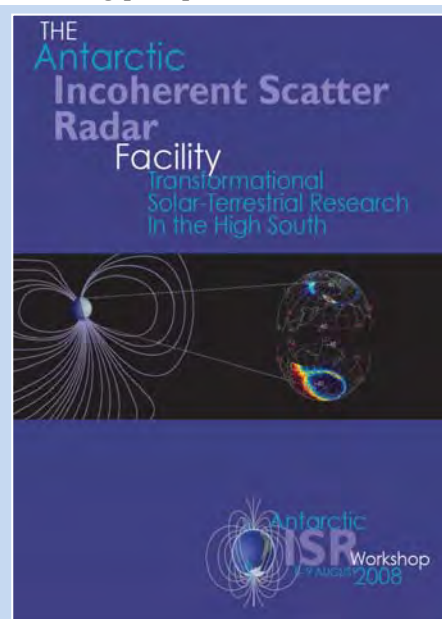


Figure 2: Antarctic Radar Report: An international workshop developed and refined the scientific imperative to gather ISR data in the southern polar cap and auroral zones, contemporaneously with matching observations in the northern hemisphere.

the auroral and sub-auroral zones move rapidly and over large distances that can exceed the field of view of any individual ISR. A ring of ISRs distributed along a meridian will provide the inter-hemispheric and global coverage not so far possible. ISRs in the high-latitude regions will monitor processes in the Earth's plasma sheet while sub-auroral and low latitude observations will reveal the effects of processes in the radiation belts, such as E-field penetration, as well as global phenomena such as Stratospheric Warmings (STRATWARMs).

So important is the study of the latter region that one of NASA's geospace missions, the upcoming Radiation Belt Storm Probes (RBSP) mission, recommended by the previous Decadal Survey as part of the Storm Geospace Network and currently scheduled for launch in May 2012, addresses it specifically. However, the ionospheric component of the Network, also recommended by the same Survey, was not realized. A latitudinal ring of ISRs can replicate many of the fundamental observations that ionospheric satellites would have accomplished, with the advantage that measurements can be performed simultaneously at all latitudes. Measurements from high-inclination orbiting satellites tracing the same meridian as the ISR ring would only provide snapshots of the ionosphere while the Magnetic Meridian Ring will cover much that the ionospheric component of the Storm Geospace Network would have provided.

(2) Tracking of the propagation of anthropogenic and natural atmospheric effects between high- and low-latitudes, between low- and high-altitudes and between hemispheres

The Earth's atmosphere is a multi-dimensional, multi-layered boundary region that responds to myriad influences, with neutral air gravity waves, planetary waves and tides forcing it from below, and solar wind inputs, magnetospheric inputs, and particle precipitation driving it from above. Unraveling the competing processes is particularly challenging at high latitudes, where the ionospheric drivers are strongly influenced by magnetospheric inputs. A systems approach is necessary. Multi-point

measurements with multiple instruments at different locations in both hemispheres constitute the appropriate method to achieve this objective.

Over the last two decades, signs of global changes have been seen over the entire planet. Near the Earth's surface, some of the changes represent global warming. However, “greenhouse cooling” is expected in response to increasing amounts of radiatively active gases, mainly carbon dioxide, in the middle (stratosphere and mesosphere) and upper (thermosphere) atmosphere. Compositional redistributions in the upper atmosphere are also expected in association with changes in the temperature profile.

Anthropogenic and natural effects in the atmosphere will be readily detected in the long term (multi-year) data records of the Magnetic Meridian Ring through changes to the composition and topology of the ionosphere, mesosphere, and lower atmosphere.

The northern and southern upper mesospheres are dynamically connected through a longitudinal summer-to-winter flow driven by gravity wave forcing. Mass conservation causes the air to rise at the summer pole with the associated adiabatic cooling driving the summer mesopause to below 130K, making it the coldest place in the Earth's vicinity. Noctilucent clouds (NLCs) and Polar Mesospheric Summer Echoes (PMSEs)—strong radar backscatter at HF, VHF, and UHF frequencies—in the ~82–85 km altitude summer mesopause are manifestations of the complex coupling processes taking place in this region, although the causes of noctilucent clouds and PMSE are far from understood. PMSE occurrence and intensity are believed to depend strongly on factors such as the temperature at ~87 km, the existence of available water vapor for ice formation in the mesopause region, the density and size of aerosols acting as

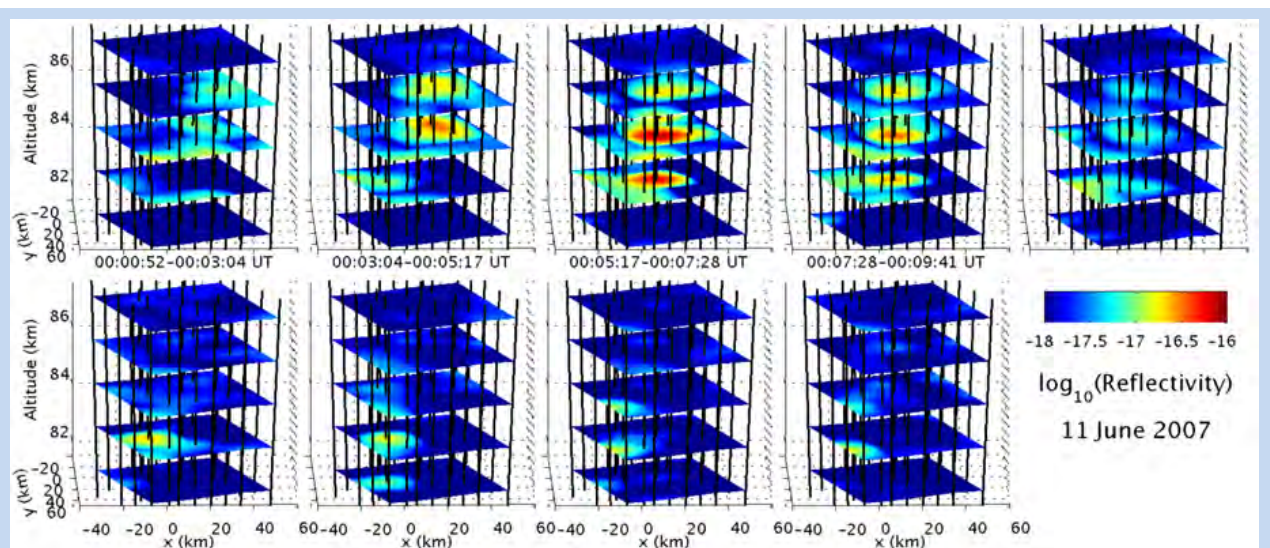


Figure 3: ISR Atmospheric Imaging: Three-dimensional snapshots of a passing layer of ionospheric irregularities. The axes are zonal distance (x , positive east), meridional distance (y , positive north), and altitude, 81.5–87.5 km in 1.5 km steps, relative to the radar location (Nicolls et al., 2007).

nucleation kernels, gravity wave activity acting as a mesopause refrigerator pump, and charged particle precipitation charging the ice crystals. PMSE intensity and occurrence are hence also North-South asymmetric.

ISRs measure atmospheric winds, tidal modes, and long-period waves, as well as estimating temperatures in the D region (with some assumptions concerning ion composition). Waves are important in driving circulation that in turn influences atmospheric temperature. Simultaneous observations of the polar mesosphere and thermosphere at similar latitudes in both hemispheres will enable new understanding of the fundamental coupling between the neutral and the charged parts of the atmosphere. Long-term series from facilities in both hemispheres will show climate trends and measure the wind circulation patterns that connect the northern and southern thermospheres, revealing the impact of observed reductions in temperatures and densities in the upper atmosphere. Among the multiple long-term effects of the

reduction is the impact that a depleted reservoir of oxygen may have on the population of the magnetosphere and the processes that have been mentioned in the previous section.

(3) Measuring seasonal effects and separating them from other effects

All the processes mentioned above are influenced by or, in some cases even determined by, asymmetries between summer and winter hemispheres, transition between dayside and night side conditions, or solar cycle. For instance, the effect of differences in ion composition, temperatures, humidity, topology – and hence wind and wave patterns – in the two hemispheres over extended timescales is little understood. We must understand how asymmetries modify the fundamental properties of the A-I-M system and we must be able to separate a given effect from all others. Simultaneous measurements in both hemispheres are needed. The long, continuous data sets that will be available eventually from the multiple ISR locations under all geomagnetic conditions for decades to come will be fundamental in documenting and explaining these processes from a systems perspective.

2: Operational Capabilities

Global understanding of climate, climate change and disruption, and the impact of space weather events on technological and societally important systems requires effective and reliable modeling to enable now- and fore-casting and to allow meaningful assessment of risk, evaluate options, and plan economic, social, and mitigation strategies. Understanding and quantifying these processes is essential if global atmospheric and climate models are to have practical application in effective now- and fore-casting, in informed evaluation of climate change and climate disruption mitigation measures, and in illuminating the global debate over how, where, and when to develop and deploy such measures. The Magnetic Meridian Ring provides the essential global scale, real-time, data input essential to ensure that models and forecasts are accurate, realistic, and above all reliable.

The Magnetic Meridian Ring ISRs, like their existing prototypes, will use high bandwidth communications to monitor and control the radars and to distribute observational data routinely to scientific and service users. The software infrastructure required to disseminate data reliably and transparently will be developed from existing systems and available capacity will be well able to address the parallel needs of initiatives such as the Small instrument Distributed Ground-based Network (DASI) recommended in the Decadal Survey [2003].

3: ROM: Construction, Operation, and Maintenance

The NRC Decadal Survey [2003] recommended the construction of an Advanced Modular ‘Movable incoherent scatter radar with supporting optical and other ground-based instruments for continuous Radar measurements of [Atmosphere-Ionosphere-Magnetosphere] interactions’. Two such ISRs have been constructed (at Poker Flat, Alaska, and Resolute Bay, Canada) and are delivering high quality data; the ISR at Poker Flat has been operating essentially continuously for more than 3.5 years. A second system is under construction at Resolute Bay, with others to follow in Argentina and elsewhere.

The proposed Magnetic Meridian Ring ISRs each include two radar faces to achieve maximum magnetic latitude coverage at each site. Each face requires a support structure, 128 radar panels, a Utility Distribution Unit UDU), an Operation and Control Center (OCC), and associated power and signal cabling. All sites also require high bandwidth internet connectivity (for control, monitoring, and data retrieval) and may require generators and diesel supply and storage facilities where adequate local main power is unavailable

The average cost of each Magnetic Meridian Ring facility will be approximately \$40M including provision for supporting instrumentation including optics. Based on actual operating costs for the existing facilities, annual funding of \$1.0-1.5M will be required to support each radar face. The total cost for seven new Magnetic Meridian Ring ISRs including operation costs for ten years (at 2010 prices) is therefore approximately \$500-550M, roughly equivalent to the cost of a single medium-sized space mission. Magnetic Meridian Ring ISRs will have an expected minimum lifetime of at least 30 years.



Figure 4: Resolute Bay, Canada: The Advanced Modular Incoherent Scatter Radar provides continuous, reliable, unattended, and autonomous operation delivering high time resolution, high precision ionospheric profiles and imaging under all conditions in real time. The modular construction uses commercially fabricated units to achieve low initial capital investment, fault tolerant and efficient operations, and reduced maintenance costs. (Photo: Craig Heinselman)

Economies of Scale:

Commercial fabrication of radar elements in such large numbers over a period of several years is expected to reduce the component cost through the economies of volume production. Such economies may be further increased if international partnerships lead to the choice of the same technology to construct other Magnetic Meridian Ring sites.

While the core Magnetic Meridian Ring would provide effective coverage, international participation could add at least a further six or seven sites, improving the overall coverage and reducing the gaps between facilities.

4: Proposed concept compared with announced Criteria

- a) Is identified as a high priority or requirement in previous studies or roadmaps (for example, the earlier NRC decadal survey):
The Magnetic Meridian Ring builds on the successful implementation of recommendations from the Decadal Survey 2003, and directly addresses the goals and requirements of both the NSF GeoVision Report [2009] and the new CEDAR strategic plan [CEDAR: The New Dimension, 2010].
- b) Makes a significant contribution to more than one of the Panel themes:
The Magnetic Meridian Ring will make significant contributions to two of the three Panel themes, viz: Atmosphere-Ionosphere-Magnetosphere Interactions and Solar Wind-Magnetosphere Interactions.
- c) Contributes to important scientific questions facing solar and space physics today:
The Magnetic Meridian Ring provides important new capabilities in Space Weather studies, anthropogenic and natural effects on global climate, and powerful new resources to disentangle seasonal, latitudinal, and hemispheric effects in global change.
- d) Contributes to applications and/or policy making (operations, applications, societal benefits):
The Magnetic Meridian Ring provides the continuously available, high quality, and appropriate data that are vital to drive current and future models and forecasts and ensure appropriate societal and technical responses to the changing, and changeable, geospace environment.
- e) Complements other observational systems or programs available:
The Magnetic Meridian Ring will complement other existing collaborations such as the NSF ISR chain, the CEDAR, SuperDARN, and LIDAR (CRRL) communities, and networks of instrumentation in Canada, Europe, and Asia. In addition, the Magnetic Meridian Ring data collection and distribution infrastructure will be available to kick-start the assimilation of data from other instrumentation as envisaged by the Small instrument Distributed Ground-based Network (DASI) recommended in the Decadal Survey [2003].

- f) Is affordable (cost-benefit):
The Magnetic Meridian Ring invests resources equivalent to a medium scale satellite mission and returns continuous high value data from multiple locations appropriate to a variety of frontier science and societal goals. With an expected lifetime of at least 30 years, the Magnetic Meridian Ring is a cost effective means to achieve transformational scientific advances and deliver crucial operational capabilities.
- g) Has an appropriate degree of readiness (technical, resources, people):
The core technologies, as well as the logistic, operational, and maintenance requirements are well developed and two operational ISRs demonstrate the effectiveness of the planned approach. A continuing program of development is also envisaged, which will ensure that the Magnetic Meridian Ring remains at the forefront of scientific and technical progress throughout the lifetime of the project.
- h) Fits with other national and international plans and activities:
ISRs have never been, and can never be, ‘one PI, one task’ systems, but rather platforms open to a wide science community. There is a long history of close collaboration between the northern hemisphere ISRs, a bond made stronger in the US since the establishment of an integrated facilities office overseeing scheduling and long- term planning for all the NSF-funded systems, among others. There is also a long tradition of mutually fruitful circumpolar collaborations among scientists at the NSF ISRs, the EISCAT radars and others, such as the Kharkov, Irkutsk, and MU radars. The US and international collaborators will provide an existing user base for a southern system, while creating new collaborations and partnerships with scientists in the Southern Hemisphere. The global nature of the Magnetic Meridian Ring of ISR facilities calls for strong international collaboration, both between nations with existing infrastructure and instruments, and also to advance the technology and science enabled by the current state-of-the-art ISRs. Many scientific breakthroughs are achievable only by combining information gathered from a network of instruments.
Several pre-existing proposals of other nations, such as the Chinese Meridian Space Weather Monitoring Project (2009) and Russian plans to update and expand the ISR at Irkutsk (private communication, 2009), are directly relevant to the completion of the Meridian Ring
It would be advantageous to widen the scope of the endeavor to include other international parties for several reasons. Adopting a trans-national approach echoes the overall attitude towards, for instance, the Antarctic continent. Moreover, involving international partners broadens the research base and enhances the effectiveness of scientific and technical collaborations in addressing societally important research programs. International partners not only widen the scope of potential research in critical space science areas but also provide important opportunities and assistance to locate facilities.

References

- ‘The Sun to the Earth —and Beyond: A Decadal Research Strategy in Solar and Space Physics’, National Research Council of the National Academies, Washington DC, 2003.
- Nicolls, M. J., C. J. Heinselman, E. A. Hope, S. Ranjan, M. C. Kelley, and J. D. Kelly (2007), Imaging of Polar Mesosphere Summer Echoes with the 450 MHz Poker Flat Advanced Modular Incoherent Scatter Radar, *Geophys. Res. Lett.*, 34, L20102, doi:10.1029/2007GL031476.
- ‘The Antarctic Incoherent Scatter Radar Facility: Report on the Antarctic ISR Workshop 8-9 August 2008’, Strømme, A., I. McCrea, A. van Eyken, E. Sanchez, and J. Kelly, 2008.
- ‘The GeoVision Report’, NSF Advisory Committee for Geosciences, NSF, Washington, 2009.
- ‘Meridian Space Weather Monitoring Project’, Center for Space Science Applied Research, Chinese Academy of Sciences (<http://english.cssar.cas.cn/op/mp/>), 2009.
- Kissinger, J. and R. L. McPherron, ‘A Bright Future for Substorms’, *Eos*, 21 September 2010
- ‘CEDAR: The New Dimension’, Thayer, J., et al., in preparation , 2010.

Using space-based GNSS radio occultation data for ionospheric and space weather applications

Abstract

This white paper addresses the scientific and operational needs for continuing global observations of the ionosphere and lower neutral atmosphere with Global Navigation Satellite System (GNSS) radio occultation (RO) data. A low Earth orbit satellite carrying a GNSS receiver measures signal phase and amplitude on two L-band frequencies that can be used to profile electron density in the ionosphere, and refractivity in the lower atmosphere. These data also provide valuable monitoring of ionospheric scintillations. The six-satellite COSMIC mission has been providing valuable data since it was launched into orbit on April 14, 2006. As of November 1, 2010, COSMIC has produced more than 2.6 million Total Electron Content (TEC) data arcs and electron density profiles, or about 1500/day on average. It has also made L-band scintillation measurements from more than 3 million line-of-sight tracks. These data are having a large impact on research and operations in the atmospheric sciences because of their complete global and diurnal sampling, high precision and accuracy, and high vertical resolution of 1 km or less. In the ionosphere, COSMIC TEC and electron density profiles are providing valuable information that is leading to improved ionospheric climatology, better understanding of ionospheric physics through model verification studies, detection of phenomena that are difficult to see with other instruments, and insight in the study of ionospheric and lower atmospheric coupling. Although COSMIC is reaching the end of its mission lifetime, it has established the value of RO measurements to ionospheric research and operations. However, due to the rapid changes during ionospheric disturbances, a higher sampling rate is needed. With future transmitting constellations being developed and deployed (e.g. Galileo), a follow-on COSMIC-2 mission consisting of a 12-satellite constellation could provide more than 14,000 occultation measurements per day, which would provide enormous benefit to the ionospheric research and operational space weather communities.

COSMIC Mission Status

The Global Positioning System (GPS) radio occultation (RO) technique was successfully demonstrated on Earth with the proof-of-concept GPS/MET mission that operated from 1995-1997. This success led to the establishment of the Formosa Satellite Mission #3/Constellation Observing System for Meteorology, Ionosphere, and Climate (FORMOSAT-3/COSMIC), hereafter referred to as COSMIC, a joint U.S./Taiwan six-satellite GPS RO mission. The main goal of the COSMIC mission was to demonstrate low latency operational use of radio occultation data. The six COSMIC satellites were launched into orbit on April 14, 2006 and were precessed differentially over 17 months into their final deployment of six planes (evenly separated by 30 degrees in ascending node) to fully sample the diurnal signal of Earth's atmosphere. The primary science instrument is the GPS RO receiver. These six receivers are providing up to 2,500 daily profiles of neutral atmospheric bending angle (BA), refractivity, temperature, pressure and humidity, and up to ~3,000 daily absolute Total Electron Content (TEC) tracking

arcs and electron density profiles (EDPs) in the ionosphere, as well as ionospheric scintillation observations. Figure 1 shows the number of neutral atmospheric profiles generated per day throughout the mission. Currently, COSMIC is producing from 1,000 to 2,000 GPS RO soundings per day. All six satellites are operating and providing data, but some degradation of the spacecraft buses and GPS RO payloads have occasionally reduced the quantity of data collected. With a 5-year spacecraft design life it is reasonable to expect continued degradation of COSMIC operation past 2010, which highlights the urgency to fund and operate a COSMIC follow-on RO mission.

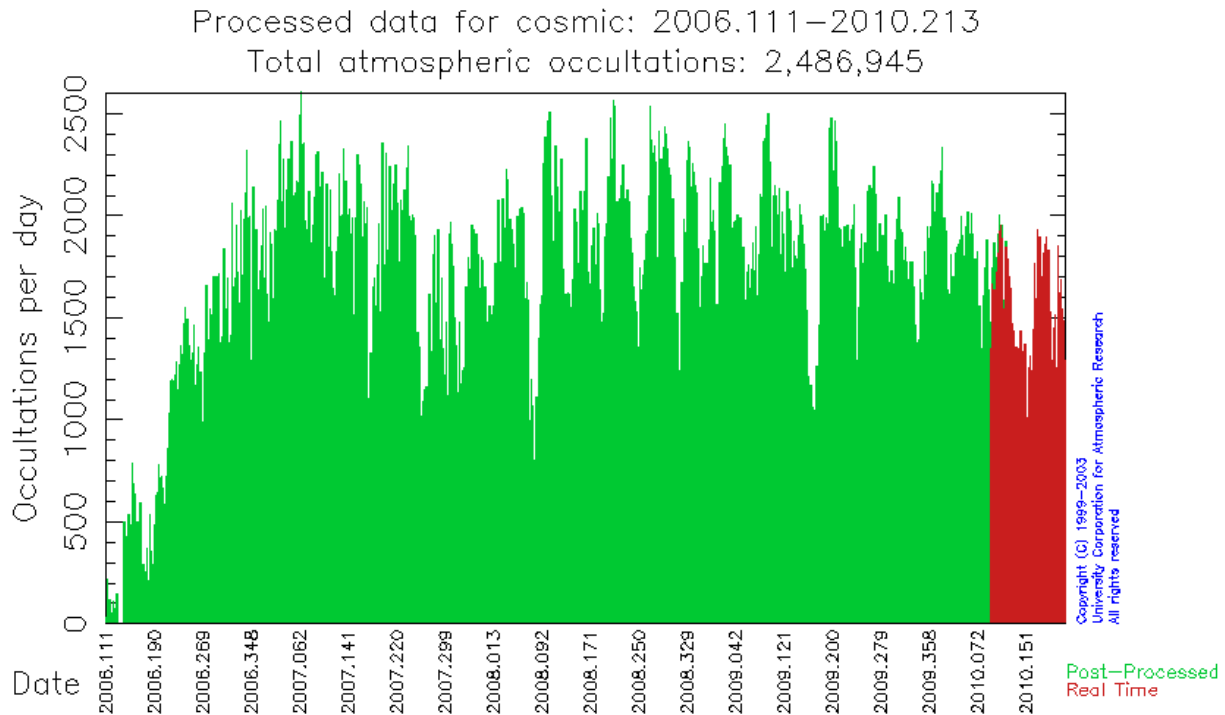


Figure 1: Daily COSMIC profile count throughout the mission. The red color shows the soundings from real-time processing. The green color shows the soundings from post-processing, which often is able to generate more sounding profiles but lags 1-2 months behind the real-time products.

Current and Future Impacts of RO on Ionospheric Research and Space Weather Applications

The COSMIC mission is generating a large number of extremely precise ionospheric observations that are useful for ionospheric research and space weather studies. The precision of the COSMIC EDPs has been quantified experimentally by comparing nearby RO soundings from different satellites (Schreiner et al., 2007). COSMIC EDPs have been used to estimate ionospheric high transition heights (O^+ and H^+ ion densities equal at this height) using COSMIC EDP data and they agreed well with C/NOFS data (Yue et al., 2010). COSMIC EDPs have also been used to investigate the coupling between the lower atmosphere and ionosphere. Yue et al (2010) investigated the ionospheric response

to the sudden stratospheric warming event of 2009. Other researchers used COSMIC TEC and EDP data recently to study the plasmaspheric response to geomagnetic storms (Pedatella, 2010), ionospheric phenomena such as mid-latitude summer nighttime anomaly (Lin et al., 2009), the Weddell Sea Anomaly (Burns et al., 2009), short-period solar cycle signals in the ionosphere (Liang et al., 2008), and ionospheric response to recurrent geomagnetic activity (which is a hot topic in recent years in solar terrestrial physics, Pedatella et al., 2009). COSMIC EDP observations are also very useful for ionospheric modeling, including improvement of existing empirical models such as International Reference Ionosphere (IRI) model (Liu et al., 2008) and constructing new empirical models (Tsai et al., 2009). Recently Komjathy et al. (2010, *J. Geophys Res.*, in press) tests the effect of COSMIC TEC data on space weather by assimilating the data into the University of Southern California (USC)/JPL Global Assimilation Ionospheric Model (GAIM). They found that assimilating COSMIC data produces higher accuracy vertical electron density information and ionospheric characteristics such as NmF2, hmF2 and TEC. This work implies that COSMIC data can have a positive effect on now-casts and short time forecasts of the ionosphere. The GPS scintillation data are also being studied to validate other RO instruments and map irregularity regions (Straus, 2009; Uma et al., 2009). Radio occultation data have previously been used to localize the irregularity regions (along the line of site) that cause the scintillation (Sokolovskiy et al., 2002). Zeng and Sokolovskiy (2010) have also recently developed a new approach for studying the sporadic E (Es) layer by GPS RO.

COSMIC data have already led to the observation of phenomena that are more difficult to see with other instruments (Pedatella, 2009; Lin et al., 2009; Burns et al., 2009). In particular there is increased interest in the study of ionospheric and lower atmospheric coupling, and RO data have unique characteristics to aid such research because they profile the full range from orbit altitude (~800 km) to Earth's surface. A clear link between weather and space weather was recently provided by Immel et al. (2006) who demonstrated an eastward propagating zonal wavenumber-3 diurnal tide (DE3) in the ionosphere that is excited by latent heat release associated with raindrop formation in the tropical troposphere. While this phenomenon was not discovered from COSMIC data, it is clearly visible in the ionospheric RO data. The upward propagation of gravity and planetary waves from the troposphere into the ionosphere and their relationship to plasma bubbles (and thus scintillation) is another area of intense interest that can benefit from the COSMIC data products. In addition, COSMIC TEC data can help study features of the plasmasphere, such as its contribution to the TEC, and plasma exchange between the ionosphere and plasmasphere during disturbed conditions.

Planned COSMIC Follow-On Mission, COSMIC-2

The National Oceanic and Atmospheric Administration (NOAA) is planning the launch and operation of an operational radio occultation mission in mid-2014. Initial funding for this follow-on to the COSMIC mission, COSMIC-2, is requested in the President's FY11 budget, but has not yet been approved. The current plan for this mission is to provide up to 14,000 soundings per day in near real-time with near uniform global and diurnal

coverage and low latency. These data will be produced with a constellation of twelve satellites; six at high inclination (72 degrees) and six at low (24 degrees) inclination.

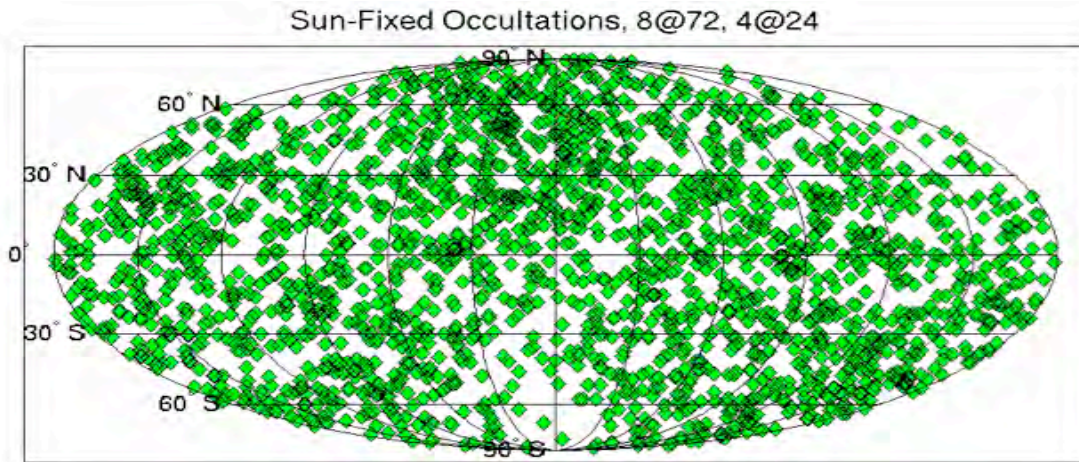


Figure 2 shows simulation results of the global distribution of soundings in a sun-fixed (i.e. local solar time) frame that are provided in a four hour period.
 Figure 2: Four hour simulation of global distribution of EDP soundings for twelve satellite COSMIC follow-on mission.

This constellation will provide slightly more soundings in the equatorial region than in the polar region. Figure 3 shows the latitudinal distribution of soundings that occur in a 500x500 km box per day for the planned constellation (in blue) and for the current COSMIC mission (in red). In addition to the EDPs, this follow-on RO mission will also provide information on absolute TEC and scintillation to all GNSS satellites in view, which will provide an enormous amount of data for ionospheric model verification, data assimilation and scintillation studies. Due to the strong impact of the COSMIC and other RO missions on neutral atmospheric, ionospheric and space weather studies, it is important for the Solar and Space Physics Decadal Survey to strongly recommend funding the planned COSMIC-2 follow-on and other future RO missions.

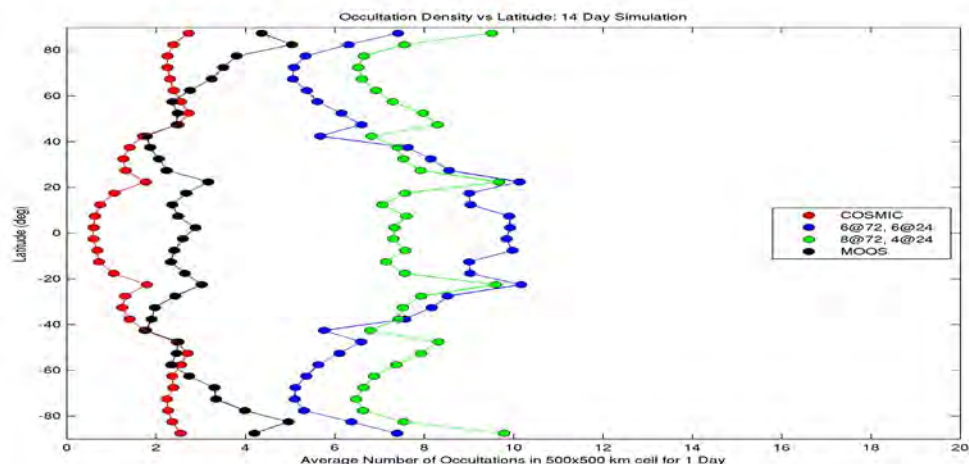


Figure 3: Simulation of latitudinal distribution of soundings (number in 500x500 km box per day) for twelve satellite COSMIC-2 follow-on.

References

- Burns, A., S. Solomon, W. Wang, A. Richmond, G. Jee, C. Lin, C. Rocken, and Y.-H. Kuo. Can the Weddell Sea anomaly and related phenomena be explained by conjugate effects? Presented at the 4th COSMIC Data Users Workshop, Oct 2009, Boulder, CO, USA, 2009.
- Uma, G., J. Y. Liu, P. S. Brahmanandam, C. H. Chen, Y. Kakinami, and C. H. Liu, A Global Observation of S4 Scintillation Index Observed by Means of FORMOSAT-3/COSMIC Presented at the 4th COSMIC Data Users Workshop, Oct 2009, Boulder, CO, USA, 2009.
- Immel, T. J., E. Sagawa, S. L. England, S. B. Henderson, M. E. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton (2006), Control of equatorial ionospheric morphology by atmospheric tides, *Geophys. Res. Lett.*, 33, L15108, doi:10.1029/2006GL026161., 2006.
- Komjathy, A., B. Wilson, X. Pi, V. Akopian, M. Dumett, B. Iijima, O. Verkhoglyadova, and A. J. Mannucci (2010), JPL/USC GAIM: On the impact of using COSMIC and ground - based GPS measurements to estimate ionospheric parameters, *J. Geophys. Res.*, 115, A02307, doi:10.1029/2009JA014420., 2010.
- Lei, J., S. Syndergaard, A. G. Burns, S. C. Solomon, W. Wang, Z. Zeng, R. G. Roble, Q. Wu, Y.-H. Kuo, J. M. Holt, S.-R. Zhang, D. L. Hysell, F. S. Rodrigues, and C. H. Lin, Comparison of COSMIC ionospheric measurements with ground-based observations and model predictions: preliminary results, *Journal of Geophysical Research*, Vol. 112, A07308, 2007.
- Liang, M.-C., K.-F. Li, R.-L. Shia, and Y. L. Yung. Short-period solar cycle signals in the ionosphere observed by FORMOSAT-3/COSMIC, *Geophys. Res. Lett.*, 35, L15818, doi:10.1029/2008GL034433, 2008.
- Lin, C., C. H. Liu, J. Y. Liu, C. H. Chen, W. Wang, and A. G. Burns. Mid-Latitude Summer Nighttime Anomaly of the Ionospheric Electron Density Observed by FORMOSAT-3/COSMIC. Presented at the 4th COSMIC Data Users Workshop, Oct 2009, Boulder, CO, USA, 2009.
- Liu, L., M. He, W. Wan, and M.-L. Zhang. Topside ionospheric scale heights retrieved from Constellation Observing System for Meteorology, Ionosphere, and Climate radio occultation measurements, *J. Geophys. Res.*, 113, A10304, doi:10.1029/2008JA013490, 2008.
- Pedatella, N. M., J. Lei, J. P. Thayer, and J. M. Forbes . Ionosphere response to recurrent geomagnetic activity: Local time dependency, *J. Geophys. Res.*, 115, A02301, doi:10.1029/2009JA014712, 2009.
- Pedatella, N. M., J. Lei, K. M. Larson, and J. M. Forbes. Observations of the ionospheric response to the 15 December 2006 geomagnetic storm: Long-duration positive storm effect, *J. Geophys. Res.*, 114, A12313, doi:10.1029/2009JA014568, 2009.
- Schreiner, W., C. Rocken, S. Sokolovskiy, S. Syndergaard, and D. Hunt. "Estimates of the precision of GPS radio occultations from the COSMIC/FORMOSAT-3 mission," *Geophys. Res. Lett.*, Vol.34, L04808. doi:10.1029/2006GL027557, 2007.

- Sokolovskiy, S., W. Schreiner, C. Rocken, and D. Hunt, "Detection of high-altitude ionospheric irregularities with GPS/MET," *Geophys. Res. Lett.*, 29(3), 3(1-4), 2002.
- Straus, P., and R. Bishop, Radio Occultation Measurements of Ionospheric Scintillation, Presented at the 4th COSMIC Data Users Workshop, Oct 2009, Boulder, CO, USA, 2009.
- Tsai, L.-C., C. H. Liu, T. Y. Hsiao, and J. Y. Huang. A near real-time phenomenological model of ionospheric electron density based on GPS radio occultation data, *Radio Sci.*, 44, RS5002, doi:10.1029/2009RS004154, 2009.
- Yue, X., W. S. Schreiner, J. Lei, C. Rocken, Y.-H. Kuo, Climatology of ionospheric high transition height derived from COSMIC satellites during the solar minimum of 2008, Presented at AGU 2009 Fall Meeting, Dec 2009.
- Yue, X., W. S. Schreiner, J. Lei, C. Rocken, D. C. Hunt, Y.-H. Kuo, and W. Wan, 2010: Global Ionospheric Response Observed by COSMIC Satellites during the January 2009 Stratospheric Sudden Warming Event. *J. Geophys. Res.*, doi:10.1029/2010JA015466, in press.
- Zeng, Z., A. Burns, W. Wang, J. Lei, S. Solomon, S. Syndergaard, L. Qian, and Y.-H. Kuo, Ionospheric annual asymmetry observed by the COSMIC radio occultation measurements and simulated by the TIEGCM, *J. Geophys. Res.*, p. A07305, vol. 113, 2008b.
- Zeng, Z., and Sokolovskiy, S.: Effect of sporadic E clouds on GPS radio occultation signals. *Geophys. Res. Lett.*, 37, doi:10.1029/2010GL044561, 2010.

Dynamical Ground-to-Space (G2S) Coupler

A Concept Paper for the NRC Space Studies Board, October 2010

David E. Siskind

Space Science Division, Naval Research Laboratory

Stephen Mende

Space Sciences Laboratory, University of California, Berkeley

Han-li Liu,

National Center for Atmospheric Research, Boulder CO

Andrew Christensen

Dixie State University, St. George Utah

1. Introduction and Motivation

One of the most exciting new paradigms in upper atmospheric science is the recognition that processes in the Heliophysical realm can be dramatically impacted by meteorological disturbances which originate in the lower atmosphere. This recognition is spurred by older work suggesting that up to a third of ionospheric variability resulted from upward propagating weather disturbances from the lower atmosphere (Forbes et al., 2000); from new observations of new modes of coupling (numerous examples are given in a recent 2010 special section of the Journal of Geophysical Research, entitled Coupling between the Upper and Lower Atmosphere) and finally, by new models (e.g. Liu et al., 2010a and b; Fuller-Rowell et al 2010) which are able to make detailed predictions of how the entire atmosphere responds to meteorological forcing from the lower atmosphere. The implication is that if realistic forecasting of space weather is ever to be achieved, these coupling mechanisms need to be understood and validated by appropriate global observations. Here we briefly describe the three atmospheric waves which are believed to couple the lower to the upper atmosphere. We then discuss programmatic implications in terms of satellite mission strategies and instrument complements which would be required for these coupling mechanisms to be understood.

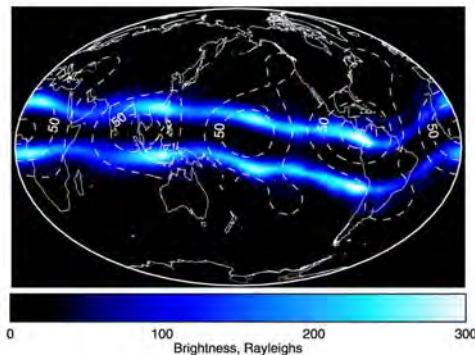
2. Modes of Dynamical Coupling

TIDES

Atmospheric solar thermal tides (as distinct from lunar/solar gravitational tides) are global-scale waves with periods that are harmonics of a solar or lunar day. Migrating tides are westward propagating waves which follow the apparent motion of the sun or moon [Forbes, 1995]. These are typically generated by the diurnal cycle of radiative heating by, for example, ozone in the stratosphere. By contrast, nonmigrating tides are linked with latent heat release from deep tropical convection which is preferentially associated with specific geographic regions [Hagan and Forbes, 2002] but which also varies diurnally.

Recently, a connection between nonmigrating tides and the ionosphere has been proposed. This link occurs in the ionospheric E region where neutral winds interact with the ionospheric plasma to set up an eastward electric field. This electric field causes plasma to drift vertically and then spread poleward, creating the so-called equatorial fountain effect and the resulting equatorial ionization anomaly (EIA). Any variations in the neutral wind from tidal oscillations will vary the E region electric field and thus affect the fountain. As shown in Figure 1 below, satellite observations of the FUV airglow clearly reveal longitudinal enhancements in the EIA which have been linked with tidal variability and variations in E region electric fields [Immel et al., 2006]. Model studies are generally supportive of this hypothesis [Hagan et al, 2007].

However, without global data on the wind and temperature variation in the E region, this linkage is circumstantial. In addition, the EIA shows much behavior and variability which hasn't been fully explained. For example, there are periods where the EIA appears "collapsed", i.e. where the



anomaly crests lie very close to the magnetic equator. These collapses show seasonal, longitudinal and even day-to-day variability. In turn, this variability is linked to the occurrence of scintillations [Basu et al., 2009]. Again, the variation of E region neutral winds and their connection to vertical $E \times B$ drifts has been emphasized as a key coupling mechanism. To fully understand this ion-neutral coupling requires neutral wind data, imagery of the airglow and data on key ionospheric variables such as ion drifts in order to track the electric field and associated F region response.

Figure 1. (From Immel et al 2006) Reconstruction of nighttime ionospheric far ultraviolet emissions from the IMAGE-FUV imager. The brightness undulations represent longitudinal variations in the equatorial ionization anomaly for 20:00 LT. The dashed circles are the amplitude of the temperature variation at 115 km due to upward-propagating tides.

Taking this a step further, if we wish to understand the variation of tides which lead to neutral wind variations, we need more data on the middle atmospheric waves induced by tropical convection and planetary waves (see next section). Recent data from ground based medium frequency (MF) radar has revealed changes in the tidal winds at the base of the thermosphere (Iimura et al., 2010) which are suggested to be linked to changes in middle atmosphere winds and heating rates. How might this effect the E region? There has been no data linking tides from their source region, through the mesosphere and into the thermosphere where the critical interaction with the plasma occurs. To address this question would require mesospheric wind data to understand the propagation from the lower atmosphere.

PLANETARY WAVES

There are two types of planetary waves which can effect the upper atmosphere: traveling waves and quasi-stationary (planetary) waves (QSPW). Traveling planetary waves are global scale oscillations associated with atmospheric eigenfrequencies. Typical wave periods are 2, 5, 10 and 16 days, with considerable spread about those values (see Lastovicka, 2006 for a review). By contrast, QSPW are forced by differences in land-sea topography and associated quasi-stationary longitudinal lower atmospheric temperature and pressure asymmetries. The latest research has revealed unexpected effects from QSPWs on the middle and upper atmosphere.

The most dramatic QSPW effects are associated with sudden stratospheric warmings (SSWs). SSWs are initiated when QSPW reach large enough amplitude to distort or break down the stratospheric circumpolar jet stream. A dramatic temperature increase can ensue. It has long been believed that SSWs have their counterparts in the mesosphere and thermosphere [Liu and Roble, 2002]. Figure 2 shows a simulated SSW from the TIMEGCM and its effects at all altitudes up to 200 km and at latitudes down to the equator. This long range coupling is known as "teleconnections" in tropospheric meteorology (<http://en.wikipedia.org/wiki/Teleconnection>); an example would be the effects of the El Nino in the Pacific on rainfall on the US East Coast. Our emerging understanding of the upper atmosphere indicates that this concept needs to be applied here. The broad morphology shown in Figure 2 has received support from additional observations [Siskind et al., 2005, 2010; Funke et al., 2010]; however, the details are often in disagreement.

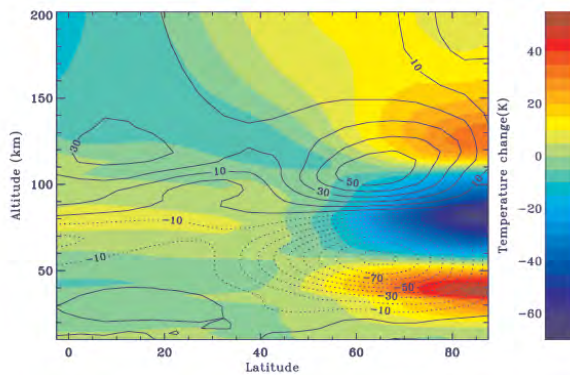


Figure 2. Difference fields of zonal mean temperature (K) (color contours) and zonal mean zonal wind (m s^{-1}) (line contours with dashed representing westward winds) due to a stratospheric warming simulated by the NCAR/Thermosphere-Ionosphere-Mesosphere-Electrodynamics General Circulation Model (TIMEGCM). From Liu and Roble (2002)

Finally, recent observations point to a link between SSWs and the equatorial ionosphere [Goncharenko et al, 2009]. Figure 3 shows that the diurnal cycle of the total electron content (TEC) was completely disrupted by the dramatic SSW of January 2009. In addition, Fuller-

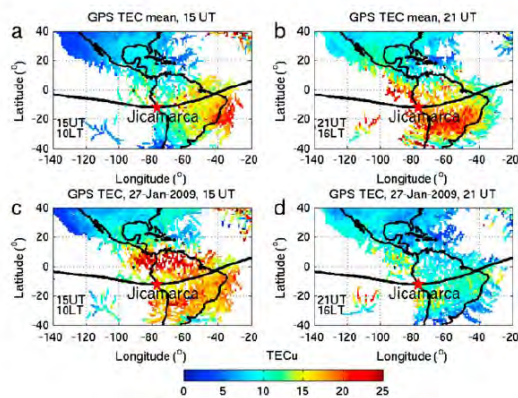


Figure 3. From Goncharenko et al [2010]. Panels (a) and (b) show the average variability of TEC for two values of UT (15 and 21). Panels (c) and (d) show dramatic deviations of TEC on January 27, 2009, the day of the dramatic warming in the NH polar stratosphere

GRAVITY WAVES

Gravity waves are atmospheric buoyancy oscillations which are typically generated in the lower atmosphere by a variety of weather disturbances such as convection, adjustment of unbalanced flow, flow over orography or shears in the jet stream. For an extensive review of gravity wave theory, observations, and parameterizations see Fritts and Alexander [2003]. These waves propagate upwards, growing exponentially, due to energy conservation and the exponential decrease in atmospheric density. At mesospheric altitudes, they can grow so large that they will break much like ocean waves on a shore. This process has long known to be critical in transporting momentum and energy from the lower atmosphere to the mesosphere. Gravity waves are responsible for closing off the mesospheric jet, for producing the cold summer mesopause (and thus the existence of thin ice clouds (e.g. Russell et al., 2010) and for driving a meridional circulation that can transport highly reactive upper atmospheric nitric oxide into the middle atmosphere [Garcia, 1987; Siskind et al., 2010]. New theoretical evidence now additionally points to a major role in the thermospheric momentum budget [Yigit et al, 2009]. Thus we must conclude that to understand the zeroth-order dynamical structure of the mesosphere and thermosphere, gravity wave drag must be precisely specified. Gravity waves have also been linked with ionospheric phenomena such as spread-F plasma instabilities [Fritts et al., 2008; Abdu et al., 2009] that can disrupt global communications as well as with travelling ionospheric

Rowell et al [2010] report a 100% increase in the total electron content (TEC) over the US on the day of the warming. Why this should be so is unclear; however, like the EIA, suspicion falls on changes in the diurnal variation of E region winds and resultant ion-neutral interactions [Liu et al., 2010; Pedatella and Forbes, 2010]. The Liu et al [2010] model makes specific predictions for the temporal and longitudinal variability of the neutral winds and ion drifts for which the requisite global data are as yet unavailable.

disturbances (Shiokawa et al., 2007), generated by either auroral heating and or from convectively generated waves at low latitudes (Vadas and Liu, 2009) One key process governing how gravity waves propagate is filtering of the wave spectrum by the background zonal and meridional wind flow. To quantify the effects of GWs in all these small and large scale atmospheric perturbations, global scale measurements of GWs and their consequences are required.

3. Science Questions

The result of these new observations and models are to re-emphasize old questions that have never been adequately addressed by a proper suite of measurements as well as animate new ones. The fundamental questions we address are:

What controls the basic state of the mesosphere, thermosphere and by implication, the ionosphere?

How does atmospheric forcing from below affect thermospheric and ionospheric variability, commonly known as space weather?

What is the relative importance of coupling due to large scale waves, wind-driven electrodynamics, composition, and temperature changes in the neutral atmosphere?

What are the sources of the waves (both large and small scale) which govern the structure and variability of the neutral atmosphere and ionosphere, how are they filtered by the mean wind environment of the middle atmosphere and lower thermosphere and what is their impact on the mesosphere, thermosphere and ionosphere?

These questions must be answered before we can hope to reliably provide forecasts of space weather effects for society at large.

3. Programmatic Implications

Given the diverse array of science questions outlined in this paper, it is unlikely that the science questions posed here could be answered by a single satellite. Indeed, to understand high latitude teleconnections on the tropical upper atmosphere, *coordinated* measurements from two satellites are likely to be needed. A satellite in high inclination (or polar) orbit is required to for observing waves produced by the aurora. Also the latitude range 40-70° is where gravity wave drag drives the mesospheric circulation [cf. Richter et al., 2008, Figure 7]. However, the low latitude measurements require dense local time coverage to adequately sample tidal variability. Detailed observations of gravity waves in the tropical region propagating from convective sources are necessary to study the triggering of ionospheric instabilities. These measurements call for a low inclination satellite.

Even for a satellite such as TIMED, which is in an inclined orbit and which does see up to 80° in each hemisphere for ½ the year, Zhang et al [2006] have shown that the local time precession is too slow to resolve temporal changes in the tropics on timescales less than 2 months. Given that the ion drift changes displayed in the models of Liu et al [2010a] and Fuller-Rowell et al [2010] occurred on timescales of less than 1 week, it is apparent that a satellite with a lower orbital inclination and more rapid local time precession is required to resolve these effects. Since this low inclination would preclude observing the temperature response displayed in Figure 3, the need for a second satellite is implied.

Other implications for measurements are

1. To understand neutral-ion coupling in the tropical dynamo region, coordinated observations are required of both neutral winds and ion drifts with sufficient diurnal, latitudinal and longitudinal resolution to see tidal changes.
2. For gravity wave effects on the mesosphere and thermosphere, there is a need to perform measurements over a wide range of spatial scales and geometries (Preusse et al., 2008). Waves in limb imagery can be observed by fluctuations in the temperature profile [Eckermann and Preusse, 1999]. Nadir infrared imaging can observe gravity waves propagating out of the stratosphere [Dewan et al., 1998]. Preusse et al., (2008) also argue that sub-limb imagery is necessary to completely categorize the spectrum. These IR observations should be combined with high spatial resolution imagery of waves in the upper mesospheric and lower thermospheric airglow [cf. Fritts et al, 2006, figure 5]. The combination of these two measurements, IR imaging and airglow imaging, would quantify the flux of gravity waves entering the mesosphere and thermosphere.
3. To track the rapid growth of QSPW and associated tidal responses, there is a need to measure meteorological state variables such as temperature and winds at both low and high latitudes with sufficient time resolution to quantify these changes.
4. There is a need to measure over a sufficient altitude range to be able to quantify the growth, filtering and ultimate breaking of waves from the lower atmosphere. This implies, for example, wind measurements from the stratopause to the upper thermosphere.
5. Finally, it is necessary to distinguish between lower atmospheric forcing and solar-terrestrial forcing. Thermospheric composition measurements such as the O/N₂ ratio have been shown to be a useful diagnostic in this regard (Stephan et al., 2008; Meier et al., 2005).

4. Possible Measurement Complements

The technology to perform the above measurements mostly exists with high heritage. Sample instrumentation might include (but need not be limited to) the following:

To measure winds, the two techniques which have been used are Fabry-Perot interferometry [Hays et al., 1993] and Stepped-Michelsen Interferometry [Shepherd et al, 1993]. A newer approach, Doppler Asymmetric Spatial Heterodyne Spectroscopy (DASH) has been proposed which combines elements of a Stepped-Michelsen Interferometer with a Spatial Heterodyne Spectrometer (Englert et al. 2010). By observing mesospheric emissions in the mesosphere and atomic emissions in the thermosphere, wind data with a precision of ~ 5 m/s can be achieved.

Temperature can be measured by the rotational temperatures (line ratios) of selected molecular emission features including the O₂(¹Δ) and O₂(¹Σ) and OH(v*) emissions. With the same interferometric techniques used to measure neutral winds, temperatures can be inferred from the widths of atomic lines such as from O(¹S) or O(¹D) that are dominated by temperature broadening. Estimated precisions of 5K or less are indicated.

As noted above, gravity waves can be imaged in the IR for the middle atmosphere or via OH and O₂ airglow imagery for the upper mesosphere.

A single channel FUV nadir high sensitivity imager which measures the 135.6 nm nightglow can provide information on the nighttime ionosphere and its response to tides and gravity waves [Coker et al., 2008]. A multi-channel FUV imager which additionally measures the N₂ Lyman-Birge Hopfield bands can be used to infer neutral composition, specifically the ratio of atomic oxygen to molecular nitrogen [Christensen et al., 2003; Stephan et al., 2008]

Ionospheric responses to the neutral wind and wave dynamics could be measured by a suite of instruments similar to those flown on the recently successful C/NOFS satellite [de la Beaujardiere et al., 2004; see also the special section in *Geophys Res Lett.*, 2009]. These instruments could include an in situ ion velocity drift meter which measures ion drift, temperature and composition [Heelis and Hanson, 1998], in situ plasma detectors to measure ion and electron densities and booms to measure the electric field. The combination of these in situ devices with

the remote sensing instrumentation can finally quantify the role of neutral wind variations on the ionospheric dynamo electric field. Note that the C/NOFS mission does not have remote sensing data on the E region winds and neutral atmosphere composition. Additionally, information on the small scale waves entering the thermosphere are required to understand how ionospheric instabilities such as ESF might be triggered.

5. Conclusion

Older versions of the Heliophysics and Sun-Earth Connections Roadmaps discussed concepts such as the Tropical ITM Coupler and the ITM Waves Coupler which are similar in many ways to that presented here (see http://sec.gsfc.nasa.gov/roadmap_appendixD.pdf); these have never been implemented. If we assume two satellites, each with approximately 125-150 kg of instrumentation, the total costs, including launch, would likely put this G2S Coupler concept in the “medium” cost bin (\$500-700M). Originally, the TIMED mission was to have been two satellites, operating in both high and low inclinations; however, this concept was dropped early on. More targeted proposals addressing a subset of the above science questions have also been proposed; even these were not accepted. These long-dormant, but still exceedingly and in fact, increasingly relevant concepts need to be resurrected and their associated questions answered before we can have a true predictive capability of space weather.

References

- Abdu, M. A., (2009) Gravity wave initiation of equatorial spread F/plasma bubble irregularities based on observational data from the SpreadFEx campaign, *Ann. Geophys.*, 27, 2607–2622,
- Basu, S. J., Day-to-day variability of the equatorial ionization anomaly and scintillations at dusk observed by GUVI and modeling by SAMI3, *JGR.*, 114, A04302, doi:10.1029/2008JA013899, 2009.
- Christensen, A. B. et al., (2003), Initial observations with the global Ultraviolet Imager (GUVI) in the NASA/TIMED satellite mission, *J. Geophys. Res.* 108(A12), 1451, doi:1029/2003JA009918.
- Coker C. et al (2009), Observations of the ionosphere using the Tiny Ionospheric Photometer, *Terr. Atmos. Ocean. Sci.*, 20, 227-235.
- Coy, L., et al (2009), Planetary wave breaking and tropospheric forcing as seen in the stratospheric sudden warming of 2006, *J. Atm. Sci.*, 66, 495-507.
- De la Beaujardiere (2004), C/NOFS: a mission to forecast scintillations, *J. Atm. and Solar-Terrestrial Physics*, 66, 1573-1591.
- Dewan, E. M. et al, (1998) MSX satellite observations of thunderstorm-generated gravity waves in mid-wave infrared images of the upper atmosphere, *Geophys. Res. Lett.*, 25, 2809-2812.
- Eckermann, S. D. and P. Preusse (1999), Global measurements of stratospheric mountain waves from space, *Science*, vol 286, 1534-1537, 1999.
- England, S. E. et al. (2006), Longitudinal variation of the ionosphere, *Geophys. Res. Lett.*, 121105, doi:10.1029/2006GL027465.
- Englert et al., (2010). Spatial Heterodyne Spectroscopy: An emerging optical technique for Heliophysics and Beyond, Concept paper submitted to NRC Space Studies Board
- Forbes, J. M. (1995), Tidal and Planetary Waves, In, *The Upper Mesosphere and lower thermosphere: A review of experiment and theory*, AGU Geophysical Monograph 87, 67-87.
- Fritts D. C. and M. J. Alexander (2003), Gravity wave dynamics and effects in the middle atmosphere, *Rev. Geophys.* Vol 41, 3-64.
- Fritts D. C. et al., (2006) Mean and variable forcing of the middle atmosphere by gravity waves, *J. Atm and solar-Terr physics*, 68, 247-265.
- Fritts, D. C., et al. (2008), Gravity wave and tidal influences on equatorial spread F based on observations during the Spread F Experiment (SpreadFEx), *Ann. Geophys.*, 26, 3235–3252, 2008,
- Fuller-Rowell, T., J. et al., (2010), A whole atmosphere model simulation of the impact of a sudden stratospheric warming on thermospheric dynamics and electrodynamics, *J. Geophys. Res.*, A00G08, doi:10.1029/2010JA015524

Funke et al, (2010) Evidence for dynamical coupling from the lower atmosphere to the thermosphere during a major stratospheric warming, *Geophys. Res. Lett.*, doi:10.1029/GL043619.

Garcia, R. R. (1987) On the mean meridional circulation of the middle atmosphere, *J. Atmos. Sci.*, 44, 3599.

Goncharenko, L. P., J. L. Chau, H.-L. Liu, and A. J. Coster (2010), Unexpected connections between the stratosphere and ionosphere, *Geophys Res. Lett.*, 37, doi:10.1029/2010GL043125.

Hagan, M. E. and J. M. Forbes, (2002), Migrating and nonmigrating diurnal tides in the middle and upper atmosphere excited by tropospheric latent heat release (2002), *J. Geophys. Res.*, 107, 4751, doi:10.1029/2001JD001236.

Hagan, M. E., A. Maute, and R. G. Roble (2009), Tropospheric tidal effects on the middle and upper atmosphere, *J. Geophys. Res.*, 114, A01302, doi:10.1029/2008JA013637.

Heelis, R. A. and W. B. Hanson (1998), Measurements of thermal ion drift velocity and temperature using planar sensors, in *Measurement Techniques in Space Plasmas: Particles*, Geophys. Monogr. Ser., vol. 102.

Hays, P. B. et al., (1993) The High-Resolution Doppler Imager on the Upper Atmosphere Research Satellite, *J. Geophys. Res.*, 10, 713-10,724.

Immel, T. J. et al., (2006), Control of equatorial ionospheric morphology by atmospheric tides, *Geophys. Res. Lett.*, L15108, doi:10.1029/2006GL026161.]

Lastovicka (2006), Forcing of the ionosphere by waves from below, *J. Atm. Solar-Terr Phys.*, 479-497.

Liu, H-L, et al (2010a), Ionospheric variability due to planetary waves and tides for solar minimum conditions, *J. Geophys. Res.*, 115, doi:10.1029/2009JA015188.

Liu, H.-L, et al (2010b) Thermospheric Extension of the Whole Atmosphere Community Climate Model, *J. Geophys. Res.*, in press.

Liu, H-L and R. G. Roble (2002), A study of a self-generated stratospheric sudden warming and its mesospheric lower thermospheric impacts using the coupled TIME-GCM/CCM3, *J. Geophys. Res.*, 107(D23), 4695, doi:10.1029/2001JD001533.

Meier, R. R., et al (2005), First look at the November 2003 superstorm with TIMED/GUVI: Comparisons with a thermospheric global circulation model, *J. Geophys. Res.*, A089S41, doi:1029/2004JA010990.

Pedatella, N.M. and J. M. Forbes (2010), Evidence for stratosphere sudden warming-ionosphere coupling due to vertically propagating tides, *Geophys. Res. Lett.*, 37, doi:10.1029/2010GL043560.

Preusse, P. et al., (2008), Transparency of the atmosphere to short horizontal wavelength gravity waves, *J. Geophys. Res.*, 113, D24104, doi:10.1029/2007JD009682.

- Richter, J. H et al.(2008), Dynamics of the middle atmosphere as simulated by the Whole Atmosphere Community Climate Model, (WACCM3), *JGR.*, 113, doi:10.1029/2007JD009269.
- Russell, J. M. III et al, (2009) The Aeronomy of Ice in the Mesosphere (AIM) mission, *J. Atm Solar Terr Phys*, 71, 289-299.
- Shepherd, G. G. et al., (1993), WINDII, the Wind Imaging Interferometer on the Upper Atmosphere Research Satellite,, *J. Geophys. Res.*, 10,725-10,750.
- Shiokawa, K. et al., Ground observation and AMIE-TIEGCM modeling of a storm time traveling ionospheric disturbance, *J. Geophys. Res.*, doi:10.1029/JA0111772, 2007.
- Siskind, D. E. et al. (2005), Observations of stratospheric warmings and mesospheric coolings by the TIMED/SABER instrument, *Geophys. Res. Lett.*, 32, doi:10.1029/2005GL02399.
- Siskind, D. E. et al (2010), Case studies of the mesospheric response to recent minor, major and extended stratospheric warmings, *J. Geophys. Res.*, D00N03, doi:10.1029/2010JD014114.
- Stephan, A. W. et al (2008), Comparison of Global Ultraviolet Imager limb and disk observations of O/N₂ during a geomagnetic storm, *J. Geophys. Res.*, 113, A01310, doi:1029/2007JA012599.
- Vadas, S. and H.-L. Liu, (2009), Generation of large-scale gravity waves and neutral winds in the thermosphere from the dissipation of convectively generated gravity waves, *J. Geophys. Res.*, 114, A10310, doi:10.1029/2009JA014108.
- Yigit, E., et al (2009), Modeling the effects of gravity wave momentum deposition on the general circulation above the turbopause, *J. Geophys. Res.*, 114, doi:10.1029/2008JD011132.
- Zhang, X., et al., (2006) Monthly tidal temperatures 20-120 km from TIMED/SABER, *J. Geophys. Res.*, 111, A10S08, doi:10.1029/2005JA011504.

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.

Gravity Wave Coupling Processes Between the Lower Atmosphere and the Mesosphere and Lower Thermosphere.

Steven M. Smith, Boston University

Summary

In order to forecast space weather phenomena we need to understand fundamentally the dynamics (e.g., energy and momentum flux budgets, chemistry, etc) of the earth's atmosphere as a whole system. In this context, one important and poorly-understood aspect is the precise role that gravity waves, originating from the lower atmosphere, play in the transport of energy and momentum up into the earth's upper atmosphere.

Introduction

Gravity waves, particularly short-period gravity waves ($T < 1$ hour) contribute a significant proportion of momentum and energy to the mesosphere and lower thermosphere (MLT) in the 80 – 110 km altitude region. These waves originate from the lower atmosphere and propagate upwards into the MLT where they deposit their momentum into the local mean flow. The resulting gravity wave deposition is realized on a global scale; it reverses the zonal MLT wind flow and it drives a strong mean meridional flow pattern that creates the coldest region on Earth – the summer mesopause.

Compared to other regions in the Earth's atmosphere, the neutral MLT region is sparsely observed and yet it is an important transition zone between the well-mixed collisionally-dominated lower atmosphere and the electrodynamically controlled region above. Above about 100 km, diffusive separation between gas species becomes dominant. The restoring buoyancy oscillations associated with the waves becomes less viable so gravity waves tend to break and be absorbed at this region.

Gravity wave propagation is governed solely by the local temperature and winds fields. The ubiquitous presence of temperature inversions and wind-shears in the stratosphere and mesosphere also provide conditions conducive to ducted propagation of the gravity waves and, as a result, the waves may propagate great distances with little or no attenuation. Factors such as these have large important implications for understanding the global transport of momentum and energy from lower levels.

Although the broad outline of how gravity waves transport energy and momentum up into the MLT is understood, what is required is the precise quantification of the amount these parameters and order to quantify how much energy is transported into the region.

Research Strategy

Quantitative measurements of MLT gravity wave coupling process require the deployment of ground-based multi-instrument clusters distributed in chains over large spatial and geographical areas.

The MLT region has several neutral airglow layers at 85–95 km which can be used as tracers of the dynamics and hence the gravity wave parameters. The required instrumentation is already available. All-sky imaging systems can provide observations of MLT gravity waves over large spatial regions and yield several important gravity wave parameters. Wind and temperature profiles can be obtained from metal resonance lidar and FPI observations. Gravity wave parameters from other altitudes can be obtained from meteor, ISR and MF radar measurements, which also will provide altitude information. The wind information will allow intrinsic wave parameters and vertical scale sizes to be determined, and importantly momentum and energy flux estimates. Momentum divergence over the altitude region will provide quantitative estimates of the deposition rates into the region. The results will provide long-sought improved parameterization schemes for global scale circulation models for the upper atmosphere.

Gravity waves originate from many sources in the lower atmosphere orography, jet streams, convection centers, frontal systems. Hence, the location of instrumentation clusters at different geographical locations will allow comparisons between gravity wave populations of differing origins. For example, orographic gravity waves near mountain ranges compared to those originating from convective systems.

Gravity waves have been implicated in the origin of the ionospheric disturbance spread-F, which causes havoc with global communications systems. Understanding the precise role of MLT gravity waves in this is a crucial societal need.

Neutral Temperature and Wind Measurements Earth's Thermosphere.

Steven M. Smith, Boston University

Summary

There is a dire need for accurate neutral temperature and wind climatologies of the earth's lower thermosphere in the altitude range 100 – 300 km. Current models such as the Mass Spectrometer Incoherent Scatter Model (MSIS), the Horizontal Wind Model (HWM) model only capture the basic features of the temperature and wind fields. In order to characterize the thermosphere and its associated phenomena such as tides, MSTIDS, spread-F, the basic state of the region must be known.

Accurate knowledge of the thermosphere and its processes is crucial to understanding the propagation of space weather processes from above and dynamical processes from below, such as gravity waves and tides.

Fabry-Perot interferometer measurements of airglow at 85-95 km and 250 km can yield neutral temperatures and winds. Metal resonance lidars and meteor radars can also provide similar measurements up to about 110–115 km.

Strategy

The required technology is readily available. Neutral winds and temperatures could be obtained routinely from a space-based system performing limb scanning operations of various airglow emissions in the UV and near-IR. Several emissions and techniques have been used previously. For example, RAIDS observations of the O₂(0-0) Atmospheric band near 765.0 nm can yield temperatures from 80–200 km altitude. A Fabry-Perot or Michelson interferometer could yield Doppler winds the same emission.

The instrumentation could be installed on a dedicated atmospheric research satellite or fly on a series long-duration sub-orbital rocket flights.

The Importance of Solar EUV and FUV Irradiance Measurements for Space Weather and Atmospheric Modeling

Submitted to: A Decadal Survey Panel on Solar and Heliospheric Physics

Contributors: M. Snow, W. E. McClintock, & F. G. Eparvier
University of Colorado/LASP

Executive Summary:

Extending the data record of solar EUV and FUV irradiances should continue to be a priority of the heliophysics community. Solar spectral irradiance in these wavelength bands is the dominant source of energy input to the upper atmosphere. Understanding both the long-term (climate) and short-term (space weather) variations in the atmosphere is critical for our technological society. NASA does not currently have any FUV irradiance measurements in the pipeline after the SORCE mission. An irradiance monitor could be manifested on a commercial satellite for a modest investment by NASA.

Introduction:

Ultraviolet solar irradiance is the primary source of energy input for the upper atmosphere (e. g. Hinteregger, [Fukui](#), & [Gilson](#), 1981). In this context, we divide the ultraviolet spectrum into the extreme ultraviolet (EUV) ranging from 10 to 120 nm, and the far ultraviolet (FUV) ranging from 120- to 200 nm. The strong Hydrogen Lyman alpha line at 121.6 nm is a natural boundary between these two regimes. Figure 1 shows the altitude of unit optical depth for these wavelengths. The solar spectral irradiance (SSI) at these wavelengths shows significant variability at all timescales that we are able to measure, all the way from very transient responses to flares through variation over decades from solar minimum to solar minimum (Snow, McClintock, & Woods, 2010). Variability on the short time scales is an important driver for space weather (Lathuillere, et al., 2002). The EUV irradiance variations can drive fluctuations in the density of the upper atmosphere (Lean, Picone, & Emmert, 2009) which affect satellite drag, global positioning, communications, navigation, etc. In fact, forecasting the solar drivers of these upper atmospheric changes was listed as a high priority in the National Space Weather Strategic Plan (1995).

Currently, there are instruments measuring SSI in this wavelength range, and the data record in the FUV extends back almost two solar cycles. There is no continuous EUV data record before 2002 when TIMED SEE began science operations. There has been a tremendous amount of work using proxies to reconstruct the ultraviolet SSI record back to 1947 or earlier for use in climate research (e. g. Krivova, Solanki, & Unruh, 2009a; Krivova et al., 2009b), as well as using these proxies to nowcast and forecast the irradiance (Lean, Picone, & Emmert, 2009).

In the near future, current EUV and FUV missions will reach the end of their lives, and there is a lack of follow-on missions in the pipeline. NOAA's next generation of EUVS will not measure the full spectrum. It will only observe certain wavelength bands and

construct the rest of the spectrum. Our only estimate of the EUV and FUV irradiance will come from proxy models unless new ultraviolet SSI missions are defined and implemented.

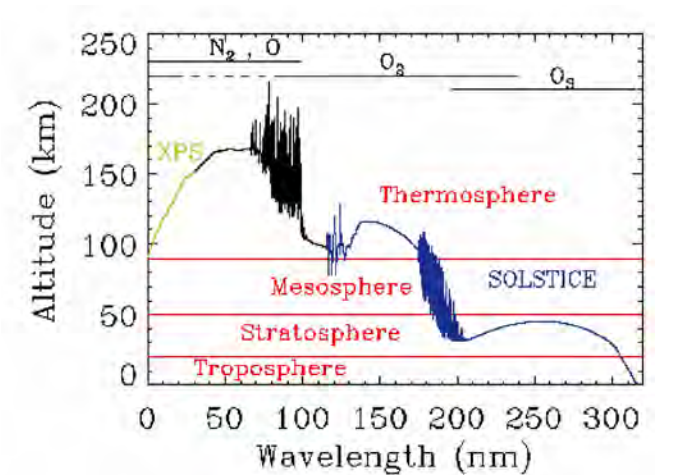


Figure 1. Altitude of optical depth unity for incoming solar spectral irradiance as a function of wavelength. The atmospheric species which are the dominant absorbers are indicated along the top. Adapted from Meier (1991).

Status of current missions and data records

We are currently in the fortunate situation of having a reliable source of ultraviolet SSI measurements available to the research community. Table 1 lists the current NASA missions measuring in this wavelength range. Additionally, there are NOAA EUV measurements from the GOES EUVS series (in selected broad bands), as well as two Russian and European instruments on the International Space Station: Russian Solar Patrol, 0.14-135 nm (Avakyan, 2006) and SOL-ACES, 16-140 nm (Schmidke et al., 2006). SOL-ACES was originally intended to observe the entire FUV range, but the flight unit can only measure up to 140 nm. The LYRA instrument on ESA's PROBA2 also measures Lyman alpha and two broad EUV bands (Hochedez et al., 2006).

Spacecraft/Instrument	Wavelength Range (nm)	Mission time-Time periodPeriod
SOHO/SEM	0-50 and 26-34	1995-present (extended mission)
TIMED/SEE	27-194	2002-2011
SDO/EVE	0.1-105 (and 121.6)	2010-present (5 year primary mission)
SORCE/SOLSTICE	115-300	2003-present (in 2 nd extended mission)

Table 1. NASA missions observing SSI in the EUV and FUV wavelength range.

There is also a wealth of historical data records of SSI in this wavelength range, primarily from the two instruments on the Upper Atmosphere Research Satellite (UARS) which

Formatted: Font: (Default) Arial, 10 pt

Formatted: Font: (Default) Arial, 10 pt

Formatted Table

Formatted: Font: (Default) Arial, 9 pt

Formatted: Font: (Default) Arial, 9 pt

Formatted: Font: (Default) Arial, 9 pt

Formatted: Font: (Default) Arial, 9 pt

Formatted: Font: (Default) Arial, 9 pt

Formatted: Font: (Default) Arial, 9 pt

operated from 1991 to 2005: SUSIM (Brueckner et al., 1993) and SOLSTICE (Rottman et al., 1993). Additionally, the Solar Mesosphere Explorer (SME) (Rottman, 1988) provided FUV measurements from 1981 to 1989. There are also EUV measurements from AE-E, 14-185 nm (Hinteregger et al., Fukui, & Gilson, 1981) and SNOE, 2-20 nm (Bailey et al., 2000).

DeLand & Cebula (2008) have worked on merging FUV datasets into a consistent composite, and their results spanning the time period of 1978 to 2005 are ~~for~~ available for download on the LASP Interactive Solar IRradiance Datacenter (LISIRD) site <http://lasp.colorado.edu/lisird>. Note that all their FUV results from 1978-1982 and 1989-1992 are synthetic values derived from the 3-component model of Woods et al. (2000).

It is important to note that the proxy model extensions are based on measurements of ultraviolet SSI that extend for less than three solar cycles. There are many indications that this most recent solar minimum was unusually low with respect to irradiance (Snow, McClintock, & Woods, 2010), solar wind (Manoharan, 2010), solar magnetic field (Mewaldt et al., 2010), etc. Extrapolating the long-term behavior of the Sun from such a small sample size of solar cycles will inevitably produce large uncertainties in reconstructed irradiances and could lead to erroneous estimates of the influence of solar irradiance on the atmosphere.

Empirical proxy models and reconstructions

The datasets described in the section above have formed the basis for both theoretical and empirical models of SSI. Domingo et al. (2009) reviewed the various empirical irradiance models including the SATIRE model of Krivova, Solanki, & Unruh (2009a). Figure 2 shows the comparison of the model results to the observations of SUSIM (green) and to a composite data record from Woods et al. (2000). The Woods composite uses data from UARS SOLSTICE, SME, AE-E, and a proxy model based on the Mg II index. For times prior to 1978, it is based on an empirical model derived from the F10.7 index. The SATIRE model seems to produce the same irradiance level at each solar minimum, possibly masking any secular trend in the Lyman alpha irradiance over many solar cycles.

Space Weather inputs

The EUV and FUV wavelength ranges are entirely absorbed in the upper atmosphere (e.g. Haigh, 2007). Ultraviolet SSI controls the production and destruction of ozone (e.g. Haigh, 1999, 2007). Solar EUV photons from flares can disrupt HF communications and GPS signals. Variations in the neutral atmosphere density from ultraviolet absorption in the upper atmosphere cause variations in spacecraft drag (Bowman, Tobiska, & Marcos, et al., 2006; Bowman et al., 2007). Solar variability near 155 nm causes variation in the molecular oxygen density in the lower thermosphere through absorption in the Schumann-Runge continuum.

Forecasting these changes in the atmosphere require near real-time knowledge of the ultraviolet SSI. Tobiska, Bouwer, & Bowman (2008) have derived a set of solar

irradiance indices to simplify the input to operational space weather models. Their new Solar Irradiance Platform (SIP) (Tobiska, Bouwer, & Bowman, 2008) is an improvement over their previous model (SOLAR2000), and is used by the Air Force and other operational customers for use in satellite tracking. Ten of the fifteen solar inputs to their model come from EUV and FUV measurements (Tobiska, Bouwer & Bowman, 2008).

Guo et al. (2007) used the accelerometer data on CHAMP and determined that the thermospheric density can be best modeled using several irradiance indices: F10.7 plus the EUV and FUV indices (S_{EUV} and S_{FUV}) from Tobiska, Bouwer, & Bowman (2008). In their estimation, the EUV contributed 40% of the variation in thermospheric density, while the FUV contributed another 28% (Guo et al., 2007). So both wavelength bands must be available for the best estimation of satellite drag coefficients.

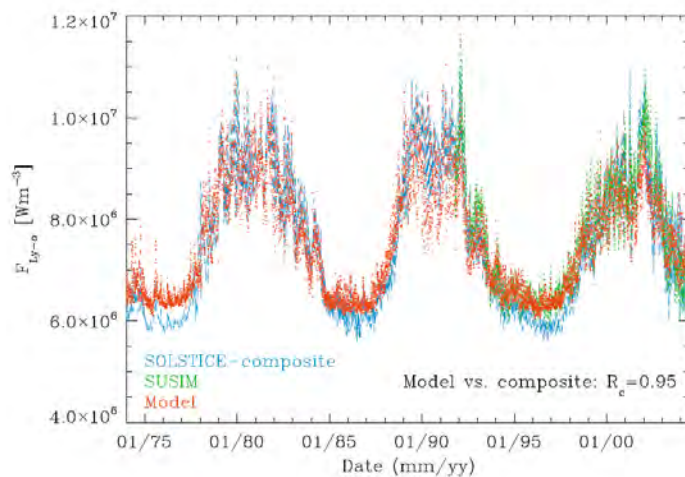


Figure 2. (Figure 6 of Krivova et al. (2009)). Solar Lyman alpha since 1974 reconstructed by the SATIRE-S model. The model captures the overall variability well, with a correlation coefficient of 0.95. However, it does poorly at reproducing the irradiance at solar minimum.

Formatted: Font: (Default) Arial, 10 pt

Use in Atmospheric Models

Solar ultraviolet irradiance's influence on climate is still somewhat difficult to model in detail. There is a clear 11-year signal in temperature and geopotential height (Langematz, Matthes, & Grenfell, 2005). Stratospheric warming and dynamics can propagate down to the troposphere (Kodera and Kuroda, 2002) where they could potentially impact climate. Another mechanism for interaction between the stratosphere and the troposphere is the quasi-biennial oscillation. Labitzke (2003) demonstrated that there is a correlation between the stability of the polar vortex and the solar cycle, and Hood (2004) has shown that there is a response of tropospheric temperatures to changes

in the ultraviolet SSI on rotational timescales. Having a consistent long-term dataset would be essential for understanding these mechanisms of climate change.

Mission Concepts

An EUV or FUV irradiance monitor need not be an expensive stand-alone mission. Most of the prior sensors in this wavelength range have been co-manifested with many other instruments. A logical host for a simple SSI instrument is a commercial satellite (<http://www.space.commerce.gov/general/commercialpurchase/hostedpayloads.shtml>). Based on our experience at LASP with the Global-scale Observations of the Limb and Disk (GOLD) mission and with the EUV sensor on the Mars Atmosphere and Volatile Evolution (MAVEN) mission, we estimate that a simple SSI monitor could be built and operated for a decade for less than \$30 million. This number includes \$10 million for costs of accommodation on the spacecraft, with the rest going towards design and fabrication of the instrument plus 10 years of operations. Besides the relatively low cost for a payload hosted on a commercial satellite, the faster time to move from design to orbit could quickly address any gaps in the FUV data record.

Conclusion:

Extending the data record of solar EUV and FUV irradiances should continue to be a priority of the heliophysics community. SSI in these wavelength bands is the dominant source of energy input to the upper atmosphere. Understanding both the long-term (climate) and short-term (space weather) variations in the atmosphere is critical for our technological society. The National Space Weather Strategic Plan highlights understanding the solar input to our space environment as a high priority. Measurements at these wavelengths can be done on a modest budget, yet yield a great value to the scientific community. It should also be noted that the NOAA plan for operational monitoring of the EUV is to measure only limited wavelength bands and then construct the rest of the spectrum through a model. The new insights already gained from less than a year of SDO/EVE data show that we still have much to learn from observing the full spectrum.

References:

[Avakyan, S.](#) (2006), Adv. Sp. Res., 37, 297.

Bailey, S. M. et al. (2000), JGR, 105, 27179.

Bowman, B. R., Tobiska, W. K., & Marcos, F. A. (2006), AIAA, 2006-6166.

Bowman, B. R. et al. (2007), JASTP, 70, 774.

Brueckner, G. E. et al. (1993), JGR, 98, 10695.

DeLand, M. T., & Cebula, R. P. (2008), JGR, 113, A11103.

Domingo, V. et al. (2009), Space Sci. Rev., 145, 337.

[Guo, J. et al. \(2007\), JGR, 112, A10308.](#)

Haigh, J. D. (1999), JASTP, 61, 63.

~~Guo, J. et al. (2007), JGR, 112, A10308.~~

Haigh, J. D. (2007), Liv. Rev. Sol. Phys.
<<http://solarphysics.livingreviews.org/Articles/lrsp-2007-2/>>

Hinteregger, H. E., Fukui, K., & Gilson, G. R. (1981), GRL, 8, 1147.

Hochedez, J.-F. et al. (2006), Adv. Sp. Res., 37, 303.

Hood, L. L. (2004), GRL, 30, 2215.

Kodera, K. & Kuroda, Y. (2002), JGR, 107, 4749.

Krivova, N. A., Solanki, S. K., & Unruh, Y. C. (2009a), JASTP,
doi:10.1016/j.jastp.2009.11.013

Krivova, N. A. et al. (2009b), JGR, 114, D00104.

Labitzke, K. (2003), Meteorol. Z., 12, 209.

Langematz, U., Matthes, K., & Grenfell, J. L. (2005), Memorie della Societa
Astronomica Italiana, 76, 868.

Lathuillere, C. et al. (2002), Ann. Geophys., 20, 1081.

[Lean, J.](#), Picone, J. M., & Emmert, J. T. (2009), JGR, 114, A07301.

Manoharan, P. K. (2010), Highlights of Astronomy, 15, 484.

Meier, R. R. (1991), Space Sci. Rev., 58, 1.

Mewaldt, R. A. et al. (2010), ApJ Lett., 723, L1.

National Space Weather Strategic Plan (1995), Off. Fed. Coord. for Met. Services,
NOAA, Silver Spring, MD.

Rottman, G. J. (1988), Adv. Sp. Res., 8, 53.

Rottman, G. J. et al. (1993), JGR, 98, 10667.

Schmidke, G. et al. (2006), Adv. Sp. Res., 37, 273.

Snow, M., McClintock, W. E., & Woods, T. N. (2010), Adv. Sp. Res., 46, 296.

[Thuillier, G. & Bruinsma \(2001\) – Appears on page 4.](#)

Tobiska, W. K., Bouwer, S. D., & Bowman, B. R. (2008), JASTP, 70, 803.

Woods, T. N. et al. (2000), JGR, 110, A01312.

Adoption of a Paradigm Shift for Space Exploration

Jan J. Sojka

Center for Atmospheric and Space Sciences
Utah State University

As the space science community wrestles with the present-day science mission's absorbing resources for much of a decade at a time and for most of the next decade, the question may legitimately be asked, "is science and society getting the best return for the resource commitments?"

Today's outstanding Space Science questions are often integrally associated with society's advancing technologies. Indeed, the developing reliances on space-based assets and evermore space weather sensitive ground based technologies are forcing science, commerce, and DoD together on common issues. In turn, this leads to potentially powerful cost effective synergisms. The most profound example of this "parasitic behavior" has been playing itself out over the past 20 years, namely GPS. To begin with, it was a necessary DoD geolocation system. Then, a worldwide geodetic community attached itself to its data streams. Worldwide deployments of dual frequency GPS-RINEX receivers followed. The last to come to the table were the space physicists who found in the "throw-away" geospace data a goldmine of scientific data. Today this goldmine, still at no cost to space science, is the basis for ionospheric space weather specifications at levels of precision unimaginable ten years earlier and most certainly unattainable with existing space science resources. The final message is that today it is the DoD and the commercial communities that use these "ionospheric space weather" products¹.

Is such a paradigm of "DoD-Commerce-Science" leveraging unreasonable?

In these three arenas, it is well appreciated that today's outstanding science questions left unanswered after fifty-some years of space exploration primarily fall into the category of systems engineering in which not only are stand-alone regions (disciplines) integrally coupled, but that all components of the "system" are spatially inhomogeneous over many orders of magnitude and whose temporal dynamics are present over many orders as well. The most dramatic space science event, and the least understood, is the Super Storm, whose impacts are felt in the DoD, Commerce, and science arenas. Yet, our knowledge of the evolution from the Sun to the Gigantic Ionospheric Current (GIC), which took Hydro-Quebec power system down for more than a day, is with the most favorable interpretation quantitatively understood but only after the event. A Carrington Event aimed at Earth would impact all, and leave the scientist blinded by all that we did not understand!

So, where is this going? What is a possible paradigm shift?

If we realistically define the outstanding space science questions of today they fall into two categories. The first is what NASA and the NSF have excelled at: go where no instrument (man/woman) has gone before, and explore. No one can argue that science and humanity have not benefited. Are there still such explorations to be done? Of course there are. The answer is yes. But behind this wave of exploration and opening science frontiers there is a trail of unfinished

¹ The enabling link between the science and commerce/DoD communities could be viewed simply as scientific data that can today be recovered just as cost effectively in real-time as via tape recorded delay. This one level of technology advance provides commerce and DoD scientific data streams in necessary and timely form. The new SDO mission is an excellent example of this joint role, as would be the ACE and other missions.

science! This is the second category and it does not come with the same one-time worldwide acclaim. Rather, the questions, the processes to be understood, and knowledge gained are difficult in quite different ways to those of the exploration mission. If indeed this second question is just a matter of “dotting the i’s and crossing the t’s,” then I would be the first to argue this second set of questions are indeed secondary. Of course, the truth is that to the contrary they are profound and they severely impact humanity and our technical world.

So how are these second space science questions to be answered?

1. Requires multi-point observations.
2. Requires interdisciplinary teams.
3. Requires resource leveraging from the highest levels of agencies that have science, commerce, and DoD interests.

What are potential models of this paradigm? GPS has become such a model, but only by accident based upon the commercial and scientific sectors innovation. The present-day SDO-type missions provide invaluable information to the commerce space weather community and potentially to the DoD. Magnetometer networks similarly provide cross agency value, but here we see the failure of kluging pieces together. Each network has its own magnetometer interface/data format! This generates a practical road block to interdisciplinary progress. GPS, irrespective of the manufacturer or operator, provides the world a standard format RINEX. The NSF has discussed multi-point systems and has a pilot DASI (Distributed Arrays of Space Instruments) system in mind. In ecology, a similar approach called NEON is viewed as the essential experiment to move forward. NASA, through STEREO and the future MMS missions, also understands that some outstanding science questions need more than one observational point.

The paradigm shift in question is to fruitfully utilize this next decade to explore the solution methods to the major science-commerce-DoD problems/questions and establish procedures for their solution. If one chooses a Carrington Event such as the Space Weather problem, it would be easy to quit before starting. However, other questions/problems exist without having to go to this extreme. The writer’s own biases will not be raised here, rather a procedure by which interagency cooperation will enable compelling relevant problems needing a scientific knowledge to enable commerce and society to enact solutions is the way forward.

Modeling and Measurement of Upper Atmosphere Climate Change

White Paper submitted to the Decadal Survey of Solar and Space Physics

Stanley C. Solomon and Liying Qian
High Altitude Observatory
National Center for Atmospheric Research

Summary

Human activity has caused subtle changes in the Earth's atmosphere over the past several decades, due to increased emissions of gases such as carbon dioxide (CO₂) and methane (CH₄). These anthropogenic changes in greenhouse gases have the potential for increasing temperature in the lower atmosphere due to their ability to absorb infrared radiation. The far-reaching consequences of these changes are now being seen in the upper atmosphere of the Earth, especially in the thermosphere. Since the prediction of *Roble and Dickinson* [1989] that a consequence of increasing CO₂ levels would be to decrease the temperature of the upper atmosphere, opposite to the response of the lower atmosphere, several investigators have attempted to quantify observed effects on the mesosphere, thermosphere, and ionosphere. In the thermosphere, one effect of a cooler atmosphere is to cause global density to decrease as the thermosphere cools and contracts. This in turn causes changes in the ionosphere, lowering ionospheric layer heights and changing their density. The thermosphere goes through natural, cyclical changes in density driven by the Sun's 11-year activity cycle, heating and expanding at solar maximum, cooling and contracting at solar minimum, which makes modeling of secular change a challenging process, due to the interaction of solar and anthropogenic effects. In addition, the evolving terrestrial magnetic field influences ionospheric change, and the role of magnetosphere-ionosphere coupling must be considered. In order to quantify and understand global change in the upper atmosphere, three-dimensional general circulation models of the thermosphere-ionosphere-mesosphere system should be integrated with global Earth system models that include the entire atmosphere, and coupled to interactive ocean and land mass models. To quantify the fundamental physics of the upper atmosphere radiative energy balance, space-based measurements of infrared emissions and cooling rates must be performed, and spectrally-resolved measurements of the full-disk solar irradiance at ultraviolet and X-ray wavelengths need to continue. Observational data of middle atmosphere, thermosphere, and ionosphere change on solar-rotational, solar-cycle, and secular trend time scales should be analyzed and used to improve model fidelity and quantify uncertain aeronomical processes.

Objectives

Global change in the upper atmosphere is an important issue for at least three basic reasons: it is important to understand the full physical scope of anthropogenic change in the entire atmosphere, in order to establish thorough understanding of lower atmosphere effects; there are practical implications for low-Earth orbit space flight, satellite drag, and the accumulation of orbiting debris; the interplay of external forcing and the natural variability of the atmosphere with changes caused by anthropogenic trace gas emissions is an inherently compelling scientific topic for heliophysics. In order to fully understand and quantify secular change of thermospheric temperature and density, we must also obtain a complete and accurate description of solar and geomagnetic forcing. An integrated program using theoretical models, solar measurements, satellite drag derived inferences of thermospheric neutral density, ionospheric measurements, and infrared observations of thermospheric cooling rates, will be able to address key questions concerning global change in the upper atmosphere:

- Is the secular change in thermospheric density observed over the past ~40 years due only to anthropogenic changes in atmospheric constituents, or are changes in solar and geomagnetic activity also implicated?
- How does solar variability, the amplitude of the solar cycle, and the length and depth of solar minimum, interact with upper atmosphere climate change?
- How will anthropogenic change in the ionosphere be manifested in layer heights and densities, *E*-region conductivities, *F*-region electrodynamics, and ion and electron temperatures?
- How will changes in the terrestrial magnetic field influence magnetosphere-ionosphere coupling, auroral dynamics, ionospheric morphology, and geospace climate?
- Do variabilities in the lower- and middle-atmosphere, such as the QBO, the El Niño southern oscillation, and stratwarms, modulate mesosphere-thermosphere-ionosphere climate?
- Will hydrogen in the exosphere increase as a consequence of global change, and what would be the effect on the plasmasphere and magnetosphere?

Background

The temperature and density of the upper atmosphere are determined by the interaction of energetics, dynamics, and composition. The absorption of solar radiation is the primary heat source, and as the wavelength of solar photons decreases, they are generally absorbed at higher altitude. Solar ultraviolet irradiance varies on periods associated with the 11-year solar cycle and the 27-day solar rotational period with the variation increasing with decreasing wavelength. Thus, the atmospheric effects of solar variability increase with altitude, with higher temperature causing higher density at a constant altitude surface, due to the expansion of the atmosphere.

Solar EUV is completely absorbed in the thermosphere due to the large ionization cross-sections of the major species. The products of ionization and dissociation initiate exothermic chemical reactions that determine time-varying neutral composition and ion composition. These photoionization and photodissociation processes thus transfer photon energy to chemical potential energy and kinetic energy that heats the neutral atmosphere, creates the ionosphere, and dissociates the major molecular atmospheric gases, N_2 and O_2 , into atoms. Orbits of satellites in the thermosphere change due to the frictional drag generated by collisions with these atoms and molecules; the amount of orbital decay is determined primarily by the neutral density. Thus, satellite orbit decay caused by atmospheric drag is both a technical problem we wish to address through the use of prognostic modeling techniques, and, through analysis of change of orbital elements, a means by which to measure the thermosphere and thereby understand its response to external forcing and internal processes. Additional variability is caused by magnetospheric energy input at high latitude, the propagation of waves and tides from the lower/middle atmosphere, and the resulting neutral winds.

Thermospheric neutral density routinely has daily changes normally on the order of 10%, but sometimes as large as a factor of two, and can have a solar cycle variation of an order of magnitude at altitudes in the 400–600 km range. Empirical models describe this variation by characterizing the temperature and density changes as a function of indexed parameters, and first-principles theoretical models, including general circulation models, have attempted to establish a theoretical description through quantification of the energy sources and sinks and the chemical and dynamical processes through which they interact [e.g., *Dickinson et al.*, 1981; *Fuller-Rowell et al.*, 1987; *Roble et al.*, 1988; *Richmond et al.*, 1992; *Roble and Ridley*, 1994].

Superimposed on this solar-driven variation is a gradual decrease in temperature and density caused by increasing CO_2 and other anthropogenic “greenhouse gases” (GHG). *Roble and Dickinson* [1989] first predicted that a consequence of increasing CO_2 levels would be to decrease the temperature of the upper atmosphere, opposite to the response of the lower atmosphere. The reason for this apparent paradox is that CO_2 and other heteronuclear molecules can emit infrared

radiation as well as absorb it. In the lower atmosphere, CO_2 absorbs radiation coming from the Earth, which excites it to higher vibrational states. Before it can re-emit the radiation, it undergoes collisions with other atmospheric gases, transferring the vibrational energy into heat. Any emission that does occur is likely to be re-absorbed by CO_2 at other altitudes. But above the tropopause the atmosphere becomes increasingly transparent to infrared radiation as densities decrease, and the radiation-excitation-collision process runs in reverse. Instead of absorbing radiation, CO_2 molecules are vibrationally excited by collisions, and then spontaneously emit in the infrared, which causes radiational cooling of the upper atmosphere.

CO_2 is the most important, but not the only, GHG with the potential to affect the upper atmosphere. CH_4 is both radiationally active in the infrared and a source of hydrogen to the mesosphere and thermosphere, which may affect polar mesospheric cloud (noctilucent cloud) frequency, odd-hydrogen in the mesosphere, atomic hydrogen in the exosphere, and hydrogen escape. Chlorine and nitrogen oxides in the stratosphere change the O_3 balance and affect both odd-oxygen chemistry and stratospheric heating. Changes in water vapor consequent to tropospheric temperature change can also propagate through the middle atmosphere. Thus, other potential anthropogenic effects include mesosphere cooling, increases in the frequency and latitudinal extent of polar mesospheric clouds, increasing atomic hydrogen in the exosphere, and ionospheric changes. Recent reviews describe these and other possible effects, and assess the observational evidence [Beig et al., 2003; Lastovicka et al., 2006; 2008].

The thermospheric density trend predicted by Roble and Dickinson [1989] has been observed in long-term studies of satellite orbits [Keating et al., 2000; Emmert et al., 2004; Marcos et al., 2005; Lean et al., 2006; Emmert et al., 2008]. These studies have unambiguously observed these trends, with the changes in upper thermospheric density ranging from $\sim 2\%$ to $\sim 5\%$ per decade, depending on altitude and solar activity. Other model studies [e.g., Akmaev and Fomichev, 1998, 2000; Akmaev et al., 2006; Gruzdev and Brasseur, 2005] similarly predicted a downward secular trend in thermosphere and mesosphere temperatures. Qian et al. [2006] explained the solar cycle dependence as due to modulation by nitric oxide cooling rates, using an updated version of the Roble et al. [1987] global mean model, and obtained good agreement with some of the observational data. However, the latest analyses [Emmert et al., 2008] confirms the large changes for solar minimum conditions, 5% per decade or more, first reported by Keating et al. Thus, the observational evidence, while confirming thermospheric cooling, indicates that models may underestimate the rate of change. Figure 1 compares observational and model estimates of the rate of thermospheric density change as a function of solar activity.

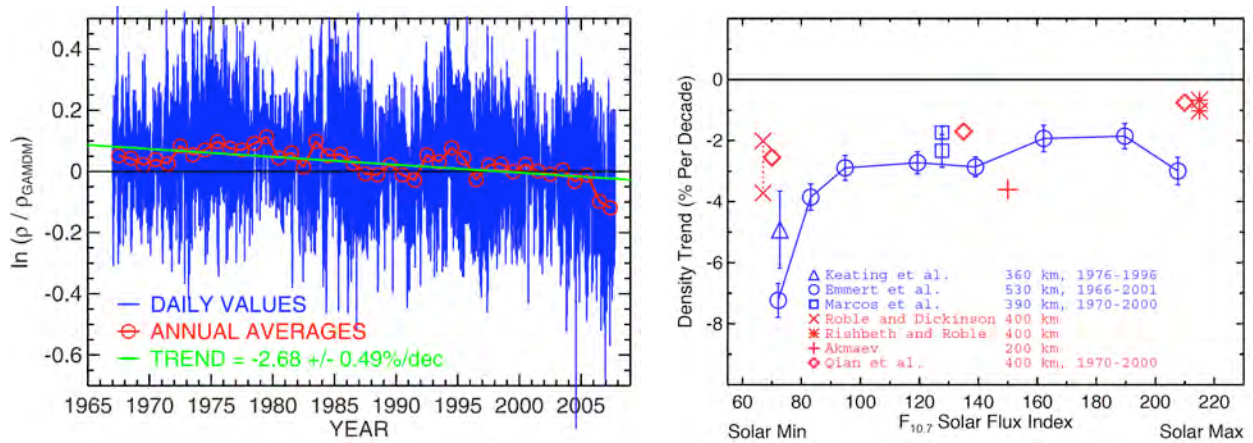


Figure 1. Left: Global average thermospheric density residuals at 400 km obtained from atmospheric drag on ~ 5000 orbiting objects, and inferred secular trend [Emmert et al., 2008]. Right: Measurement and model estimates of thermospheric density trends as a function of solar activity [Lastovicka et al., 2008].

Additional evidence comes from radar measurements of ion temperature. *Holt and Zhang* [2008] found a trend of -47 K per decade in ion temperatures at 350–400 km measured by the incoherent scatter radar at Millstone Hill, MA. This implies a neutral temperature decrease significantly larger than expected from model simulations or seen in satellite drag measurements, and, if substantiated, would have significant implications for the ionosphere as well.

Changes in the ionosphere should accompany the neutral atmosphere changes, including changes in the height of ionospheric layers due to cooling and contraction, and small changes in peak density [*Rishbeth and Roble*, 1992]. Since ionospheric measurements, especially ionosondes, have a long history, these records could provide a means of assessing secular change. Peak densities at E and F region altitudes (N_mE and N_mF_2) are easier to measure in a stable long-term manner than the peak altitudes (h_mE and h_mF_2), since they are based on critical frequency. However, model predictions of peak density change reveal a complex scenario, due to the interaction of atmospheric contraction with changes in composition and chemical rates. The E and F_1 regions are expected to actually increase slightly in density as they reduce in altitude, and very small changes are anticipated in N_mF_2 . Observational evidence is reviewed by *Lastovicka et al.* [2008]; model results and the underlying theory are described by *Qian et al.* [2008; 2009].

Ionospheric secular change is also complicated by evolution of the terrestrial magnetic field. Over the past century, the dipole strength has decreased about 6%, and non-dipole features of the field also evolve, affecting magnetic declination and inclination angles and the locations of the magnetic poles and equator. The magnetic equator in the Atlantic sector has moved northward as much as 15° in the past century, and the equatorial ionization anomaly moved correspondingly. The electrical conductivity of the ionosphere depends on the strength and direction of the geomagnetic field [e.g., *Takeda*, 1996], so the ionospheric dynamo and consequent geomagnetic perturbations are affected by field changes. The strength and orientation of the geomagnetic dipole determine how the solar wind and interplanetary magnetic field interact with the magnetosphere, and therefore influence the auroral particle precipitation and high-latitude Joule heating. All of these changes affect the ionosphere [e.g., *Cnossen and Richmond*, 2008], its interaction with the magnetosphere, and its influence on the thermosphere. In order to understand how observations of long-term trends in the ionosphere relate to thermospheric changes, we need to be able to identify and quantify the effects associated with changes in the geomagnetic field.

An additional complexity now appears to have been added to the problem of quantifying anthropogenic change cast against a background of considerable natural variation. During 2008 and 2009, solar activity became extremely low, and the onset of solar cycle 24 was late and weak [e.g., *Russell et al.*, 2010]. A variety of evidence concerning the near-Earth space environment, including ionospheric parameters, radiation belt levels, and thermospheric density, indicated that conditions were significantly quieter than the “usual” solar minimum, and the two-year period between mid-2007 and mid-2009 was one of the longest and quietest solar minima on record. There is evidence from space-based solar measurements that the solar EUV irradiance was also anomalously low during this time, or at least lower than the previous solar minimum [*Didkovsky et al.*, 2010]. And, analyses of satellite orbits show that the upper thermosphere was significantly cooler and less dense than the previous several solar minima, and indeed cooler than any other such period since the beginning of space flight [*Emmert et al.*, 2010]. Figure 2 shows the satellite drag and solar EUV measurements. Model simulations supporting this are shown in *Solomon et al.* [2010]. Clearly, the 2008-2009 solar minimum was well outside of the expected range due to anthropogenic effects, and solar measurements are commensurate with the atmospheric response. A much smaller change is seen, however, in standard solar proxy indices such as $F_{10.7}$. Since comparison of successive solar minima was employed by *Keating et al.* [2000] to make the original detection of thermospheric cooling, and could be a valuable technique for re-

moving solar cycle effects, the ramifications of this finding are significant. Quantification of past and future solar EUV levels will be a crucial aspect of understanding anthropogenic global change in the upper atmosphere.

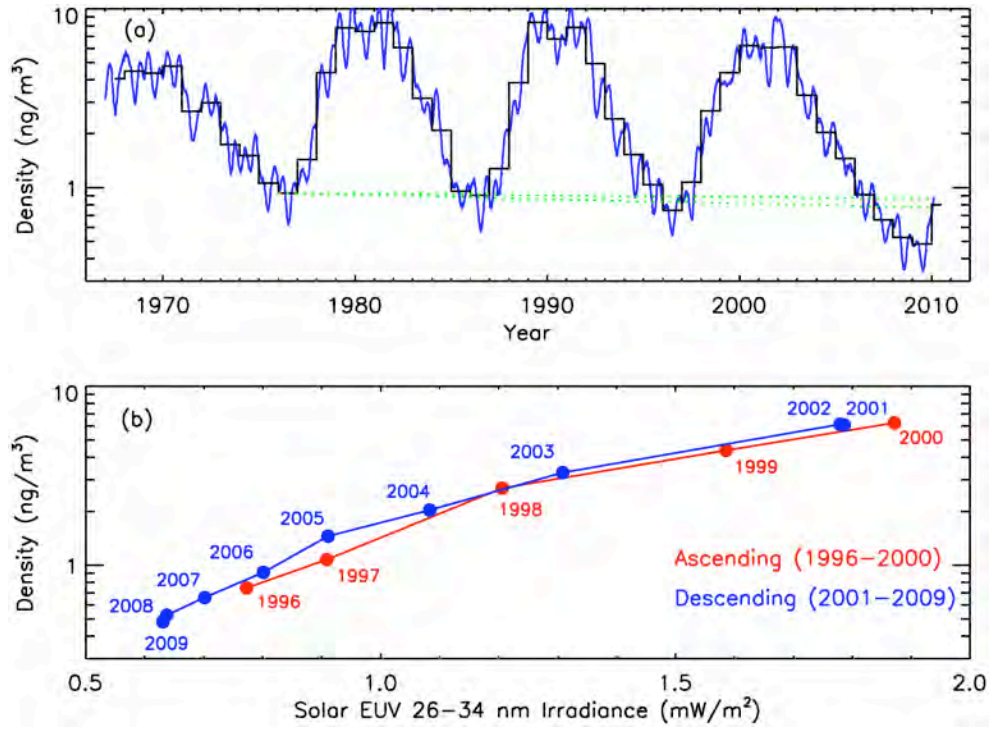


Figure 2. (a) Global mean thermospheric density at 400 km, obtained from satellite orbital parameters over four solar cycles. Blue: 81-day centered running mean. Black: annual average. Green dotted lines: envelope of expected decrease due to increasing CO₂ levels, in the range of 2% to 5% per decade, starting with the 1976 annual average. (b) Global mean thermospheric density annual average plotted as a function of the 26–34 nm solar EUV irradiance annual average measured by the SOHO SEM instrument for the ascending (red) and descending (blue) phases of solar cycle 23 [Solomon *et al.*, 2010].

Necessary Measurements

In addition to ongoing monitoring of satellite drag, ionospheric parameters, geomagnetic activity indices, and greenhouse gas concentrations, future measurements of the major heating and cooling terms in the upper atmosphere will be required.

Solar Inputs

The Solar EUV Experiment (SEE) [Woods *et al.*, 1998; 2005] on the TIMED satellite provides measurements in the spectral range 0.1–195 nm, and the SORCE X-ray Photometer System (XPS) augments this data set. Another broad-band data set is provided by the Space Environment Monitor (SEM) on the SOHO spacecraft [Didkovsky *et al.*, 2010] (see above). Because the soft X-ray measurements on TIMED and SORCE are broad-band photometric measurements, similar to the earlier SNOE measurements [Bailey *et al.*, 2000], considerable uncertainty as to their spectral interpretation remains. However, SDO/EVE now measures the solar spectrum from 6 to 106 nm at 0.1 nm resolution with a 10 s cadence [Woods, *et al.*, 2010]. These measurements will be the key to resolving issues in the soft X-ray range, understanding flare irradiance variability, and validating solar spectral irradiance models. In the FUV, measurements by instruments on the UARS satellite [Woods and Rottman, 2002] and the SOLSTICE instrument on board the SORCE satellite [Rottman, 2005, 2006] measure the solar spectrum from H Lyman-alpha (121.6

nm) through 300 nm. In addition, the SIM instrument on SORCE covers the range 200 nm through the infrared. Low-resolution instruments on the NOAA GOES satellites now augment this measurement set. It will be extremely important to obtain a comprehensive inter-calibration between TIMED SEE, SOHO/SEM, and the new SDO/EVE data, in order to quantify and understand solar irradiance during the recent solar minimum, and its deviation from past solar minima. For the future, maintaining a continuous program of both high-resolution solar irradiance measurements throughout the spectrum, and low-resolution monitoring (by NOAA GOES) as well, will be crucial to understanding upper atmosphere climate, and also the small but non-negligible solar perturbations of stratosphere and troposphere climate.

Infrared Cooling Rates

Considerable progress has been made understanding the role of radiatively active minor species in the upper atmosphere, due to satellite observations such as from UARS, CRISTA, SNOE, TIMED, and others. Global integrated thermospheric cooling rates for CO₂ and nitric oxide (NO) obtained from measurements by the TIMED/SABER instrument (updated from the data of *Mlynczak et al.*, 2007) are shown in figure 3. The decline in nitric oxide (NO) cooling during the course of the solar cycle has been successfully simulated, but there are discrepancies between models and measurements of the CO₂ cooling rates, as discussed below. These rates are derived from zonal-average rates as a function of latitude and altitude. Infrared measurements will also be valuable for verification of the model altitude profiles of the CO₂ 15- μ m emission and NO 5.3- μ m emission, which are critical for climate change studies. Comprehensive future measurements of atmospheric infrared emissions will be necessary to understand the evolution of the terrestrial climate system at all altitudes.

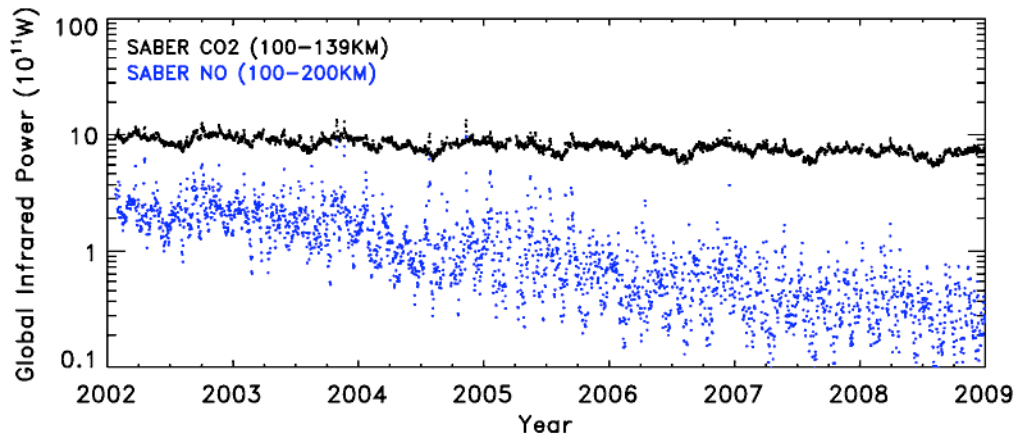


Figure 3. Global integrated thermospheric cooling rates by carbon dioxide and nitric oxide derived from SABER data, over the course of the TIMED mission.

Model Development

Aeronomical updates to the original *Roble and Dickinson* [1989] estimates of thermospheric change now produce significantly more NO, in better agreement with observations. In the earlier studies, NO did not enter significantly into the energy budget of the thermosphere and nearly all of the cooling was due to CO₂. Current work [*Qian et al.*, 2006; 2008; 2009] indicates significant changes to the predicted global change response, primarily based on increased cooling from the NO 5.3 μ m band. Since NO radiation is much larger during solar maximum, NO cooling acts as a regulator, causing the predicted global change to become smaller at solar maximum than at solar minimum due to the decreased importance of CO₂ cooling at solar maximum. These estimates agree with the overall rate of change found by *Marcos et al.* [2005], and exhibit a similar solar cycle dependence found by *Emmert et al.* [2004], but recent results by *Emmert et*

al. [2008] show even more rapid change, especially at solar minimum. Therefore, having established the importance of NO cooling, it is necessary to re-evaluate CO₂ cooling. The laboratory values for the O-CO₂ excitation rate used in the NCAR TIME-GCM and global mean models have historically been lower than those derived from atmospheric measurements or employed in planetary atmosphere models. Global cooling rate data from the TIMED/SABER (c.f., figure 3) agree with model cooling rates for NO, but are considerably higher, about a factor of two, for CO₂. This offers support for the higher collisional excitation rates, and could explain why the model apparently underestimates observed changes. However, changing the model CO₂ cooling rate will force a revision of solar heating, conductance, and other parameters controlling the energy budget, in order to retain agreement with observed temperatures and their solar cycle dependence.

In addition to a strong upper atmosphere and geospace modeling program, it will be necessary to continue the development of theoretical and empirical models of solar irradiance variability, especially throughout the ultraviolet.

Data Analysis

Long-term analysis of climate trends will require diverse programs to analyze existing data sources and exploit proposed new measurements. The following overview is not an exhaustive list but describes the scope of the effort.

Mesosphere

Polar mesospheric cloud observational analysis from the AIM mission, space-based temperature profiles from limb radiometers, and ground-based airglow and lidar observations will be important components of trend analysis and lower-atmosphere interactions. Changes to the hydrogen budget of the mesosphere as a consequence of temperature change and CH₄ increases should be monitored.

Thermosphere

Global-average neutral density derived from satellite drag will continue to be a primary tool for measuring the integrated climate of the thermosphere. New empirical descriptions of the spatial distribution of thermospheric density and temperature, based on data assimilation approaches to satellite drag analysis, can augment this. Ultraviolet remote sensing of thermosphere/ionosphere properties is extremely valuable as a global observational tool, and strongly complement the infrared and solar measurements discussed above.

Ionosphere

Because the ionosphere is much more sensitive to highly variable electromagnetic forces than the thermosphere, resulting in a considerably more complex climatology, the primary focus for ionospheric studies will be to estimate the relative importance of different influences and to understand the underlying physical processes. The predicted responses of the ionospheric N_mF_2 , h_mF_2 , and total electron content (TEC) at different geographic locations to changes in greenhouse gases, solar EUV, geomagnetic activity, and the evolving terrestrial magnetic field need to be modeled and compared with long-term data bases of ground-based ionosonde and radar observations, and future measurements such as from the COSMIC-2 constellation.

Exosphere

Calculations of the atomic hydrogen budget, which is a product of CH₄ dissociation, predict a large increase in exospheric H, which could ultimately be detectable by ground-based H _{α} observations. Understanding processes governing the rate of hydrogen escape is important for studies of anthropogenic climate change, but also has implications for climate evolution on geologic time scales. Additionally, changes to H density in the exosphere would effect plasmaspheric H⁺, potentially ion outflow, and the magnetosphere.

References

- Akmaev, R. A., and V. I. Fomichev, Cooling of the mesosphere and lower thermosphere due to doubling of CO₂. *Ann. Geophys.*, 16, 1501–1512, 1998.
- Akmaev, R. A., and V. I. Fomichev, A model estimate of cooling in the mesosphere and lower thermosphere due to the CO₂ increase over the last 3–4 decades. *Geophys. Res. Lett.*, 27, 2113–2116, 2000.
- Akmaev, R. A., V. I. Fomichev, and X. Zhu, Impact of middle-atmospheric composition changes on greenhouse cooling in the upper atmosphere. *J. Atmos. Terr. Phys.*, 68, 1879–1889, 2006.
- Bailey, S. M., T. N. Woods, C. A. Barth, S. C. Solomon, L. R. Canfield, and R. Korde, Measurements of the solar soft X-ray irradiance by the Student Nitric Oxide Explorer: First analysis and underflight calibrations. *J. Geophys. Res.*, 105, 27179, 2000.
- Beig, G., P. Keckhut, R.P. Lowe, et al., Review of mesospheric temperature trends, *Reviews of Geophysics*, 41(4), 1015, 2003.
- Cnossen, I., and A. D. Richmond, Modelling the effects of changes in the Earth's magnetic field from 1957 to 1997 on the ionospheric hmF2 and foF2 parameters. *J. Atmos. Solar-Terr. Phys.*, 70, 1512–1524, 2008.
- Dickinson, R. E., E. C. Ridley, and R. G. Roble, A three-dimensional general circulation model of the thermosphere, *J. Geophys. Res.*, 86, 1499–1512, 1981.
- Didkovsky, L. V., D. L. Judge, and S. R. Wieman, Minima of solar cycles 22/23 and 23/24 as seen in SOHO/CELIAS/SEM absolute solar EUV flux, in *ASP Conference Series, SOHO-23: Understanding a Peculiar Solar Minimum*, S. R. Cranmer, J. T. Hoeksema, and J. Kohl, eds., Vol. 428, p. 73, 2010. http://arxiv.org/PS_cache/arxiv/pdf/0911/0911.0870v1.pdf.
- Emmert, J. T., J. M. Picone, J. L. Lean and S. H. Knowles, Compelling evidence of a secular decrease in density, *J. Geophys. Res.*, 109, A02301, 2004.
- Emmert J. T., J. M. Picone, R. R. Meier, Thermospheric global average density trends, 1967–2007, derived from orbits of 5000 near-Earth objects, *Geophys. Res. Lett.*, 35, L05101, doi:10.1029/2007GL032809, 2008.
- Emmert, J. T., J. L. Lean, and J. M. Picone, Observations of record-low thermospheric densities during the 2008 solar minimum, *Geophys. Res. Lett.*, 37, L12102, doi:10.1029/2010GL043671, 2010.
- Fuller-Rowell, T. J., D. Rees, S. Quegan, R. J. Moffett, and G. J. Bailey, Interactions between neutral thermospheric composition and the polar thermosphere using a coupled global model, *J. Geophys. Res.*, 92, 7744, 1987.
- Gruzdev, A. N., and G. P. Brasseur, Long-term changes in the mesosphere calculated by a two-dimensional model, *J. Geophys. Res.*, 110, D03304, 2005.
- Holt, J. M., and S. R. Zhang, Long-term temperature trends in the ionosphere above Millstone Hill, *Geophys. Res. Lett.*, 35, L05813, doi:10.1029/2007GL031148, 2008.
- Keating, G. M., R. H. Tolson, and M. S. Bradford, Evidence of long-term global decline in the Earth's thermospheric densities apparently related to anthropogenic effects, *Geophys. Res. Lett.*, 27, 1523–1526, May 15, 2000.
- Lean, J. L., J. M. Picone, J. T. Emmert, and G. Moore, Thermospheric densities derived from spacecraft orbits: Application to the Starshine satellites, *J. Geophys. Res.*, 111, A04301, doi:10.1029/2005JA011399, 2006.
- Laštovička, J., R. A. Akmaev, G. Beig, J. Bremer, and J. T. Emmert, Global change in the upper atmosphere, *Science*, 314, 1253, 2006.
- Laštovička, J., R. A. Akmaev, G. Beig, J. Bremer, J. T. Emmert, C. Jacobi, M. J. Jarvis, G. Nedoluha, Y. I. Portnyagin, and T. Ulich, Emerging pattern of global change in the upper atmosphere and ionosphere, *Ann. Geophys.*, 26, 1255, 2008.

- Marcos, F. A., J. O. Wise, M. J. Kendra, N. J. Grossbard and B. R. Bowman, Detection of a long-term decrease in thermospheric neutral density, *Geophys. Res. Lett.*, 32, L04103, doi:10.1029/2004GL021269, 2005.
- Mlynczak M. G., F. J. Martin-Torres, B. T. Marshall, R. E. Thompson, J. Williams, T. Turpin, D. P. Kratz, J. M. Russell III, T. N. Woods, and L. L. Gordley, Evidence for a solar cycle influence on the infrared energy budget and radiative cooling of the thermosphere, *J. Geophys. Res.*, 112, A12302, doi:10.1029/2006JA012194, 2007.
- Qian, L., R. G. Roble, S. C. Solomon, and T. J. Kane, Calculated and observed climate change in the thermosphere, and a prediction for solar cycle 24, *Geophys. Res. Lett.*, 33, L23705, doi:10.1029/2006GL027185, 2006.
- Qian, L., R. G. Roble, S. C. Solomon, and T. J. Kane, Model simulations of global change in the ionosphere, *Geophys. Res. Lett.*, 35, L07811, doi:10.1029/2007GL033156, 2008.
- Qian, L., A. G. Burns, S. C. Solomon, and R. G. Roble, The effect of carbon dioxide cooling on trends in the F2-layer ionosphere, *J. Atmos. Sol.-Terr. Phys.*, 71, 1592, doi:10.1016/j.jastp.2009.03.006, 2009.
- Richmond, A. D., E. C. Ridley, and R. G. Roble, A thermosphere/ionosphere general circulation model with coupled electrodynamics, *Geophys. Res. Lett.*, 19, 601, 1992.
- Rishbeth, H., and R. G. Roble, Cooling of the upper atmosphere by enhanced greenhouse gases- Modeling of thermospheric and ionospheric effects, *Planet. Space Sci.*, 40, 1011-1026, 1992.
- Roble, R. G., E. C. Ridley, and R. E. Dickinson, On the global mean structure of the thermosphere, *J. Geophys. Res.*, 92, 8745-8758, 1987.
- Roble, R. G., E. C. Ridley, A. D. Richmond, and R. E. Dickinson, A coupled thermosphere/ionosphere general circulation model, *Geophys. Res. Lett.*, 15, 1525, 1988.
- Roble, R. G. and R. E. Dickinson, How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere? *Geophys. Res. Lett.*, 16, 1441-1144, 1989.
- Roble, R. G. and E. C. Ridley, Thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (TIME-GCM): Equinox solar min simulations, 30-500 km, *Geophys. Res. Lett.*, 21, 417, 1994.
- Rottman, G., The SORCE mission, *Solar Physics*, 230 (1-2), 7, 2005.
- Rottman, G., Measurement of total and spectral solar irradiance, *Space Sci. Rev.*, 125, 39, 2006.
- Russell, C. T., J. G. Luhmann, and L. K. Jian, How unprecedented a solar minimum?, *Rev. Geophys.*, 48, RG2004, doi:10.1029/2009RG000316, 2010.
- Solomon, S. C., T. N. Woods, L. V. Didkovsky, J. T. Emmert, and L. Qian, Anomalously low solar extreme-ultraviolet irradiance and thermospheric density during solar minimum, *Geophys. Res. Lett.*, 37, L16103, doi:10.1029/2010GL044468, 2010.
- Takeda, M., Effects of the strength of the geomagnetic main field strength on the dynamo action in the ionosphere, *J. Geophys. Res.*, 101, 7875, 1996.
- Woods, T. N., S. M. Bailey, F. Eparvier, G. Lawrence, J. Lean, W. McClintock, R. Roble, G. Rottman, S. C. Solomon, W. K. Tobiska, G. Ucker, and O. R. White, TIMED Solar EUV Experiment, *Proc. SPIE*, 3442, 180, 1998.
- Woods, T. N., and G. J. Rottman, Solar ultraviolet variability over time periods of aeronomic interest, in *Comparative Aeronomy in the Solar System*, AGU Monograph, 2002.
- Woods, T. N., F. G. Eparvier, S. M. Bailey, P. C. Chamberlain, J. Lean, G. J. Rottman, S. C. Solomon, W. K. Tobiska, and D. Woodraska, The Solar EUV Experiment (SEE): Mission overview and first results, *J. Geophys. Res.*, 110, A01312, doi:10.1029/2004JA010765, 2005.
- Woods, T. N., et al., Extreme Ultraviolet Variability Experiment (EVE) on the Solar Dynamics Observatory (SDO): Overview of Science Objectives, Instrument Design, Data Products, and Model Developments", *Solar Physics*, doi: 10.1007/s11207-009-9487-6, 2010.

Dynamic Geospace Coupling Mission

James F. Spann¹, Lynn Kistler², Robert Strangeway³, Larry Paxton⁴, Dennis Gallagher¹, Don Mitchell⁴, Nicky Fox⁴, Laila Andersson⁵, Stephen Mende⁶, Damien Chua⁷, Matthew Fillingim⁶

¹ National Aeronautics and Space Administration, Marshall Space Flight Center

² University of New Hampshire

³ University of California at Los Angeles

⁴ The Johns Hopkins University Applied Physics Laboratory

⁵ Colorado University

⁶ The University of California at Berkeley

⁷ Naval Research Laboratory

Science Goal

To understand how magnetospheric dynamics provides energy into the coupled ionosphere-magnetosphere system.

Science Rationale

The Dynamic Geospace Coupling science mission addresses fundamental questions related to plasma processes that determine how energy and momentum from the solar wind propagate downward through geospace to Earth. Understanding the dominant plasma acceleration and transport processes operative in the Earth's geospace region has general applicability across the Sun-Earth-Heliosphere coupled system. The number one recommendation for Heliophysics in the 2009 roadmap entitled *Heliophysics/Solar and Space Physics of a New Era*, is to "Implement a science target queue to address the most urgent heliophysics science problems facing the Nation" – see Executive Summary and page 64. The mission described in this whitepaper fulfills this recommendation.

The question that specifically motivates this science mission is how mass and energy flows in magnetosphere-ionosphere systems determine and control their coupled behavior. Particle acceleration and particle transport are two of the most fundamental processes in the plasma universe occurring in essentially every region of the heliosphere from the Sun to the heliopause, in the solar wind, in planetary magnetospheres and ionospheres, in cometary tails, and many others. The same processes that have been discovered by Heliophysics missions form a foundation for understanding astrophysical systems. It is in the magnetosphere-ionosphere system where the basic physical processes of particle acceleration and transport have the most direct impact on human activity through space weather effects on satellite operations, through disruption of electromagnetic signals passing through the ionosphere, and through dramatic reconfigurations of the electrodynamic currents that connect the Earth to the heliosphere through the ionosphere and magnetosphere. The DGC science goal is specifically formulated to understand how particle acceleration and transport couple the ring current and inner plasma sheet (the primary repositories of energetic particles in the magnetosphere) to the ionosphere through current fields, and the flow of mass and energy.

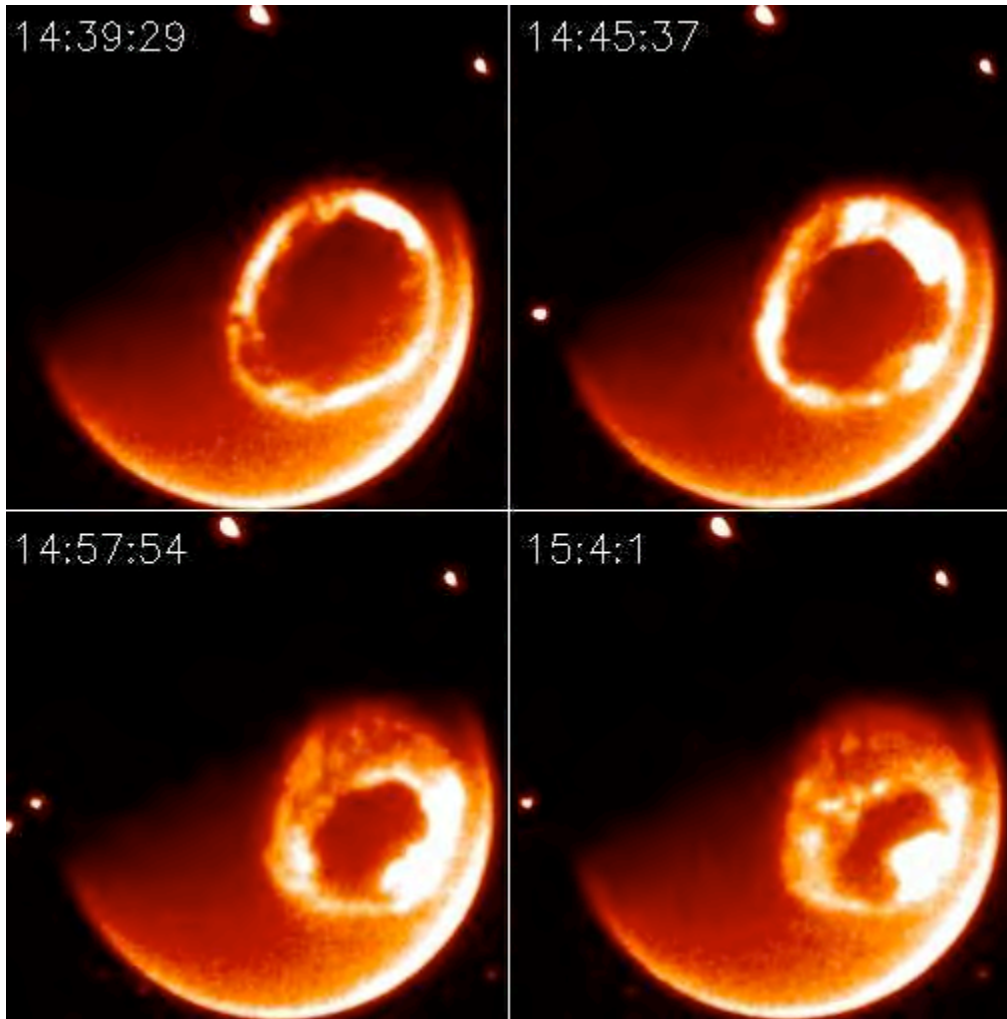


Figure Caption: Evolution of the aurora during the July 15, 2000, Bastille Day magnetic storm. 1423UT–1512UT

The next step in answering these questions will be achieved through simultaneous measurement of the dynamics of the ionospheric and magnetospheric auroral regions over global scales with high time and spatial resolution. Magnetospheric dynamics can be understood through global energetic neutral atom (ENA) imaging of the ring current and extreme ultraviolet imaging (EUV) of the plasmasphere, complemented by in situ measurements across spatial scales characteristic of global circulation while ionospheric dynamics can be remotely probed through auroral imaging. The aurorae are much more than a passive response to the deposition of energy into the ionosphere by magnetospheric particles and currents. Auroral acceleration and heating change both ionospheric and magnetospheric currents as well as provide ionospheric plasma to the magnetosphere magnetosphere through light and heavy ion outflow. The plasmasphere is similarly an active participant in the formation, dynamics, and decay of the ring current. Wave particle processes and the wave propagation of energy are strongly influenced by plasmaspheric densities and composition. The ever-changing distribution of plasmaspheric plasma is also a direct measure of the electrodynamic coupling of

magnetospheric high-energy plasma with the ionosphere. Achieving insight into the nature of these processes, their linked responses to solar wind driving, and the interrelationships between different regions is the key to understanding dynamic geospace coupling.

Mapping to NASA Heliophysics Roadmap Research Focus Areas, and NRC Decadal Challenges

The Dynamic Geospace Coupling science mission primarily falls under research focus areas H2 to understand changes in the Earth's magnetosphere, ionosphere, and upper atmosphere to enable specification, prediction, and mitigation of their effects and F2 to understand the plasma processes that accelerate and transport particles.

The relevant to Decadal Survey Challenges are:

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.

Measurements

Measurements over a range of spatial and temporal scales could include:

- Composition and morphology of the inner magnetosphere and plasma sheet
- Composition, conductivity and morphology of the ionosphere in the polar regions
- Large scale conjugate observations of the aurora
- High temporal/spatial/spectral resolution observation of the aurora
- Knowledge of the E and B fields in the inner magnetosphere extending into the ionosphere

To understand coupling between the ionosphere and magnetosphere systems it is necessary to make measurements of the dynamics of both. Magnetospheric dynamics are best measured by global magnetospheric imaging of the ring current and plasma sheet supplemented by probes at multiple locations within the system separated spatially on scales that are characteristic of global circulation. This implies global ENA imaging, presumably from the same platform as that which hosts the EUV/FUV imaging, supplemented by in situ measurements of plasma, energetic particles, electric, and magnetic fields at as many points separated by distances of Earth radii as feasible. While ionospheric dynamics can be studied in detail by in situ measurements, another valuable tool is remote sensing through auroral imaging. Global imaging provides system wide context while high spatial and temporal resolution imaging of more limited regions probe detailed coupling. Simultaneous, conjugate ENA and auroral imaging provides critical information on global ring current evolution and pitch angles, and on how ionospheric conditions differ in the two hemispheres and what those differences can tell us about

magnetospheric coupling—in particular, the flow of particles and currents between different parts of the system.

An implementation concept to achieve the goals of this mission is based on having two polar HEO or Molniya orbiting spacecraft manifested primarily with remote sensing FUV, EUV and ENA imagers/spectrometers, complimented by two or three LEO equatorial spacecraft manifested primarily with particle, waves and plasma instruments. Furthermore, it is advantageous to have a plasma instrument and a magnetometer on each imaging observatory, and to have one wide-FOV imager on each LEO in situ observatory.

Enhancing Technologies

Technology investments to enhance the return of the Dynamic Geospace Coupling mission include:

- High spatial and temporal auroral imaging in the EUV/FUV wavelengths
- Improved sensitivity and resolution for ENA optics/sensors
- Electric field instruments with lower spacecraft impact than current systems
- Low-resource plasma and energetic particle measurements with ion composition
- Innovative systems to neutralize spacecraft potential, and technologies
- Processes to reduce cost and risk in multispacecraft missions

Enhancing Pre-mission R&A Focus Areas

The science return from the Dynamic Geospace Coupling mission could be significantly enhanced through advance research and analysis of data collected from simultaneous measurements of magnetospheric dynamics and auroral processes. Theory and modeling of global electrodynamics, the effects of particle precipitation, and expectations for auroral conjugacy would also provide high value.

Example science questions for the Dynamic Geospace Coupling Mission are:

- How do magnetospheric particle and field signatures relate to visible/UV auroral forms?
- To what extent are auroral forms conjugate?
- What magnetospheric parameters control the degree of conjugacy?
- What do nonconjugate auroral forms tell us about the ionospheric control of magnetosphere-ionosphere coupling?
- How does magnetosphere-ionosphere coupling control particle outflow from the auroral regions?
- What are the source of currents that couple the magnetosphere and ionosphere?
- What is the energy input from the magnetosphere to the ionosphere through precipitating particles?
- To what extent do the ionosphere and magnetosphere “drive” the dynamics of the other region?

Global Ionosphere-Thermosphere-Mesosphere (ITM) Mapping Across Temporal and Spatial Scales

A White Paper for the NRC Decadal Survey of Solar and Space Physics

Andrew Stephan, Scott Budzien, Ken Dymond, and Damien Chua
NRL Space Science Division

Overview

In order to fulfill the pressing need for accurate near-Earth space weather forecasts, it is essential that future measurements include both temporal and spatial aspects of the evolution of the ionosphere and thermosphere. A combination of high altitude global images and low Earth orbit altitude profiles from simple, in-the-medium sensors is an optimal scenario for creating continuous, routine space weather maps for both scientific and operational interests. The method presented here adapts the vast knowledge gained using ultraviolet airglow into a suggestion for a next-generation, near-Earth space weather mapping network.

Why the Ionosphere, Thermosphere, and Mesosphere?

The ionosphere-thermosphere-mesosphere (ITM) region of the terrestrial atmosphere is a complex and dynamic environment influenced by solar radiation, energy transfer, winds, waves, tides, electric and magnetic fields, and plasma processes. Recent measurements showing how coupling to other regions also influences dynamics in the ITM [e.g. Immel et al., 2006; Luhr, et al, 2007; Hagan et al., 2007] has exposed the need for a full, three-dimensional characterization of this region. Yet the true level of complexity in the ITM system remains undiscovered primarily because the fundamental components of this region are undersampled on the temporal and spatial scales that are necessary to expose these details.

The solar and space physics research community has been driven over the past decade toward answering scientific questions that have a high level of practical application and relevance. The result of that focus has been the proliferation of research captured under the heading of “space weather.” This phrase is chosen intentionally to highlight the commonality between space physics research and the focused effort of the meteorological community to make weather forecasting comprehensive, accurate, and accessible. Similar to how weather measurements and forecasts affect our daily lives, there also is a need for space weather measurements to inform missions, operations, and assets in the near-Earth space environment that facilitate modern life.

Current measurements are able to connect solar system events to isolated changes in the three-dimensional volume of the near-Earth space weather environment. The understanding of how that connection evolves in space and time is lost as sensors move away from the initial perturbation. Similar problems have been resolved in other areas of solar-terrestrial research. In particular, the solar community has developed results and

advances in temporal and spatial imaging of the Sun and its upper atmosphere all the way to Earth. Yet our understanding of ITM space weather, which is needed to connect those solar and interplanetary phenomena to terrestrial effects, is still in its infancy. Recent measurements and models have helped, but to reach the maturity on par with successful meteorological forecasts for rain, wind, and temperature, space weather requires significant, long-term commitment to making the measurements that track the development of the ITM system across temporal and spatial scales to rigorously expose the changes of this medium through which spacecraft, both manned and unmanned, must traverse. With a futuristic vision toward routine travel in the space environment, these forecasts must become as comprehensive, accurate, and accessible as the weather maps of today.

Measurement Approach

There are a wide variety of methods for mapping the ITM system. The measurement concept put forth in this document suggests a two pronged attack to filling the deficiency in making routine space weather maps: (1) a comprehensive, rapid global imaging system from geosynchronous or other high-altitude orbit, and (2) a complement of small, low-Earth orbiting sensors to provide the third dimension for contextual interpretation of the global images. In order to provide the global coverage, a remote sensing technique from a space-based platform is essential. Current data from recent remote sensing missions such as the NASA Thermosphere Ionosphere Mesosphere Energetics and Dynamics (TIMED) mission demonstrate the current state-of-the-art capability. These sensors make measurements of strips of the ITM system over the course of a day that are meshed together to create a global map as the Earth rotates under the satellite orbital plane. At the orbital period of 90 minutes, the sampling rate is not sufficient to properly address changes in the global ITM environment that occur on timescales as small as a few minutes. In addition, the spatial sampling also is limited by orbit characteristics as 15 orbits per day distributed around the globe yield effective zonal sampling of more than 20 degrees. Thus these types of maps are effectively a composite of temporal and spatial changes but produce the illusion of a global measurement. While the data are representative of the high-quality measurements and results that can be achieved, they must be interpreted in the context of various models that are available. Only a combination of comprehensive, dedicated spatial and temporal imaging can provide the same rigor in space weather maps as that found in the meteorological community.

The technique of using ultraviolet (UV) airglow for remote sensing of the ITM is a strong tool for providing the measurements necessary to meet the needs proposed in this document. The UV spectral range from 10-400 nm has many emissions that tie directly to the constituents of the upper atmosphere [Meier, 1991], and background contamination is extremely low in most cases due to absorption of these emissions in the lower atmosphere below 100 km. Indeed, many discoveries about the composition, temperature, and dynamics across the ITM have been made using UV airglow. However, these missions have historically been single-sensor missions that, while continually providing new information about the state of the ITM, are not configured to discern temporal from spatial changes. A sensor suite unifying global and local measurements

would work together to provide the imagery necessary to provide routine space weather maps of the ITM system at temporal and spatial resolutions necessary to meet the needs of future users. There are many well-developed alternate techniques for monitoring the ITM and these data will continue to provide complementary information for comparison. However, without delving into the many strengths and weaknesses of these numerous alternate approaches, the combination of UV sensors provides the most complete, proven technique for meeting the comprehensive requirements for measurements that are needed.

First, a global UV imager would provide the means to characterize the temporal changes in the global ITM system, including its response to solar and geomagnetic activity. This sensor could be placed in geosynchronous orbit, at a Lagrangian vantage point, or even on the Moon, with the capability to provide high temporal resolution disk imaging to expose the cadence of the response of the whole ITM system, on timescales of the order of a few minutes. Fittingly, this imager will be a modern implementation of the far ultraviolet camera flown to the Moon during the Apollo 16 mission [Carruthers and Page, 1972, 1976] that demonstrated for the first time the possibility of obtaining such a global ITM image, including auroras, the ionospheric tropical arcs, and even the hydrogen geocorona. This concept was taken to the next level when DE-1 returned over a half a million images of the ultraviolet airglow and aurora, exposing many new surprising results [see Frank and Craven, 1988]. Spectroscopic techniques have advanced significantly since then but no mission has yet been designed to take advantage of this capability. To achieve the desired resolution and cadence for the next generation might require additional innovation in optics and sensor development and the data volume will be large, but the return on investment will be substantial.

Second, small UV sensors will provide essential, targeted information to further define the changes to the IT system on all scales. This is a necessary complement to provide the spatial aspects not resolved with global imaging. Much like local Doppler radar enhances satellite imagery for meteorological purposes, these smaller “field” sensors would provide the second half of this imaging suite. The ITM volume, and the coupling, has important structure in all three dimensions and the users of this information occupy the entire volume. These sensors are an essential cog to put the large-scale maps into context with the vertical structure and interaction that also defines the system. While functionally they would be similar to those already flying on current missions, it is necessary to update these current sensors to meet added demands. A handful of sensors in orbits at key local times would be required to appropriately complement the global images. Naturally a larger fleet would produce an even better data product, so with the high demand for access to space it is beneficial to take advantage of innovative optical designs that improve the possibility of incorporating a sensor into available slots on any mission. Small-scale sensors have already been developed for high-sensitivity UV measurements of the ITM system from a relatively small package [e.g. Cotton, et al., 1994, 2000]. These Limb-imaging Ionospheric and Thermospheric EUV Spectrographs (LITES) use a simple optical design with a single optical element to obtain high-sensitivity limb profiles of EUV and FUV emissions. Low-cost, light-weight, high-sensitivity sensors like LITES would be an ideal complement to the flagship global UV imager.

Application to Science and Society

The measurements obtained by this UV sensor suite would be complementary to many other areas of NASA research. The images would provide key details about the propagation of solar events once they reach the terrestrial atmosphere. Images that extend from the surface of the Sun out into the coronal region and into the near-Earth region are now routinely collected. Likewise, the lower atmosphere receives routine coverage from missions that map changes in the composition as they affect life on Earth. A similar effort has not yet been initiated for the upper atmosphere of Earth where the impacts of these events are realized. The ITM, while effectively ignored in comparison to these two other high-profile areas of study, must now be considered on par with these other solar-terrestrial regions as global lifestyles depend on satellites and space-based platforms for communication and information transfer, and space travel and tourism rapidly raises the commercial and social relevance of this region. This importance is even recognized by the American Meteorological Society that recently added a Space Weather section to its portfolio, in part because of the impact space weather has on transpolar airline flights, the ubiquitous applications for Global Positioning Satellites, and the lifetime and risk assessments for satellites and all space-based assets. The cost to replace these resources, the timeline for development of these replacements, and the inability to easily access these resources for repair, all make the ITM region an equally compelling region to add a routine monitoring capability. While either the global disk imager or the suite of smaller local sensors alone would provide significant gains toward filling this data void, the combination provides an exponential gain that cannot be matched.

References

1. Carruthers, G. R. and T. Page [1972], Apollo 16 far-ultraviolet camera/spectrograph: Earth observations, *Science*, **177**, 788.
2. Carruthers, G. R. and T. Page [1976], Apollo 16 far-ultraviolet imagery of the polar auroras, tropical airglow belts, and general airglow, *J. Geophys. Res.*, **81**, 483.
3. Cotton, D. M., T. Cook, and S. Chakrabarti [1994], A single element imaging spectrograph, *Appl. Opt.*, **33**, 1958.
4. Cotton, D. M., A. Stephan, T. Cook, J. Vickers, V. Taylor and S. Chakrabarti [2000], TESS: The tomographic EUV spectrographs, *Appl. Opt.*, **39**, 3991.
5. Frank, L. A., and J. D. Craven [1988], Imaging results from Dynamics Explorer I, *Rev. Geophys.*, **26**, 249.

6. Hagan, M. E., A. Maute, R. G. Roble, A. D. Richmond, T. J. Immel, and S. L. England [2007], Connections between deep tropical clouds and the Earth's ionosphere, *Geophys. Res. Lett.*, **34**, L20109, doi:10.1029/2007GL030142.
7. Immel, T. J., E. Sagawa, S. L. England, S. B. Henderson, M. E. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton [2006], Control of equatorial ionospheric morphology by atmospheric tides, , *Geophys. Res. Lett.*, **33**, L15108, doi:10.1029/2006GL026161.
8. Luhr, H. K., C. Hausler, and C. Stolle [2007], Longitudinal variation of F region electron density and thermospheric zonal wind caused by atmospheric tides, *Geophys. Res. Lett.*, **34**, L16102, doi:10.1029/2007GL030639.
9. Meier, R. R. [1991], Ultraviolet spectroscopy and remote sensing of the upper atmosphere, *Space Sci. Rev.*, **58**, 1-185.

Direct Observations of Global-Scale Transport in the Lower Thermosphere

White Paper submitted to the Decadal Survey
Atmosphere-Ionosphere-Magnetosphere Interactions Panel
November 2010

M.H. Stevens, NRL

R.R. Meier, GMU

S.I. Azeem and G. Crowley, Astraspace, Inc.

H.-L. Liu and J. Yue, NCAR

1. Motivation.

Water vapor is a common exhaust product of liquid fueled launch vehicles and a large portion of this exhaust is deposited into the lower thermosphere. Early observational and theoretical studies explored the possible effects of this water vapor injected into the lower thermosphere, focusing primarily on local phenomena [Debus et al., 1964; Hicks et al., 1976; Forbes, 1980; Hicks et al., 1999]. Yet over the past decade, there have been many new ground-based and satellite-borne observations of water vapor exhaust from the space shuttle and from other launch vehicles that demonstrate that the effects can be global-scale [Stevens et al., 2002; Stevens et al., 2003; Siskind et al., 2003; Stevens et al., 2005a; Stevens et al., 2005b; Meier et al., 2010a; Kelley et al., 2010; Pumphrey et al., 2010]. One of the more surprising results of these studies was that the space shuttle's main engine plume could be transported in only one to three days to the polar summer mesosphere to form large bursts of mesospheric clouds [Stevens et al., 2003; Stevens et al., 2005a; Stevens et al., 2005b; Kelley et al., 2010]. These bursts can contribute 10-20% to the polar mesospheric cloud (PMC) mass for one season. This is important because PMCs have been suggested as indicators of climate change and long-term trends since 1980 are reported to be less than 1%/year [DeLand et al., 2007].

The shuttle injects the largest amount of water vapor of any launch vehicle currently operating. On the shuttle's ascent, about 300 metric tons of water vapor are deposited into a narrow altitude region between 100-115 km. Several unresolved science questions have emerged through direct observations of this plume [Meier et al., 2010b]:

- 1) What mechanisms can account for the rapid expansion of atomic hydrogen, which grows to more 1000 km in less than a day?
- 2) How are some of the plumes transported to Polar Regions at rates exceeding wind climatologies by as much as a factor of four?
- 3) How does the injection of water vapor with mass mixing ratios approaching unity during the main engine burn affect atmospheric chemistry, heating and ionization balance?
- 4) How does the water vapor injected between 100-115 km get transported down to polar mesospheric cloud altitudes (80-85 km)?
- 5) What is the contribution of space traffic to the climatology of polar mesospheric clouds (PMCs), thought by some to be an indicator of global change in the upper atmosphere?

These remarkably diverse questions challenge our basic understanding of dynamics, chemistry and microphysics in the upper atmosphere. Although the shuttle program is ending, the U.S. space program will continue as will space traffic worldwide, offering continued insight to global-scale transport processes in the Earth's upper atmosphere. Answers to the questions above require a mission that can monitor the evolution of a concentrated water parcel at high temporal and spatial resolution. We describe such a mission below.

2. Measurement Approach.

Table 1 provides a summary of some of the shuttle plume observations, where the average maximum observed meridional wind speed is 37 m/s. Thus the reported meridional wind speeds from shuttle plume observations are typically well in excess of those reported from satellite wind climatologies or general circulation model calculations, which are at most 10-15 m/s [Stevens et al., 2005b; Liu, 2007].

Direct global-scale wind measurements in this region of the atmosphere (80-115 km) are sparse and subject to their own observing limitations. Ground-based chemical tracer experiments routinely show winds in excess of 100 m/s in the 100-115 km region of the atmosphere [Larsen, 2002]. However, these ground-based observations are necessarily limited in time and space so that the horizontal extent and temporal variability of these winds are difficult to directly quantify. Satellite wind climatologies allow for a global-scale picture of tidal variability in the lower thermosphere [e.g. McLandress, 1996]. However, many of these studies require averaging over weeks or even months so that shorter duration events and phenomena such as planetary waves cannot easily be captured in the results [Yue and Liu, 2010; Niciejewski et al., 2010].

Table 1

Shuttle	Date	Transport Time	Northward Speed	Reference
STS-66	Nov., 1994	1.1 days	26 to 40 m/s	Stevens et al. (2002)
STS-85	Aug., 1997	1.3 days	25 to 45 m/s	Stevens et al. (2003)
STS-109*	Mar., 2002	2.3 days	9 to -32 m/s	Meier et al. (2010b)
STS-110*	Apr., 2002	1.6 days	6 to -67 m/s	Meier et al. (2010b)
STS-111*	Jun., 2002	1.6 days	9 to -20 m/s	Meier et al. (2010b)
STS-112	Oct., 2002	0.4 days	7 to -44 m/s	Siskind et al. (2005)
STS-112*	Oct., 2002	1.8 days	1 to -33 m/s	Meier et al. (2010b)
STS-113	Nov., 2002	0.4 days	-45 to -60 m/s	Siskind et al. (2005)
STS-113*	Nov., 2002	1.7 days	-6 to -19 m/s	Meier et al. (2010b)
STS-107	Jan., 2003	2.2 days	-44 m/s	Stevens et al. (2005)
STS-107*	Jan., 2003	2.1 days	-7 to -47 m/s	Meier et al. (2010b)
STS-121*	Jul., 2006	1.9 days	-2 to -16 m/s	Meier et al. (2010b)
STS-115*	Sep., 2006	2.2 days	1 to -8 m/s	Meier et al. (2010b)

*Uses average elapsed time of observation closest to 2 days from launch and the typical injection latitude of 33 N.

Aside from the rapid transport speed of the shuttle plume, locating the plume has also proven to be challenging. Ground-based instruments require that the plume be transported overhead by chance [Stevens et al., 2003; Stevens et al., 2005b; Kelley et al., 2010]. Satellite-borne limb scanning instruments can be sensitive to the water vapor [Siskind et al., 2005] but cannot constrain the horizontal extent of the plume. A water vapor plume expands to a diameter of ~3 km before equilibrating to the ambient temperature and pressure [Vlasov and Grushin, 1996; Stevens et al., 2003] and rapidly diffuses horizontally and vertically after injection, allowing the plume to be imaged in the nadir. This is currently done by observing the hydrogen (H) Lyman- α feature near 1216 Å, which results from scattered solar

radiation following the dissociation of the plume water vapor [Stevens et al., 2005b; Meier et al., 2010a]. Figure 1 shows an example of the evolution of the H Lyman- α from a shuttle main engine plume three days as viewed by the Global Ultraviolet Imager on the TIMED satellite [Meier et al., 2010b].

To understand what controls the rapid global-scale transport, a coordinated experiment with uninterrupted observation of the entire Earth's disk is needed. To circumvent the current limitations outlined above we require the following:

- 1) A satellite instrument in geosynchronous orbit that can continuously image the disk at ~ 10 km horizontal resolution. This avoids the limitation of satellites in low Earth orbit, which require that observations typically be made near a launch site only once or twice a day.
- 2) Deposition of exhaust products between 80-115 km systematically throughout the diurnal cycle and throughout the year from a launch site in view of the satellite instrument. Although space traffic can produce this routinely, this could also be done artificially from a mid-latitude launch site and would not require dedicating the observations to changeable schedules of orbital launch vehicles.
- 3) Observation of a tracer within the plume. This could be H Lyman- α , which has been demonstrated to be effective in locating the shuttle plume in numerous studies already [Hicks et al., 1999; Stevens et al., 2005b; Meier et al., 2010a; Meier et al., 2010b]. However, we need not limit our analysis to H Lyman- α as many other exhaust products can be substituted and targeted in the observations. Other common exhaust products include CO₂, CO, H₂, N₂ and HCl [AIAA, 1991]. Although water vapor is the most favorable species to observe for studies of PMCs, the most viable candidate for global-scale transport may not be associated with an exhaust product. Since we are primarily interested in the ultimate fate of an exhaust plume from a fixed location, any long-lasting tracer that can be detected in the nadir between 100-115 km could suffice, including one injected by a rocket launch dedicated to this study. A trade study is needed to identify the best candidate for observation.

3. Scientific Impact

As discussed by Yue and Liu [2010], there are at least four important components to the global-scale transport of water vapor exhaust plumes: 1) The quasi two-day wave (QTDW) [see also Niciejewski et al., 2010], 2) the migrating diurnal tide, 3) the migrating semi-diurnal tide and 4) the mean meridional wind. In their TIME-GCM modeling study, Yue and Liu scaled the amplitude of the QTDW (doubled) and semi-diurnal tide (tripled) to more realistically represent the observed amplitudes of these components before running an off-line parcel advection calculation. These adjustments increased the predicted meridional motion of a water parcel to more realistically represent the observations of the STS-107 main engine plume in January, 2003.

These results, however, apply to only one case study and there is little observational data to help us determine the ultimate fate of lower thermospheric exhaust products from the shuttle or other launch vehicles. A coordinated study over at least a year as outlined in Section 2 above will have the following scientific impact:

- 1) Quantify the amplitude of the QTDW in altitude, latitude and day of year in both the zonal and meridional directions. The QTDW is known to peak near solstice so any other types of planetary waves that may peak at different times of the year are also of interest.
- 2) Quantify the amplitude of the diurnal and semi-diurnal tides in altitude, latitude and day of year for both the zonal and meridional components of the wind.
- 3) Define the horizontal and vertical evolution of a water vapor exhaust plume with time between 80-115 km altitude and determine whether this can be explained by classical molecular diffusion alone or whether some other process is controlling it, such as eddy diffusion. This in turn will help to quantify the general transport features in this region of the atmosphere, which is critical for exchange between the mesosphere and the thermosphere.

Closure on the first two questions will allow us to determine when exhaust plumes can travel to the pole and whether such conditions can routinely be met. Closure on all three of these questions would contribute significantly to a study of the chemistry within a launch vehicle plume to determine whether the local environment is more or less conducive to PMC formation than an unperturbed atmosphere. The ultimate goal of such a study is to not only understand the global-scale dynamics of the Earth's upper atmosphere but to determine whether space traffic can contribute to long-term PMC trends.

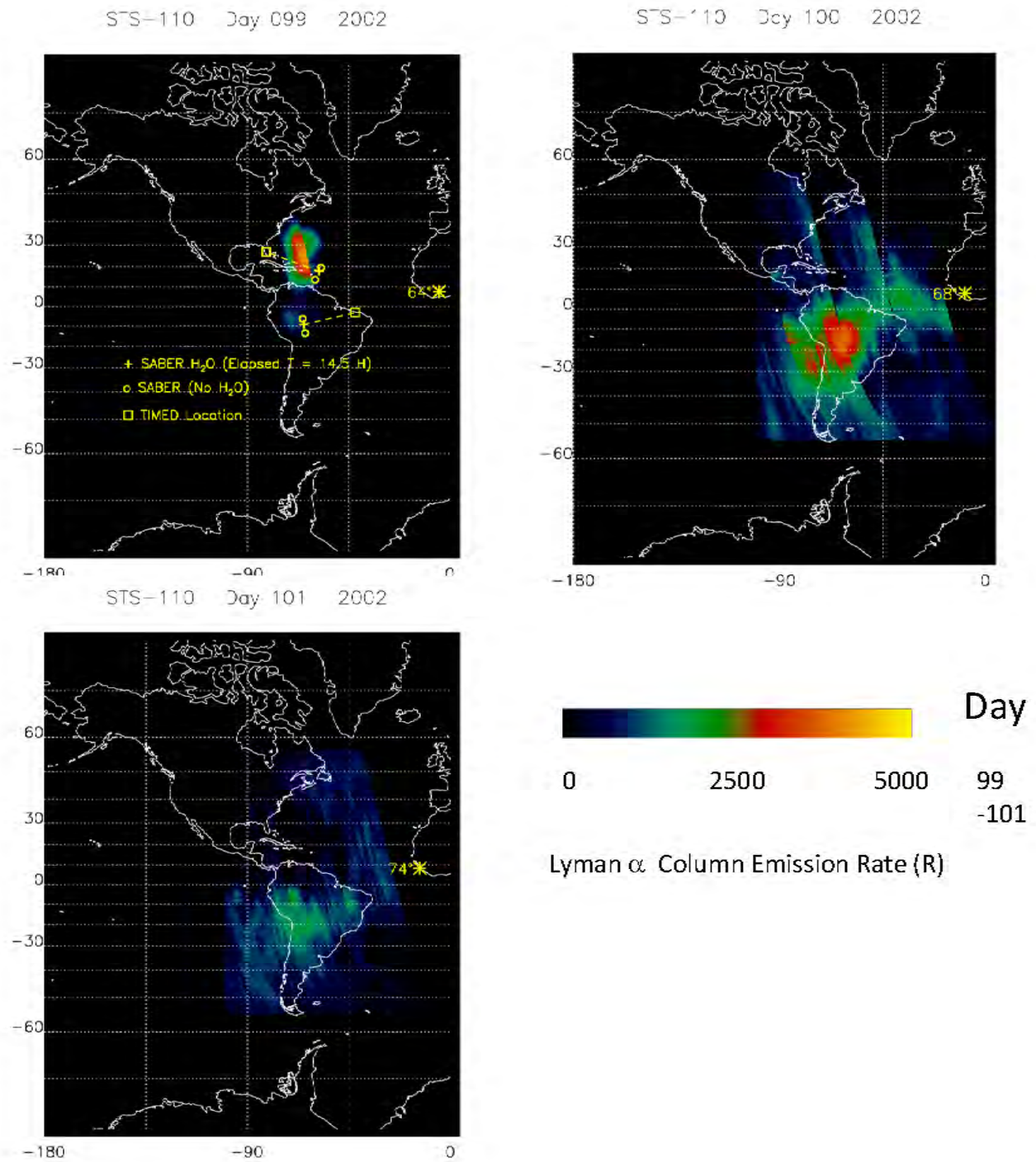


Figure 1. A three day sequence of the STS-110 main engine exhaust plume in April, 2002. The daytime observations show solar scattered H Lyman- α emission after subtraction of the bright Rayleigh scattered background. Near simultaneous water vapor observations by SABER are overplotted on day 99. The rapid southward motion across the equator implies sustained wind speeds of up to 67 m/s.

References

- American Institute of Aeronautics and Astronautics (AIAA), Atmospheric Effects of Chemical Rocket Propulsion, 52 pp., New York, 1991.
- Debus, K.M. (1964), "A preliminary review of the upper atmosphere observations made during the Saturn high water experiment", Proceedings of the XIIIth International Astronautical Congress, Springer Verlag, New York, 182-196.
- DeLand, M.T. et al. (2007), Latitude-dependent long-term variations in polar mesospheric clouds from SBUV version 3 PMC data, *J. Geophys. Res.*, 112, D10315, doi:10.1029/2006JD007857.
- Forbes, J.M. (1980), Upper atmosphere modifications due to chronic discharges of water vapor from space launch vehicle exhausts, in *Space Systems and Their Interactions with Earth's Space Environment*, edited by H.B. Garrett and C.P. Pike, *Progress in Astronautics and Aeronautics*, 71, 78-98.
- Hicks, G.T., T.A. Chubb and R.R. Meier, Far ultraviolet studies of missile trails (1976), NRL Rep. 8061, Nav. Res. Lab., Washington, DC.
- Hicks, G.T. et al. (1999), Observations of hydrogen Lyman α emission from missile trails, *J. Geophys. Res.*, 104, 10101-10109.
- Kelley, M.C. et al. (2010), Radar, lidar, and optical observations in the polar summer mesosphere shortly after a space shuttle launch, *J. Geophys. Res.*, 115, A05304, doi:10.1029/2009JA014938.
- Larsen, M.F. (2002), Winds and shears in the mesosphere and lower thermosphere: Results from four decades of chemical release wind measurements, *J. Geophys. Res.*, 107, 1215, 10.1029/2001JA000218.
- Liu, H.-L. (2007), On the large wind shear and fast meridional transport above the mesopause, *Geophys. Res. Lett.*, 34, L08815, doi:10.1029/2006GL028789.
- McLandress, C. (1996), Combined mesosphere/thermosphere winds using WINDII and HRDI data from the Upper Atmosphere Research Satellite, *J. Geophys. Res.*, 101, 10441-10453.
- Meier, R.R. et al. (2010a), Can molecular diffusion explain Space Shuttle plume spreading? (2010), *Geophys. Res. Lett.*, 37, L08101, doi:10.1029/2010GL042868.
- Meier, R.R. et al. (2010b), A study of space shuttle plumes in the lower thermosphere, in preparation for submission to *J. Geophys. Res.*
- Niciejewski, R. et al. (2010), Verification of large-scale rapid transport in the lower thermosphere: tracking the exhaust plume of STS-107 from launch to the Antarctic, Submitted to *J. Geophys. Res.*
- Pumphrey, H.C. et al. (2010), Observation of the exhaust plume from the space shuttle main engine using the Microwave Limb Sounder, *Atmos. Meas. Tech. Discuss.*, 3, 3971-3986.

- Siskind, D.E. et al., Signatures of shuttle and rocket exhaust plumes in TIMED/SABER radiance data (2005), *Geophys. Res. Lett.*, 30, 1819, doi:10.1029/2003GL017627.
- Stevens, M.H. et al. (2002), OH observations of space shuttle exhaust, *Geophys. Res. Lett.*, 29, 10.1029/2002GL015079.
- Stevens, M.H. et al. (2003), Polar mesospheric clouds formed from space shuttle exhaust, *Geophys. Res. Lett.*, 30, 1546, doi:10.1029/2003GL017249.
- Stevens, M.H. et al. (2005a), The polar mesospheric cloud mass in the Arctic summer, *J. Geophys. Res.*, 110, A02306, doi:10.1029/2004JA010566.
- Stevens, M.H. et al. (2005b), Antarctic mesospheric clouds formed from space shuttle exhaust (2005), *Geophys. Res. Lett.*, 32, L13810, doi:10.1029/2005GL023054.
- Vlasov, M.N. and V.V. Grushin (1996), A model of global distribution of the exhaust gases from rocket engines in the upper atmosphere, *Cosmic. Res.*, 34, 26-31.
- Yue, J. and H.-L. Liu, Fast meridional transport in the lower thermosphere by planetary-scale waves (2010), *J. Atm. Sol.-Terr. Phys.*, in press.

Effective Aperture Behavior on the Earth and Sun

R. A. Stoneback

University of Texas at Dallas

W. B. Hanson Center for Space Sciences

I. Summary

A model has been developed that treats regions of open field lines on the Earth and Sun as apertures (holes) in a conducting plasma. This association has been made due to the striking similarity of electromagnetic fields in the Earth's polar ionosphere and those observed in a numerical model of a circular aperture in a plane conductor. The effect of an effective aperture in the polar ionosphere is included by incorporating an impedance description of electromagnetic transmission through an aperture. Using this method, two different behaviors for Magnetosphere-Ionosphere (MI) coupling are predicted based upon the electric and magnetic aperture dipole moments. The average frequency and latitude distribution of these two modes is consistent with structured and unstructured Pc1 oscillations. Further, coupling via the magnetic dipole leads to a beat frequency with a period of 100s, consistent with the general modulation period of structured oscillations. Electric dipole behavior does not excite this beat frequency, consistent with the association to unstructured waves.

Applying the same aperture model to the Sun, effective aperture behavior is expected in coronal holes at high latitudes. The impedance of the aperture is expected in the polar chromosphere, increasing the impedance of waves traveling out towards the corona. The additional impedance of the aperture provides a good match to the impedance of the corona near a 5 minute period, leading to a peak in transmitted power across the transition region sufficient to heat the corona.

Though the effective aperture model predicts behaviors consistent with observations, a number of best guesses have been employed in the model construction. The unknowns present in the model construction can be determined by using a combination of numerical and experimental techniques. Basic properties of an aperture in the complex plasma environments around the Earth and Sun can be studied numerically, while interaction properties of an aperture can be studied in the lab using high frequency electrical techniques. Once the properties of the model have been established, further comparison to observations can be made to determine the validity of the effective aperture model. Given that this simple model offers potential solutions to long standing problems in both Earth and Solar science at a low cost, I propose that investigations of effective aperture behavior should be included in the decadal survey.

II. Science Summary

The equivalence of an aperture to the Earth's polar ionosphere is motivated by the similarity of electromagnetic fields produced by each system. Figure 1 illustrates the electric field normal to the aperture when backed by a conductor and excited by a plane wave. This distribution is highly similar to the Birkeland currents in the polar region [Korth]. The boundary of the aperture is associated with the boundary between open and closed magnetic field lines. Figure 2 depicts the electric and magnetic fields within an aperture, generally consistent with fields in the polar ionosphere. While electric fields at the boundary between open and closed field lines go to zero, the pictured electric field increases at the aperture boundary. This change is due to the different boundary conditions in each system, though this does not alter the application of the aperture impedance to polar cap [Stoneback, 2011]. Figure 3 illustrates the convection pattern obtained using the aperture electric field and assuming a magnetic field into the page. A frequency is chosen such that the electric field goes to zero at the aperture boundary,

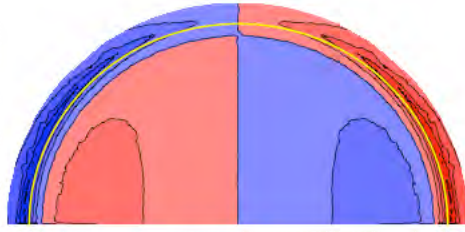


Fig. 1: Normal electric field in an aperture backed by a conductor. Yellow boundary is aperture edge. Field reversal occurs before boundary due to numerical effects. Distribution of electric field consistent with observed polar cap currents.

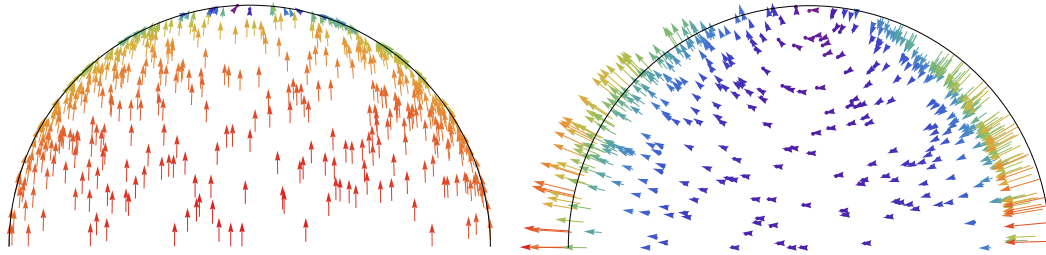


Fig. 2: Tangential magnetic (left) and electric (right) fields in an aperture at low frequencies.

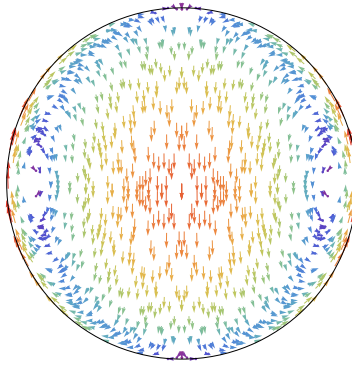


Fig. 3: Convection pattern using aperture electric field just above aperture resonance.

producing a pattern quite similar to expectations for solar quiet during southward IMF [Heelis].

The equivalence of polar cap fields to an aperture arises due to Pedersen and Hall conductivities in the ionosphere. These conductivities support currents like those employed in low frequency aperture diffraction solutions. Bethe calculated the dipole moments of an aperture for low frequency diffraction by replacing the aperture with a conductor containing magnetic currents. The conductor replacing the aperture simplifies boundary conditions while the magnetic currents recreate the fields originally produced by the aperture. Pedersen currents in the ionosphere serve as the conductor replacing the aperture while Hall currents $J_{Hall} = \nabla \times \vec{E} \times \hat{z}$ are equivalent to the magnetic currents $\vec{M} = \vec{E} \times \hat{z}$ replacing the aperture. The potential across the polar cap applied by the solar wind is functionally equivalent to a plane wave exciting an aperture, thus the Earth's polar caps are treated as apertures.

The proposed aperture behavior of the polar ionosphere can be incorporated in theoretical models by using a new dipole impedance description of an aperture [Stoneback, 2010b]. This model determines power radiated through the aperture by incorporating standard relations for waves at the interface between two media. The dipole impedance of the aperture is calculated by taking a parallel combination of electric and magnetic dipole impedances,

$$Z_{mag} = \frac{1}{2} \left(\frac{S k_b^2}{Z_b} + \frac{S}{j \omega \mu_b} \right)^{-1} \quad (1)$$

$$Z_{elec} = 2 \left(\frac{S l^2 k_b^4}{Z_b} + \frac{j \epsilon_b S l^2}{\alpha_e} \right)^{-1} \quad (2)$$

$$Z_{ap} = \frac{Z_{mag} Z_{elec}}{Z_{mag} + Z_{elec}} \quad (3)$$

where S is the aperture area, μ_b, ϵ_b are the magnetic and electric aperture polarizabilities, l is the thickness of the aperture and $Z_b, \mu_b, \epsilon_b, k_b$ are the wave impedance, permeability, permittivity and wave number in region **b**. It is presumed the wave source is in region **a** on one side of the aperture, radiating through the aperture into region **b**. Power transmitted into the aperture from the magnetic field in the incident wave and radiated is given by

$$P_b = \frac{4 \mu_a^2}{(\mu_a + \mu_b)^2} \frac{4 Z_{inc} Z_{ap}}{|Z_{inc} + Z_{ap}|^2} \frac{Z_{inc} S |\vec{H}_{inc} \cdot \vec{x}|^2}{2} \quad (4)$$

where Z_{inc} is the impedance of the incident wave and $\vec{x}$ is in the plane of the aperture.

The first fractional term in (4) accounts for boundary conditions in an aperture when bounded by different media. The second fractional grouping is the power transmitted normally across an impedance discontinuity. Due to aperture properties [Stoneback, 2010b], power can only be transmitted normally into the aperture. The reduction in radiated power for an oblique wave is accounted for in the last fractional grouping, the incident power term over the aperture area. Only the magnetic field component in the plane on the aperture can couple with the magnetic dipole, accounted for with the dot product. A similar calculation yields the power for the electric dipole term. Presuming a rotationally symmetric aperture, then the radiated power is

$$P_b = \frac{4 \epsilon_b^2}{(\epsilon_a + \epsilon_b)^2} \frac{4 Z_{inc} Z_{ap}}{|Z_{inc} + Z_{ap}|^2} \frac{Z_{inc} S |\vec{E}_{inc} \cdot \vec{n}|^2}{2} \frac{1}{4} \quad (5)$$

where $\vec{n}$ is normal to the aperture. An additional factor of 1/4 appears since the electric dipole acts as a parallel combination of magnetic dipoles [Stoneback, 2010b]. Note that the boundary condition terms at the beginning of (4) and (5) apply to power radiated by the aperture since the aperture must drive fields in both regions at once. Since the polar ionosphere does not contain an actual aperture, these boundary conditions would not normally be considered for waves upon the ionosphere. Thus, power transmitted into the polar ionosphere, even when including the aperture impedance, is presumed to not

include these material constraints. As noted, power radiated by this coupling as a dipole field is limited by these conditions. Power transmitted into the system but not radiated as part of an aperture dipole field is presumed to continue as an Alfvén wave.

The aperture impedance has been determined when surrounded by materials with constant properties [Stoneback, 2010b]. As a first order look at applications to space science, an average of plasma properties is taken in a volume surrounding the aperture to determine the aperture impedance. The thickness of this volume is taken to be the minimum thickness of an aperture [Stoneback, 2010b] with a radial size larger than the aperture radius to account for image currents around the aperture. The impedance of an aperture in a space environment can be determined directly by numerical investigations, discussed in detail later.

A. Earth Science

It is postulated that the polar aperture interacts with the cavity formed between the Earth and ionosphere. Below the Schumann resonances, this cavity is approximated here as a capacitive impedance. The interaction of the aperture and the Earth Ionosphere Cavity (EIC) is accounted for by combining the impedance of the aperture and EIC. For long incident wavelengths, the aperture impedance is inductive, thus the interaction with EIC forms an LC resonator that may be described as an electromagnetic Helmholtz resonator [Gady'l'shin; Rao; Stoneback, 2009; Stoneback, 2010a]. The resonance of this system is analogous to the sound produced when blowing over an empty wine bottle.

It is postulated that the proximity of a conductive Earth to the bottom-side ionosphere alters the presented aperture impedance. It is further presumed that the effect upon the aperture impedance is different for the electric and magnetic aperture dipoles. The supposition that the two dipoles will behave differently is based upon the underlying series and parallel impedance relationships for each dipole [Stoneback, 2010b]. Thus, it is postulated that two behaviors are possible for waves in a given region incident upon the polar ionosphere.

Figure 4(a) illustrates the predicted transmitted power for a wave in the magnetosphere incident upon the ionosphere at high latitudes, using an average polar

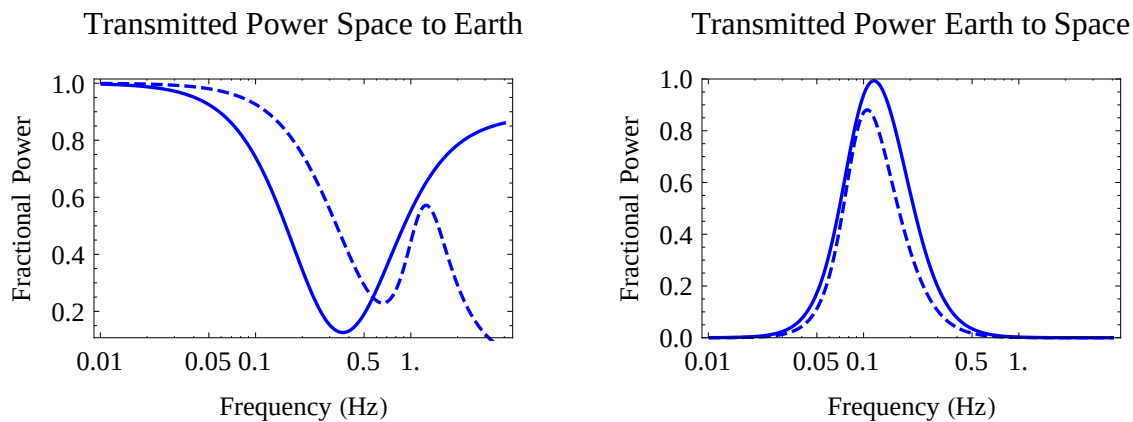


Fig. 4. Left (a): Power transmitted into ionosphere for a magnetospheric wave source. Dashed line - magnetic dipole, solid - electric dipole. Right (b): Power transmitted into ionosphere for a wave source below the ionosphere.

cap size covering 15 degrees latitude. Electric dipole mode (solid) has a minimum near 0.36 Hz while the magnetic dipole mode (dashed) has a relative minimum near 0.7 Hz and a relative maximum near 1.2 Hz. These minima in transmitted power have corresponding maxima in transmitted electric field, driving ionospheric Hall currents. The transmitted power for the electric dipole mode closely follows measurements of waves incident upon the ionosphere made by FAST [Dombeck, Fig. 3].

Figure 4(b) illustrates the transmitted power for a wave source below the ionosphere. The proposed aperture impedance is now determined by properties of the ionosphere, increasing the capacitance of the aperture above the capacitance of the EIC, lowering the frequency range compared to magnetospheric sources. Power transmitted into the EIC by the magnetic dipole will excite these two modes, separated in frequency by .01 Hz, leading to a beat period of 100s. Waves from the magnetosphere coupling through the electric dipole are excluded from exciting this EIC mode due to boundary conditions upon radiated power (5). Magnetospheric sources have $I_a \gg I_b$ (5) and no power is radiated into EIC. Thus, waves coupling with the magnetic dipole can be structured, while waves from the electric dipole mode will be unstructured.

Structured waves are Pc oscillations with a regular modulation on the order of 100s that are detected over all latitudes. The average frequency of these oscillations measured at high latitudes [Mursula] is consistent with magnetic dipole coupling along with the beat period expected for magnetic dipole coupling.

Unstructured oscillations do not display the modulation of structured oscillations and are primarily confined to high latitudes. The average frequency of unstructured oscillations is consistent with the minimum in power for electric dipole coupling [Mursula]. In addition, coupling with the electric dipole is limited to regions with field lines that have applied potentials, generally confining this behavior to high latitudes. Given that the electric dipole mode also does not excite the beat period near 100s, this mode is consistent with unstructured oscillations.

B.Solar Science

The Pedersen and Hall conductivities in the Earth's polar ionosphere are also present in the Sun's chromosphere, more accurately modeled as Hall and ambipolar diffusion [Pandey]. Using an average of the Alfvén wave speed around the Sun, the impedance of a solar aperture is calculated using a polar coronal hole radius spanning 30 degrees of latitude. The presence of Hall diffusion in the chromosphere increases the impedance of waves in this region by the aperture impedance. Using an average wave speed in the chromosphere of 125 km/s to generate an average chromosphere impedance Z_{ch} , incident upon the corona with an average Alfvén wave speed of 5.25 Mm/s and impedance Z_{co} . The average for the corona is obtained by averaging the wave speed from the photosphere out to 1.4Rsun [Pekunlu, Tomczyk], based upon the minimum effective thickness (.4 Rsun) of the effective solar aperture. Power transmitted from the chromosphere with impedance $Z_{ch} + Z_{ap}$ to the corona Z_{co} ,

$$P = \operatorname{Re} \left\{ \frac{4(Z_{ch} + Z_{ap})(Z_{co}^*)}{|Z_{ch} + Z_{ap} + Z_{co}|^2} \right\}$$
 is illustrated in Fig. 5. The aperture impedance used in Fig. 5

also includes terms here for magnetoacoustic modes based upon the coupling determined for the Earth, additional details are available [Stoneback, 2011]. The peak in transmitted power near 5 minutes for the magnetic dipole mode provides enough power to heat the solar corona. Accounting for both polar caps in both hemispheres and the fact that only half of the transmitted power can be radiated into a single region [Stoneback, 2010b], the polar apertures transmit enough power to obtain 100 W/m² when averaged over the surface area of the Sun if there is approximately 2 kW/m² in 5 minute oscillations in the polar chromosphere. Accounting for the reduced transmission away from the peak, this is generally consistent with estimates of power densities of 4-7 kW/m² in the chromosphere [Pontieu].

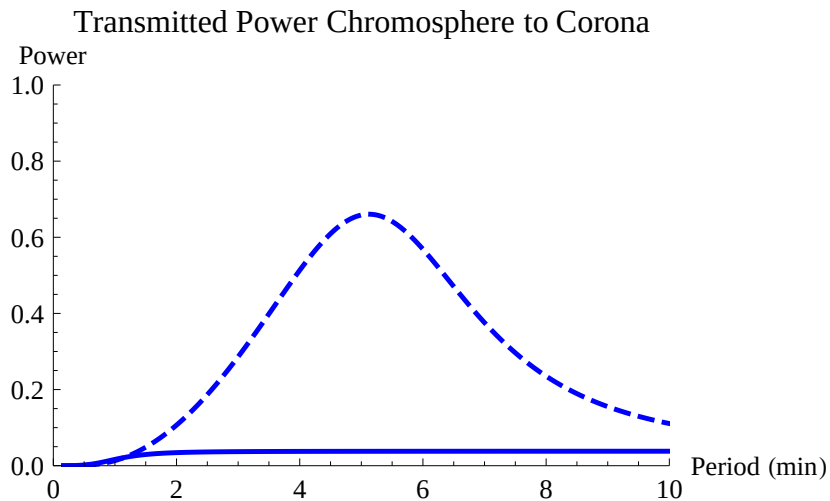


Fig. 5. Power transmitted from chromosphere to corona in the polar coronal hole presuming the chromosphere has an additional impedance due to effective aperture behavior. Dashed - Magnetic dipole mode. Solid - Electric dipole mode

III. Opportunities and Approach

A single model has been developed that predicts behaviors consistent with observations on both the Earth and Sun. This model provides a possible solution to the coronal heating issue, a long standing research problem in solar physics. Further, it may offer insight on the behavior of sunspots, another long standing topic of interest. For the Earth, the model provides a possible mechanism for magnetosphere-ionosphere coupling that produces two classes of behavior consistent with structured and unstructured oscillations. The fact that this single model operates successfully over two different scale sizes and plasma environments provides strong support that research should continue on this model.

The constructed model required some educated guess on the behavior of an aperture in a space environment. Before the aperture model can be accurately compared to observations, unknowns within the model need to be determined. For

example, the behavior of an aperture in a complicated plasma environment, the impedance of an aperture backed by a conductor, the coupling of an aperture with a capacitor, as well as the interaction of multiple apertures in a given system needs to be determined. Further, the coupling of an aperture and capacitor is expected to change depending upon whether a wave source excites the capacitor or the aperture. These properties can be determined through a combination of numerical and experimental work. Once the model is verified, comparisons can be made to observations with more confidence.

Knowledge of aperture properties in the environments around the Earth and Sun can be obtained through numerical work. The Finite Element Method (FEM) has already been used to determine the aperture impedance [Stoneback, 2010b]. Plasma distributions provided by Solar and Earth numerical models, such as TIE-GCM for the Earth, could be used to calculate the environment that surrounds the Earth. A physical aperture in a conductor will replace the polar cap and the impedance will be calculated for an incident wave. Within ANSYS, the impedance of the aperture is obtained directly from the fields within the aperture. The impedance of an aperture backed by a conductor can be obtained similarly.

Interaction properties of multiple apertures in a system also need to be investigated. The current model only considers a single polar aperture. The interaction of apertures can be investigated experimentally using high frequency measurement techniques. At sizes appropriate for the laboratory, the resonance of an aperture and capacitor is expected to occur above 1 MHz. The behavior of EM waves incident upon a parallel plate capacitor with an aperture can be precisely measured using modern high frequency vector network analyzers. Network analyzers can measure both the magnitude and phase of waves reflected from a given electric device. Measurements of this sort must be performed in an EM anechoic chamber. A suitable chamber is located in the same building as the UT Dallas W. B. Hanson Center for Space Sciences.

IV.Costs

\$50K for computer for FEM and corresponding software every 2 years, \$250k

\$50k for network analyzer and associated equipment.

\$150K/yr for 1 full-time researcher with overhead * 10 yrs = \$1.5 million

\$60k/yr for 1 full time grad student, 2 students over 10 yrs = \$600k

\$120k/yr for personnel (with overhead) to construct devices to be measured, 2 yrs = \$240k

Net cost over 10 years less than \$3 million. Two full time researchers and 4 graduate students could be supported for \$5 million over 10 years.

V. Decadal Survey Criteria

The effective aperture model has demonstrated potential for making contributions to long standing problems in Earth and solar science. The proposed topics of study are well-posed and may be completed affordably by a small team of researchers within 10 years. Tools necessary to complete this research already exist and personnel at the University of Texas at Dallas already have the requisite experience [Stoneback, 2009; 2010ab; 2011] with the hardware and software needed to investigate aperture properties and applications to space science.

VI. References

Bethe. Theory of Diffraction by Small Holes. *Physical Review* (1944) vol. 66 (7-8) pp. 163-182

Dombeck et al. Alfvén waves and Poynting flux observed simultaneously by Polar and FAST in the plasma sheet boundary layer. *Journal of Geophysical Research* (2005) vol. 110 (A12) pp. 8

Gadyl'shin, R. On acoustic Helmholtz resonator and on its electromagnetic analog. *Journal of Mathematical Physics* (1994) vol. 35 (7) pp. 3464-3481

Heelis et al. A Model of the High-Latitude Ionospheric Convection Pattern. *Journal of Geophysical Research* (1982) vol. 87 (A8) pp. 6339-6345

Korth et al. Comparison of Birkeland current observations during two magnetic cloud events with MHD simulations. *Annales Geophysicae* (2008)

Mursula et al. Properties of structured and unstructured Pc1 pulsations at high latitudes: Variation over the 21st solar cycle. *Geophysical Monograph* (1994) pp. 409-415

Pandey et al. Waves in the solar photosphere. *Monthly Notices RAS* (2008) vol. 386 (3) pp. 1635-1643

Pekunlu et al. Alfvén waves in the inner polar coronal hole. *Monthly Notices RAS* (2002) vol. 336 pp. 1195-1200

De Pontieu et al. Chromospheric Alfvénic Waves Strong Enough to Power the Solar Wind. *Science* (2007) vol. 318 (5856) pp. 1574-1577

Rao and Barakat. Plane wave scattering from a cylindrical Helmholtz resonator: Dirichlet case. *Optics Communications* (1999) vol. 164 pp. 223-232

Stoneback, R. Applications of the electromagnetic Helmholtz resonator. Dissertation, University of Texas at Dallas, May 2009

Stoneback, R. (a) Electromagnetic musical instruments. United States Patents #7,777,118 - 7,777,120; August, 2010.

Stoneback, R. (b) The dipole impedance of an aperture. *Progress in Electromagnetics Research B*, vol. 26, 401-423, 2010.

Stoneback, R. and R. Heelis. Effective aperture behavior of open magnetic field lines. Manuscript under preparation, available upon request. Expected publication 2011.

Tomczyk et al. Alfvén Waves in the Solar Corona. *Science* (2007) vol. 317 (5842) pp. 1192-1196

CUBESATS IN HELIOPHYSICS RESEARCH

Charles Swenson, Utah State University, Charles.Swenson@usu.edu

David Klumppar, University of Montana, klump@physics.montana.edu

Robert Lin, University of California Berkeley rlin@ssl.berkeley.edu

Aaron Ridley, University of Michigan, Ridley@umich.edu

Gary Swenson, University of Illinois, swenson1@illinois.edu

Miguel Larson, Clemson University, mlarsen@clemson.edu

Geoff Crowley, Astraspace, gcrowley@astraspace.net

Rick Doe, SRI, Doe@sri.com

Douglas Rowland, NASA Goddard, Douglas.E.Rowland@nasa.gov

Vassilis Angelopoulos, University of California Los Angeles, Vassilis@ucla.edu

Ennio Sanchez, SRI, Ennio.Sanchez@sri.com

Scott Palo, University of Colorado, Scott.Palo@colorado.edu

Gregory Earle, University of Texas Dallas, Earle@utdallas.edu

Chad Fish, Utah State University Space Dynamics Laboratory, Chad.Fish@sdl.usu.edu

The most significant advances in solar and space physics, or Heliophysics, over the next decade are most likely to derive from new observational techniques. The connection between advances in scientific understanding and technology has historically been demonstrated across many disciplines and time. There are clear ties between advances in our understanding of Heliophysics processes and the deployment of new sensing techniques, from new vantage points, which fuel new discoveries. Sensors such as X-ray and UV-imaging, GPS based measurements, energetic neutral atom imagers and others along with the ability to place these sensors above the atmosphere and within regions of interest have revolutionized Heliophysics science. The study of the Heliophysics system requires multipoint observations to develop understanding of the coupling between disparate regions: solar-wind, magnetosphere, ionosphere, thermosphere, mesosphere on a planetary scale. Multipoint measurements are also needed to develop understanding of the various scalars or vector field signatures (i.e gradients, divergence) that arise from coupling processes that occur across temporal and spatial scales within localized regions. Such multipoint scientific investigations have been presented in the most recent NASA Roadmap and it is likely that these and other science objectives will be expanded upon in the current decadal study process. It is clear that the Heliophysics community now needs multi-point measurements from within the space environment to make progress on important scientific questions.

Therefore significant attention must be paid to maintaining the many satellites of and the continued development of the Heliophysics System Observatory under NASA's care. In addition it is imperative that promising new technologies that will enable multipoint measurements from the space vantage point be further developed.

Remote imaging is one source of multi-point measurements of Heliophysics systems but not all measurement parameters of interest can be observed through remote sensing techniques. Some examples that are not well observed by remote imaging are electric field patterns and currents

flowing along magnetic field lines both of which are important quantities for understanding the coupling of regions. It is also clear that significant scientific advances can be made by placing remote imaging sensors at multiple points to make distributed observations of globally coherent phenomena such as atmospheric tides and auroral storms, or to improve the observations through advanced signal processing methods such as tomography or improved temporal/spatial resolutions. The development of platforms and the required infrastructure which enables both *in situ* and remote sensing from multiple points within the space environment must now be a high priority for the advancement of the science of Heliophysics.

The resources that will be available over the next decades for all areas of Heliophysics research have limits and it is therefore important that the community find ways to leverage the costs of developing new technologies to advance science. The high cost of access to space, at first review, is a serious impediment to making multipoint measurements within the space environment or in other words in deploying constellations of traditional satellites. It is therefore desirable to develop much smaller and lower-cost sensor/satellite systems such that the largest number of distributed measurements can be economically made in the space environment. The smaller the mass and volume of the sensor/satellite the larger the number will be that can be deployed from a single launch vehicle. The prospect of creating miniaturized sensors and satellite systems is good given the enormous investment of commercial, medical, and defense industries in producing highly capable, portable and low-power battery-operated consumer electronics, in-situ composition probes, and novel reconnaissance sensors. The advancement represented by these technologies has direct application in developing small sensor/satellite system for Heliophysics research. The application of portable consumer electronics technology for Heliophysics exploration must be well represented in the current decadal survey process.

Affordable constellations are not the only observational tool enabled by smaller and lower-cost sensor/satellite systems. With them it becomes feasible to put "almost disposable" platforms into heretofore sparsely or un-sampled locations or regions where it is currently not economical to place a larger more expensive satellites. Deployed into very low Earth orbit, these small low-cost platforms could carry instruments into the lower ionosphere/thermosphere, for example. The region between 100 km and about 250 km in the Earth's atmosphere is not conducive to long lifetime orbits but could be monitored nearly continuously by periodically deploying small satellites from the International Space Station, for instance. The region is currently explored in part by remote sensing methods and has been punctured by sounding rockets. Similarly in the solar chromosphere it is difficult to get closer than about 10 solar radii (e.g. Solar Probe Plus) since the harsh environmental conditions limit the satellite's lifetime. With low cost, disposable satellites, direct entry missions into the solar atmosphere could be conceived. Large missions tasked to explore comets, asteroids or planets could take along a number of small sub-satellites that could be deployed on arrival at these bodies to provide multipoint observations while communicating with the "mother ship" to relay the data to Earth. These are but a few examples of short lifetime targeted missions with potentially high scientific value. Small low-cost satellites can be placed into short lifetime trajectories lasting only a few days or a few weeks for scientific purposes which would not be feasible for larger more-expensive satellites.

The CubeSat standard for picosatellites was developed in the late 1990's for the use of the academic community in teaching space systems engineering to the next generation. It has since become widely accepted both internationally and by a broad spectrum of organizations due to the low-costs and relatively easy access to launch services which promoting the standard have engendered. The distinguishing characteristic of CubeSats is the mechanical standard for containerized launch services and how the picosatellites are opportunistically paired with those launch vehicles that provide deployment containers. The Poly Picosatellite Orbital Deployer (P-POD) developed by Cal Poly is an example of a widely accepted containerized launch system for secondary payloads. Most launch vehicles in the United States have designed support for multiple P-Pod containers which each delivering a 3 liter (10 x 10 x 30 cm) volume weighing no more than 4 kg to orbit. CubeSats have developed a significant industrial base with several aerospace companies providing subsystems to entire spacecraft buses of varying capabilities. Current estimates place the number of CubeSat developers at over 100 worldwide including governments, industry and academia. The NASA Flight Projects Office has promoted the launch of CubeSats on NASA vehicles through the Education Launch of Nanosatellite (ELaNA) program in cooperation with the NASA Office of Education. The NSF has supported the scientific development of picosatellites through the CubeSat-based Science Missions for Space Weather and Atmospheric Research program since 2008. As of Fall 2010, six NSF CubeSat missions are in development spanning remote detection and *in-situ* investigations of critical space weather state variables. Several commercial sector companies and labs have demonstrated miniaturized optical, particle, and RF sensors purpose-built for CubeSat-based research missions. CubeSat technologies, including standards based launch containers for secondary payloads, represent a significant opportunity both for developing the required technology and for achieving multipoint observations from within the space environment.

The advancement of solar and space physics, or Heliophysics, requires a skilled and trained pool of experimentalists, engineers and managers in addition to scientific modelers and theorists. The Heliophysics science community must give consideration to the organization and training of the human elements which drive scientific progress. "Doing" is the key element in developing the experimental, engineering and management skills for a scientific community. Therefore it is important to have a spectrum of experimental programs with differing costs and expectations such that these skills can be affordably developed. Small CubeSat missions are one way to develop scientific teams such that they are better prepared to manage and accomplish larger programs. It is important that these missions have strong scientific purposes and leadership such that the community fully reaps the training benefits as well as the scientific benefit of the investment. Experimental programs that are strongly driven by scientific objectives provide significantly different training experiences than those driven merely as technical demonstrations. Secondary payloads such as CubeSat represent not only an important platform for scientific investigations but also a platform to develop the experimental skills of the community, but only if these programs are strongly tied into the Heliophysics scientific agenda.

The application of CubeSats or a proposed standard for larger scale nanosats, 6 to 12 liter volume spacecraft, for multipoint measurement in support of Heliophysics science will require technology development and policy evolution. We strongly suggest that the technologies be developed with strong coupling to the Heliophysics scientific community.

- 1) Constellation deployment and maintenance: technologies and strategies need to be developed for deploying a large number of CubeSats or nanosats in a cost effective way to achieve the required constellation configuration. It is apparent that some scientific objectives can be achieved with a large constellation of essentially uncontrolled spacecraft orbits where as other studies would benefit from semi or fully controlled orbits, via propulsion capabilities, within the constellation. The development of a low-cost capability to launch or position individual CubeSats to specific LEO destinations is also desirable and would enable a number of targeted scientific studies.
- 2) Communications: In order for CubeSats or nanosats to become a significant contributor scientifically they must provide downlink capabilities similar to existing scientific satellites. The architectures for offloading data from a large number of spacecraft, frequency regulatory frameworks and ground station infrastructures must be developed.
- 3) Sensors: miniature space sensors suitable for deployment on CubeSats and appropriate for the scientific investigation need to be developed. The model developed by NASA's Planetary Instrument Development Program may be appropriate for developing scientific sensors for CubeSats or nanosats. The scientific community must learn how to make use of the commercial and defense-related technologies enabling low-power, compact systems and apply them to instrumentation for Heliophysics.
- 4) Risk Culture: CubeSat or nanosat-based constellations may achieve significant scientific breakthroughs despite failures associated with individual sensors or spacecraft due to the large number of elements that continue to function. The constellation, when considered as a whole, would therefore still be a very low risk scientific endeavor. An acceptance of increased risk tolerance for individual spacecraft within a constellation will need to emerge within NASA to fully realize the promise of low-cost CubeSat constellations.

Remote Sensing the Upper Atmosphere with Lidar from Space

by Gary Swenson

Electrical and Computer Engineering, University of Illinois, Champaign, Urbana, IL

Lidar (Light Detection And Ranging) remote sensing of the atmosphere has now been realized for tropospheric remote sensing of clouds and aerosols on the NASA CALIPSO satellite as part of the A-Train constellation (ref: CALIPSO:Link). NASA also demonstrated cloud lidar on the Space Shuttle with the LITE experiment (ref: LITE:Link). NASA has invested and developed lidar systems for tropospheric observations, successfully. Lidar technology is now ready for development and incorporation into Space Science platforms for upper atmospheric/ionospheric remote sensing.

Background: Ground based lidar studies of the upper atmosphere

The Earth's upper atmosphere (Upper stratosphere and lower mesosphere, 30-75 km) have been studied from ground based observatories with Rayleigh lidars, primarily using 532 and 355 nm transmitters. Resonance lidar has been used on metals (Na, K, Ca, Ca^+ , and Fe constituents), all of which reside in a globally distributed layer the 80-110 km altitude region as a result of meteor ablation (See Ground Based LINK to more that 50 globally distributed Rayleigh LIDARS). A study by the NSF Coupling of Energetic and Dynamic Atmospheric Regions (CEDAR) community summarizes a compilation of scientific accomplishments, effective 2005, by the CEDAR:LIDAR committee as reported in (ref: CEDAR: LIDAR). The resonance LIDARS are primarily Doppler systems, measuring winds and temperature at relatively high temporal (seconds) and spatial (fraction of a km) resolutions from ground and occasionally aircraft platforms. The systems are invaluable in the measurement and studies of mesospheric and lower thermospheric dynamics, composition, chemistry, and thermodynamics. Dynamic studies have been a primary research focus in the understanding of smallscale (gravity) buoyancy waves, tides, as well as planetary waves, and their effects on the middle and upper atmospheric region. The 80-110 km is also an atmospheric region where chemiluminescent emissions from recombination chemistry provides rich remote sensing information on wave dynamics, correlative to that provided by lidar.

Ground based lidar has been used to study the mesopause region (80-105 km) using Na resonance by the U of Illinois (C. Gardner) and Colorado State University (J. She). An example of Na resonance lidar data from ground based measurements made by the University of Illinois resonance system is shown in Fig 1, below [similar to that described by Li et al., 2007]. The Illinois system, currently operating in Chile, is described by Carlson et al. [2005]. The measurements by lidar have made important contributions to the understanding of a number of investigations including the intrinsic properties of small scale waves, tides, and planetary waves, metal densities, momentum, heat, constituent fluxes, and a host of contributions, often with correlative measurements by radars and imagers from this chemically active atmospheric region.

A demonstration of community motivation

Members of the Scientific Community have organized a number of mission plans and proposed to the highly competitive Explorer program to study coupling of

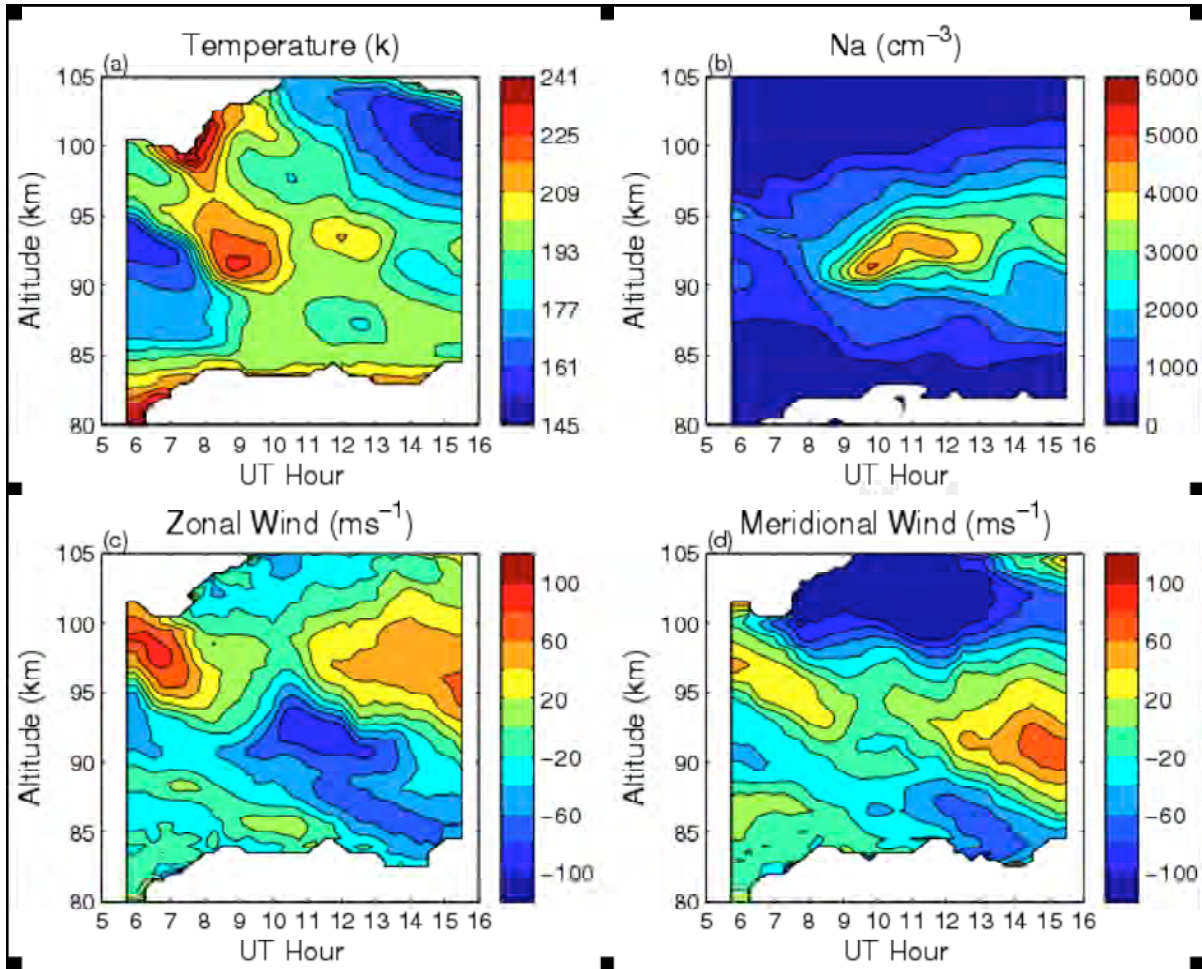


Figure 1. Time-height contours of (a) temperature, (b) Na density, (c) zonal wind and (d) meridional wind. This was taken at Maui HA, with a 1.5 W Na wind/temperature transmitter.

waves into the upper mesosphere and thermosphere from the lower atmosphere using passive remote sensing methods. These investigations from space have not been achieved to date. Among these experiments include the TIPE experiment proposed to AO No. 92-OSSA-1, (Accepted but descoped from the NASA TIMED mission due to \$), WAVES Explorers to AO-98-OSS-03 [1998] and AO 01-OSS-03 [2001], as well as Thermospheric Waves Explorer (TWEX), proposed to AO 03-OSS-02 [2003]. The problems associated with wave coupling from the lower to the upper atmosphere remain as important global atmospheric and ionospheric

problems to be solved. Ground based research is ongoing and has made progress, but global perspectives are not understood [e.g. Ground based measurements: Link].

Mesospheric Resonance LIDAR, and Na, 589 nm.

LIDAR would enable important technology to make significant enhancements to the global investigations of atmospheric coupling, upward through the mesopause boundary into the thermosphere. The ESA has performed a major study to perform metal lidars (sampling the mesopause region) from space in the past 2 years. Their community is taking this very seriously, with US technologist and technologies contributing to the ideas.

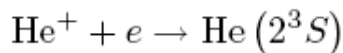
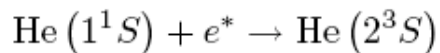
Mesospheric sampling from space would ideally be performed in the FUV where Earth backgrounds and albedo's protected by the ozone shield, although traditional visible resonance schemes using visible wavelengths can also be engineered to function, just as CALYPSO has.

Technology is rapidly advancing in the area of diode laser technologies and all of the special needs of reliability and space ruggedness for technical readiness. Solid state technologies are rapidly evolving including fiber amplifiers. A very special technology that is very exciting for the metals is **the vapor alkali, a new class of lasers** suited to Na and K lasers, recently demonstrated by Readle et al. [2009, 2010] for Cs and Rb. For Na, this technology would require a 532 nm pump laser (already in space) and a gas cell. LIDAR transmitters don't get any simpler in technology or components than this. This US innovation could put this remote sensing capability, with reasonable wall plug to transmitted power. Power efficiency is among the major considerations in orbiting active LIDAR experiments.

Thermospheric Resonance LIDAR, and He 1083 nm.

Emerging technologies are enabling thermospheric LIDAR development, including a He resonance LIDAR capable of sensing in the 250-700 km altitude region [Carlson et al., 2008]. A ground based laser has been built and demonstration on the atmosphere is in process. This will demonstrate thermospheric lidar applications for the first time.

The principles of resonance with the He lidar is the ground state helium is pumped into the 2^3S state by photoelectrons at low and mid latitude, and that state is metastable and very long lived to form a resonant population. Associative recombination of He ions also produce the metastable.



This population has been simulated initially by Gerard et al., 1997, and later by Waldrop et al., 2006, based on the analysis of Waldrop, 2005 [See Fig 2.].

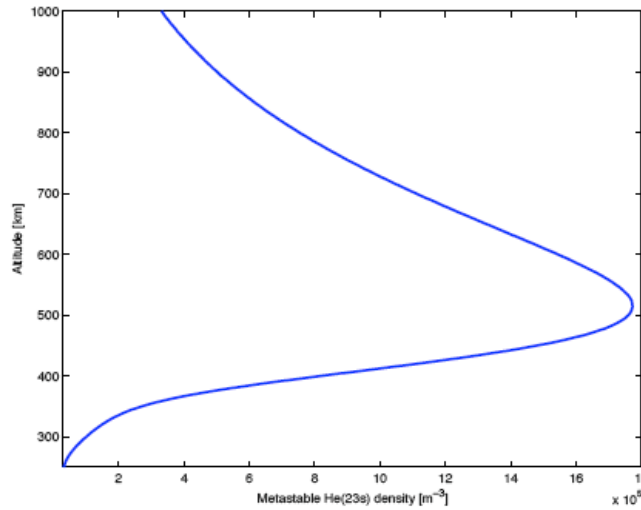


Figure 2: Metastable He(2³S) density [m⁻³] as a function of altitude for midlatitude, Northern hemisphere, solar maximum conditions with a solar zenith angle of 90°. From [Waldrop, 2006]

The He resonance lidar development employs a solid state transmitter, The laser has been built, and stages of amplification are planned to generate 50W. This technology can be further developed into a pulsed laser, with the ability to perform altitude sensing of the thermosphere between 300 and 800 km. 3-frequency tuning of the transmitter enables Doppler receiving, for measurements of altitude profiles of winds and temperature in the thermosphere, at spatial scales to sample waves and structure, not here-to-fore able to be measured. Sampling vertical resolutions on the order of km, at time scales of 10's of seconds, can be accomplished with a relatively low power system (few watts, average power). He metastable populations described in Fig 2 have been used to describe S/N of > 20 from ground based lidar [Carlson, 2008].

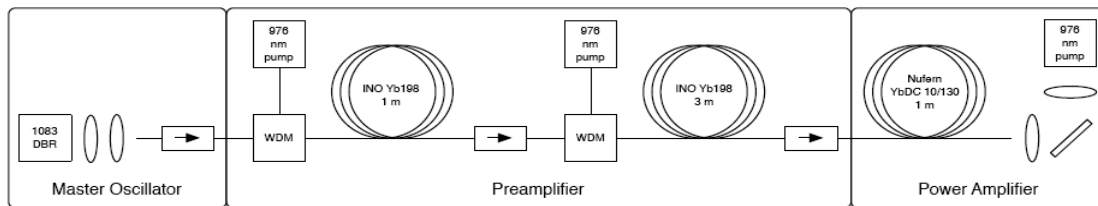


Figure 3. A solid state laser has been developed to generate 10 W of CW, at 1083 nm for ground based lidar testing using bistatic imaging of thermospheric resonance.

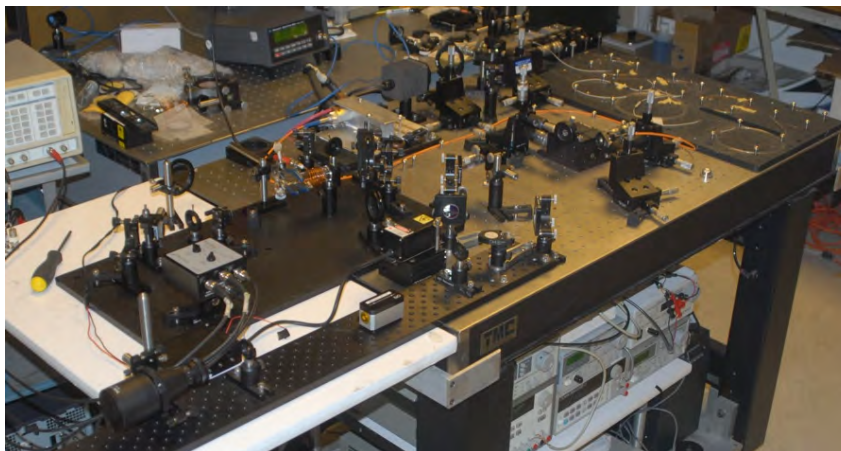


Figure 4. An image of serial number 1, of a He 1083 resonance laser on a tabletop breadboard assembly at the U of Illinois, Remote Sensing and Space Science group.

Summary

Key lidar technologies in resonance lidar with metals and He are enabling a new dimension of upper atmospheric remote sensing, not unlike the lower atmosphere community is experiencing from CALYPSO. Lidar designed to study upper atmospheric enables a new dimension of spatial and temporal sampling of density, temperature, and wind. It is timely to develop simulations and carefully evolve the hardware which is practical to employ in space to investigate our upper atmosphere.

Acknowledgement: The current lidar systems described were supported by funding from the NSF, Geospace

References

CALIPSO LINK: <http://www-calipso.larc.nasa.gov/>

Carlson, C. G., P. D. Dragic, B. W. Graf, R. K. Price, J. J. Coleman, and G. R. Swenson, "High power Yb-doped fiber laser-based LIDAR for space weather," SPIE, Vol. 6873, 68730K, 2008.

Carlson, C. G., P. D. Dragic, L. Waldrop, and G. R. Swenson, A He resonance fluorescence Doppler lidar for thermospheric investigations from 300-700 km, 24th International Laser Radar Conference, Boulder, CO, June 23-27, 2008.

Carlson, C. G., G. R. Swenson, P. Dragic, and A. Liu, Maui mesosphere and lower thermosphere (MALT) Na lidar-enterprise, SPIE, Sep, 2005.

CEDAR: Lidar beyond Phase III, Accomplishments, Requirements, and Goals, with 241 publications, March, 2004

<http://cedarweb.hao.ucar.edu/wiki/index.php/Community:Documents>

Gerrard, A. J., T. J. Kane, D. D. Meisel, J. P. Thayer, and R. B. Kerr, "Investigation of a

resonance lidar for measurement of thermospheric metastable helium,” J. Atmos. Solar-Terr. Phys. 59(16), pp. 2023–2035, 1997.

Ground Based Instruments LINK:

http://cedarweb.hao.ucar.edu/wiki/index.php/Documents:Ground_Based_Instruments

Li, F. G. R. Swenson, A. Z. Liu, M. Taylor, and Y. Zhao, Investigation of a "wall" wave event, J. Geophys. Res., 112, D04104, doi:04110.01029/02006JD007213, 2007.

LITE: <http://www.nasa.gov/centers/langley/news/factsheets/LITE.html>

Readle, J. D., J. T. Verdeyen, J. G. Eden, S. J. Davis, K. L. Gabally-Kinney, W. T. Rawlins, W. J. Kessler, Cs 894.3 nm laser pumped by photoassociation of Cs-Kr pairs: excitation of the Cs D2 blue and red satellites, Opt. Lett., 34, 3638-3640, 2009.

Readle, J. D., C. J. Wagner, J. T. Verdeyen, T. M. Spinka, D. L. Carroll, and J. G. Eden, Excimer-pumped alkali vapor lasers: A new class of photoassociation lasers, Proc. Of SPIE, Vol. 7851, High Energy/Average Power Lasers and Intense Beam Applications IVdoi:10.1117/12.842496, 2010.

Swenson, G., P. Dragic, A. Z. Liu, L. Waldrop, and J. M. C. Plane (2005), Lidar and the Mars upper atmosphere, paper presented at American Geophysical Union Fall Meeting, San Francisco, CA, Dec. 5-9, 2005.

Swenson, G. R., P. Dragic, L. Waldrop, J. M. C. Plane, C. Carlson, and A. Z. Liu (2006), Mars and earth upper atmosphere sensing simulations of Rayleigh and Na resonance lidar, paper presented at 23rd International Laser Radar Conference, Nara, Japan.

Waldrop, L. S., R. B. Kerr, S. A. Gonzalez, M. P. Sulzer, J. Noto, and F. Kamalabadi. Generation of metastable helium and the 1083nm emission in the upper thermosphere. J. Geophys. Res., 110(A08304), 2005.

Waldrop, L. S. Private Communication, 2006.

Electrodynamics Observations with Numerous Satellites

E. R. Talaat, J.-H. Yee, L. Paxton, D. Morrison, B. J. Anderson, T. Sotirelis
The Johns Hopkins University Applied Physics Laboratory

R. Heelis, P. Anderson
University of Texas, Dallas

A. Mannucci
Jet Propulsion Laboratory

T. Fuller-Rowell
University of Colorado

Dave Hysell
Cornell University

Summary

How the neutral atmosphere interacts with the ionosphere is not only fundamental to understanding the mechanisms of drivers and effects in the Earth's atmosphere as well as all other planetary and stellar atmospheres. Detailed understanding of these processes are still lacking because of limitations of current and past observations. While significant advances have occurred in measuring the upper atmosphere and ionosphere over the last three decades, the science returns from these measurements have come as either snapshots in time that are localized in space or as global climatologies that are low in resolution in both space and time. Furthermore, comprehensive, simultaneous measurements of critical parameters necessary for science closure are missing. As such, the science enabled has been limited by the measurement capabilities or heavily leveraged by global models for interpretation and compromised by the inherent uncertainties and approximations that are implicit within. As a result key science questions related to ion-neutral coupling are impossible to address with current groundbased and spaceborne assets. A suite of small satellites distributed in local time at F-region altitudes with essential instrumentation would be able address ion-neutral coupling and the roles of the disturbance dynamo, tidal dynamos and magnetospheric penetration electric fields in determining the global electrodynamics.

Background

While both ground-based and spaceborne efforts over the last three decades have greatly expanded our knowledge base in the ionosphere thermosphere region, key issues still remain unresolved due to lack of contextual and coincident measurements. The region above 150 km has been sampled in-situ by the AE and DE series of satellites, even though these satellites had insufficient instrumentation as well as sporadic coverage. Further up, in the region above 200 km, remote sensing instruments have measured in detail various auroral and airglow features, particularly in the spectral region between 120 and 750 nm to obtain information about neutral and ion composition (for instance, TIMED). However, these spacecraft

have lacked the complete instrumentation and spatial/temporal coverage/resolution to fully address fundamental processes in terrestrial ion-neutral coupling: the disturbance dynamo that occurs as the whole upper atmosphere and ionosphere composition and circulation is modified through Joule heating caused by the imposition of magnetosphere currents and particles at high latitudes; the tidal dynamo in which both in-situ solar-forced atmospheric tides and propagating tides from the lower atmosphere significantly modify electrodynamics, particularly at low and mid-latitudes; and penetration electric fields from high to low latitudes that occur during geomagnetic storms. During disturbed or “storm” times our predictions reveal that our basic understanding is flawed: similar storm conditions produce wildly different responses in winds, temperatures and densities – and we don’t understand why. The underlying causes of the observed variability simply cannot be determined because of lack of fundamental measurements. With increasing human activity in space, accurate modeling and prediction of the behavior of the thermosphere-ionosphere system in relation to Space Weather is of increasing significance, yet its effects cannot be modeled adequately. Specifically, to understand these processes simultaneous measurements of ion and neutral density and temperature, ion drift (E-field), and neutral wind are required globally.

Enabled Science Focus

A suite of small satellites distributed in local time at F-region altitudes would address ion-neutral coupling and the roles of the disturbance dynamo, tidal dynamos and magnetospheric penetration electric fields in determining the global electrodynamics. The Electrodynamics Observations with Numerous Satellites (EONS) mission has been optimized to address these outstanding science questions.

In order to address fundamental issues in ion-neutral coupling, EONS has three goals:

- 1) Determine how ions and neutrals are coupled in the upper atmosphere
- 2) Determine the structure of the interaction region and test fundamental simplifying assumptions of the physics
- 3) Build the first climatology of the spatial-temporal variability of ion-neutral coupling

Goal 1) The components of the ionosphere/thermosphere system are coupled to each other through their relative motion as well as externally. Exactly how this coupling occurs is not known. EONS will determine by direct measurement whether the ions are tied to the neutral through ohmic dissipation or through collisional drag.

Ohmic dissipation is the redistribution of energy and momentum through the loss of currents passing through a resistive medium (in this case, the ionosphere). Collisional drag arises from the difference in the velocities of the ions and the neutrals. EONS will measure all the key parameters necessary to determine which

mechanism dominates the transfer of energy from the solar wind/magnetosphere at mid- and high-latitudes.

Goal 2) In order to solve the differential equations that describe the energy, momentum and mass continuity equations of the ionosphere and thermosphere, approximations and linearizations are made. These models all predict that the upper atmosphere is isothermal above about 200km, the wind velocity fields are uniform with altitude as is the electric field. EONS will verify that the physics of the collisional processes that couple the ions and neutrals are understood.

Goal 3) We lack a systematic description of the temporal and spatial variability of the magnitude of ion-neutral coupling. This is a fundamental model output reflected in the ExB force that redistributes ions in the equatorial region. An adequate first-principles model would have to be able to demonstrate that it could reproduce this behavior. Failure to do so indicates failure of the model.

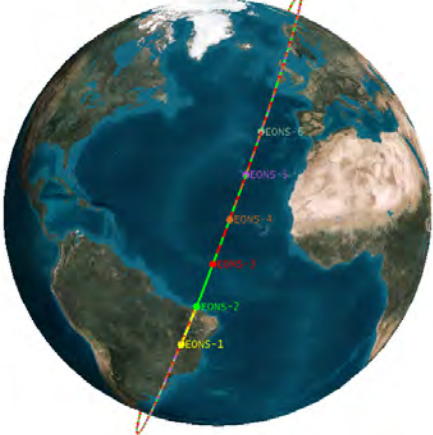
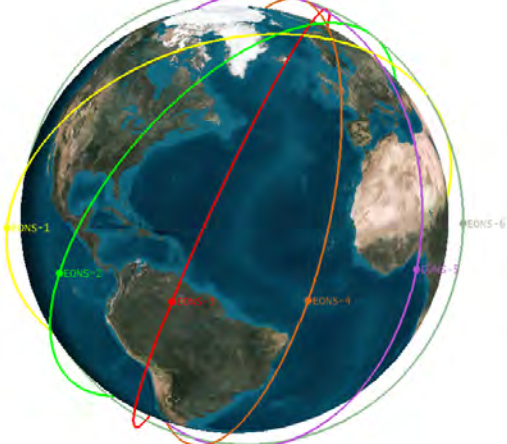
Mission Implementation

EONS instrument complement and the parameters measured include

Instrument	Measured Parameters	Heritage
Accelerometer	Neutral density, cross-track winds	CHAMP GRACE <i>SWARM</i>
GPS Receiver	Atm. Drag, Orbital Determination, Electron Density Profiles, TEC	CHAMP SAC-C COSMIC
Neutral Mass Spectrometer	Neutral composition, temperature	<i>Bepi- Colombo</i>
Retarding Potential Analyzer/Ion Drift Meter/Langmuir Probe	Ion drifts, temp. & density Electron temp. & density	DE, DMSP C/NOFS <i>SWARM</i>
Magnetometer	Currents	Many <i>SWARM</i>
Ram Wind Sensor	Along-track neutral winds	<i>C/NOFS</i>

Airglow Imager/Photometer Suite	NO, N(⁴ S), Temp., H (Lyman α)	<i>TIMED</i>

EONS can be implemented in three different mission phases, each one specifically designed to address key aspects of ion-neutral coupling.

	Mission Phase 1 – Pearls on a String Duration: ~6 months Specific Science Goals: •Penetration Electric Fields •Disturbance Dynamo Features: •High temporal cadence •High spatial resolution in orbit plane
	Mission Phase 2 – Slowly phasing Duration: ~18 months Specific Science Goals: •Penetration Electric Fields •Disturbance Dynamo •Tidal Dynamo (later half) Features: •High temporal cadence in 1 orbit plane •High spatial resolution in 1 orbit plane •Distributed in local time •Differential altitude studies for five ~1 month period after drop of each satellite
	Mission Phase 3 – Distributed in local time Duration: ~12 months Specific Science Goals: •Tidal Dynamo •Disturbance Dynamo •Penetration Electric Fields Features: •Diurnal coverage in local time •Global mapping of response of ionosphere and thermosphere

On understanding the origin of Plasma Density Variability within the polar cap

The Polar ionospheric regions encompass both the auroral oval and the polar cap regions. In these regions, the unique topology of the magnetic field lines, extending almost vertically upward, creates a variety of phenomena that are not encountered in other parts of the Earth. This document describes two ionospheric processes that contribute to the high variability of the plasma density within the polar cap. These projects are outlined and justified.

The polar region is commonly populated by macroscale (hundreds of km) and mesoscale (km and smaller) structures whose formation is not completely understood. One type of large-scale structures consists of the **polar cap patches**. These structures are the framework of km and sub-km scale irregularities that are responsible for the degradation of satellite communications. The polar cap patches consist of plasma densities enhanced by a factor between 2 and 10 that drift across the polar cap from the dayside to midnight. They cover an area of 1000 km², and extend between 100 and 1000 km altitude. In the last 20 years, we have identified up to eight mechanisms that can generate polar cap patches [Sojka, *et al.*, 1993, Rodger, 1994, Valladares *et al.*, 1994a, 1998, 1999, Lockwood and Carlson, 1992, Carlson *et al.*, 2002, MacDougall and Jayachandran, 2007, Bust and Crowley 2007]. Some of the patch formation mechanisms consist of cusp-related processes that are able to produce small-scale (km) irregularities at the time patches are formed. Because cusp processes respond rapidly to the variability of the IMF and solar wind conditions, it is expected that the IMF can control, indirectly, the patch formation and the initial density (and velocity) structuring. Some of the formation mechanisms also imply conjugate interplay, and simultaneous formation of patches in opposite hemispheres. Other formation mechanisms require a fragmentation of the TOI to form the discrete entities. Based on the large number of patch formation mechanisms, the question that we need to answer is: **Which one is the dominant patch formation mechanism and what is the level of structuring that is originated by this mechanism?**

We are still uncertain of the dominant mechanism, if a single one really exists, or if different mechanisms are more effective at different times, at different locations, or even in different hemispheres. Another intriguing and still unexplained feature of the polar cap patches is the asymmetric conjugacy of the seasonal distribution of patches in both hemispheres. Coley and Heelis [1998] introduced data from the DE satellite showing that in the northern hemisphere (NH) the occurrence probability maximizes during the December solstice in a broad time interval between 12 and 22 h UT, when 4-5 patches are seen per hour. This time interval of patch occurrence is consistent with previous ground-based observations and the time the cusp is near its most dayward location. For the southern hemisphere (SH) Coley and Heelis [1998] indicated that the patches occur again in the local winter (June) solstice, with a sharp maximum at about 1230 UT, and another smaller broad maximum around 1900-2200 UT. Note that there exists an important discrepancy with the NH. In the SH, the cusp is at its most sunlit location at about 0300 UT, a time of minimum patch occurrence. It is also mentioned that the number of patches in the southern hemisphere is 10 per hour; this is twice as large as the occurrence seen in the northern hemisphere.

Presently, continuous plasma density measurements are conducted at Resolute Bay, Qaanaaq and Nord (Polar cap stations), and at Sondrestrom, and Svalbard (polar cap/auroral oval stations) using imagers, digisondes, and scintillation receivers. Similar instrumentation should be placed in the southern hemisphere at or near the center of the polar cap (e.g. Vostok Station). The SuperDARN network consists of several coherent radars operating in the Arctic and Antarctic and able to measure the drift velocities, which gives an indication of the formation mechanism and the level of plasma structuring. About 100 GPS receivers are placed along the Coasts of Greenland and Antarctica to measure the evolution of TEC enhancements as the patches transit across the polar caps. Satellites in polar orbits, containing instrumentation similar to DMSP will be required to provide precise, but localized measurements of density and electric fields.

It is known that the Earth's tilt angle produces a strong seasonal distribution of the density over the polar caps and almost always eliminates the possibility of simultaneous formation of patches in both hemispheres during the solstices. However, during major storms, when storm enhanced densities (SED) penetrate both polar caps, patches will be generated and will traverse across both polar caps in spite of the

seasonal differences between the hemispheres. Recently, a new incoherent scatter radar has been installed at Resolute Bay. This AMISR radar augments the capability to probe the polar cap densities with 2 other IS radars that operate near the polar cap-auroral oval boundary (Sondrestrom, Svalbard EISCAT). These radars do not operate continuously, but should be able to run in campaign basis and give the evolution of the density and velocities associated with the polar cap patches. The result of these investigations will lead us to an understanding on how the patches are formed in different hemispheres and to develop procedures to forecast the occurrence and evolution of the plasma structuring within polar cap patches.

Other types of structures commonly seen in the polar cap are the **polar cap (Sun-aligned) arcs** and the transpolar auroras (theta aurora). These arcs have been mainly observed to develop when the IMF B_z component is directed northward. The polar cap auroras probably map to the magnetotail; therefore their observations can provide significant insights into the evolution of the magnetotail as the solar wind parameters change continuously. However, basic investigations need to be done because our knowledge on the interaction of the polar cap ionosphere and the magnetotail is limited. It is necessary to define **how the magnetotail topology in some cases allows the development of transpolar arcs and in other instances favors the occurrence of the weak, but numerous F-region polar cap arcs.**

Several conceptual models have been proposed [Reiff and Burch, 1985; Kullen, 2000] to explain the formation of transpolar arcs. These models make predictions about the IMF control of the conjugacy of the theta aurora, its location with respect to the polar cap cells, the current system associated with the aurora, and the movement and temporal evolution of the arcs. There are two models: the Reiff and Burch [1985] model (called here the reclosure cell model, RCM) and the twisted magnetotail model (TMM). The RCM model uses the hypothesis of anti-parallel merging between northward-directed interplanetary field lines and dayside closed and open tail lobe field lines merging at locations poleward of the cusp. The RCM model indicates that if closed field lines are merged with northward IMF containing $B_z > |B_y|$, a pair of field lines, one connected to the NH and the other to the SH are opened. The RCM model indicates that the direction of motion of the transpolar arc will be in the direction of B_y in the northern hemisphere. The arc motion is due to the dayside merging process that brings newly opened flux tubes into the polar cap [Valladares *et al.*, 1994b], displacing not only the transpolar arc but the whole reclosure cell as one entity. The location of the transpolar arc is non-symmetric between hemispheres. For a negative value of B_y the NH transpolar arc will be embedded in the dusk cell, but in the SH, it will reside within the dawn cell.

The TMM model is based on the observational fact that a nonzero IMF B_y is associated with a twist of the tail plasma sheet [Kaymaz *et al.*, 1994]. These authors indicated that the sense of tail rotation depends on the sign of B_y . Viewing from the tail toward the Earth, the magnetotail twists clockwise for positive B_y and counter-clockwise for negative B_y . The TMM model relies on an IMF B_y sign change to produce a double rotation of the magnetotail, creating in this way a tail twisted in two opposite directions. According to the TMM model, the formation and evolution of the transpolar arc occurs following a unique sequence of events. Initially, if B_y is nonzero and steady for tens of minutes, the whole plasma sheet becomes twisted in one direction, mapping to the ionosphere as a widened and poleward-expanded auroral oval. The B_y sign flip affects first the near-Earth tail producing a separation of a bar of closed field lines from the expanded oval. This separation is produced by closed field lines originating in the highly twisted far-tail plasma sheet that bifurcate the near-Earth tail lobes and the polar cap. As the new twisting propagates to the far tail, the closed field line region will move poleward. Eventually, the closed region will reach the opposite side of the auroral oval. The TMM model argues that the closed field line region is the location where the transpolar arc occurs. MHD simulations of the TMM model [Kullen and Janhunen, 2004] have in fact corroborated these arguments. While the direction of the arc motion is identical to the predictions of the RCM model, the simultaneous appearance of transpolar auroras in opposite hemisphere is quite different to the RCM model that postulates a conjugacy dependent on the B_x sign.

The RCM and TMM models make different predictions on the fashion that the polar cap cells evolve in time. Utilizing satellite data and the Royal Institute of technology (KTH) model of potential patterns,

Cumnock and Blomberg [2004] demonstrated that during the transit of the transpolar arc, a new cell develops between the newly separated arc and the poleward boundary of the auroral oval. This cell possesses a region of sunward flow adjacent to the sunward flow region that contains the transpolar arc. As the transpolar arc moves closer to the other side of the polar cap, the new cell expands and fills a substantial part of the polar cap. About the currents associated with the transpolar arc, the RCM model stipulates that a field-aligned current will be located in the dawnward side of the sunward flow region. The TMM model describes a current configuration associated with a transpolar arc having a pair of upward and downward field-aligned currents, with the upward current on the dawnside.

The preceding paragraphs suggest to conduct very precise measurements of the aurora at the moment that the arc separates from the oval, and to follow the arc evolution during the early part of its transit across the polar cap. They also imply to simultaneously observe the variability of the polar cap convection pattern and to determine the boundaries of the cell in which the arc is embedded. To observe the aurora the best instrument is a ground-based wavelength-tunable all-sky imaging photometer (ASIP) capable to record the 630.0, 557.7 and 427.8 nm lines. Valladares and Carlson, [1991] used the imager located in Sondrestrom, Greenland to show that a single ground-based imager is able to identify transpolar arcs if simultaneous observations of the red and blue lines are obtained. In addition to the Sondrestrom imager, there are many other imagers covering the polar cap-oval boundary such as the ones in Alaska, Canada, Greenland and Svalbard. It is proposed to augment the existing network of ground-based imagers by deploying imagers at Franz Joseph land and in the Severnaya Zemlya archipelago, both located in the Russian Arctic region. These two additional imagers will give the ability to observe the arcs when they separate from the auroral oval between 100° W and 170° E longitude.

Estimation of the drift velocity within the polar cap and the polar cap-oval boundary can be carried out with the SuperDARN coherent radars. There are also three incoherent scatter radars, located at Resolute Bay, Sondrestrom and Svalbard. The last 2 radars, placed near 75° magnetic latitude, are quite optimum to probe the area where the transpolar arcs separate from the oval.

Imagers on-board satellites have been used to resolve transpolar auroras [Frank *et al.*, 1982, 1986]. But they do not measure the polar cap arcs with the spatial resolution and sensitivity that is necessary to resolve thin arcs. However, it is undeniable that satellites orbiting over different polar caps will be instrumental to study the conjugacy of the polar cap arcs. Conjugate observations can firmly resolve the prevailing formation model of transpolar arcs. They have demonstrated in the past that transpolar arcs occur in opposite hemispheres and move in opposite directions [Mizera *et al.*, 1987; Obara *et al.*, 1988; Craven *et al.*, 1991]. This result is in accord with both models of transpolar arcs described above. Unfortunately, solar illumination and differences in ionospheric conductivity between the hemispheres make the conjugate observations very limited and restricted to a few weeks a year.

It is possible that the appearance of polar cap arcs is associated with the development of localized and/or transient cells that may not be strictly related to variability in the IMF or other solar wind parameters, but instead to the occurrence of wavelike or instability processes in the magnetotail. Images from this network of ASIPs, used in conjunction with data collected by satellites orbiting in the magnetosphere, should be able to study this hypothesis.

The instrumentation described here will also make it possible to study other plasma structuring processes that develop within the polar cap such as the polar hole and the plasma outflows.

References

- Basu, S., Su. Basu, J.J. Sojka, R.W. Schunk, and E. MacKenzie, Macroscale modeling and mesoscale observations of plasma density structures in the polar cap, *Geophys. Res. Lett.*, 22, 881, 1995.
- Basu, Su. and C. E. Valladares, Global aspects of plasma structures, *J. of Atmos. and Solar-Terr. Phys.*, 61, 127-139, 1999.
- Carlson, H.C., V.B. Wickwar, E.J. Weber, J. Buchau, J.G. Moore, and W. Whiting, Plasma characteristics of polar cap F-layer arcs, *Geophys. Res. Lett.*, 11, 895, 1984.
- Carlson, H.C., et al., ESR mapping of polar cap patches in the dark cusp, *Geophys. Res. Lett.*, 29 (10), 1386, doi:10.1029/2001GL014087, 2002.

- Coley, W.R., and R.A. Heelis, Seasonal and universal time distributions of patches in the northern and southern polar caps, *J. Geophys. Res.*, 103, 29229, 1998.
- Craven, J. D., J. S. Murphree, L. A. Frank, and L. L. Cogger, Simultaneous optical observations of transpolar arcs in the two polar caps, *Geophys. Res. Lett.*, 18, 2297, 1991.
- Cumnock, J.A., and L.G. Blomberg, Transpolar arc evolution and associated potential patterns, *Ann. Geophys.*, 22, 2004.
- Frank, L.A., J.D. Craven, J.L. Burch, and J.D. Winningham, Polar views of the Earth's aurora with Dynamic Explorer, *Geophys. Res. Lett.*, 9, 1001, 1982.
- Frank L.A., et al., The theta aurora, *J. Geophys. Res.*, 91, 3177, 1986.
- Hones, E.W., Jr., J.D. Craven, L.A. Frank, D.S. Evans, and P.T. Newell, The horse collar aurora: A frequent pattern of the aurora in quiet times, *Geophys. Res. Lett.*, 16, 37, 1989.
- Kaymaz, Z., G.L. Siscoe, J.G. Luhmann, R.P. Lepping, and C.T. Russell, Interplanetary magnetic field control of magnetotail magnetic field geometry: IMP-8 observations, *J. Geophys. Res.*, 99, 11113, 1994.
- Kullen, A., The connection between transpolar arcs and magnetotail rotation, *Geophys. Res. Lett.*, 27, 73, 2000.
- Kullen, A., and Janhunen, Relation of polar auroral arcs to magnetotail twisting and IMF rotation: A systematic MHD simulation study, *Ann. Geophys.*, 22, 1, 2004.
- Mizera, P. F., D. J. Gorney, and D. S. Evans, On the conjugacy of the aurora: high and low latitudes, *Geophys. Res. Lett.*, 14, 190, 1987.
- Obara, T., M. Kitayama, T. Mukai, N. Kaya, J. S. Murphree, and L. L. Cogger, Simultaneous observations of Sun-aligned polar cap arcs in both hemispheres by EXOS-C and Viking, *Geophys. Res. Lett.*, 15, 713, 1988.
- Reiff, P.H., and J.L. Burch, IMF B_y -dependent dayside plasma flow and Birkeland currents in the dayside magnetosphere, 2, A global model from northward and southward IMF, *J. Geophys. Res.*, 90, 1595, 1985.
- Rodger, A.S., M. Pinnock, J.R. Dudeney, K.B. Baker, and R.A. Greenwald, A new mechanism for polar patch formation, *J. Geophys. Res.*, 99, 6425, 1994.
- Sojka, J.J., M.D. Bowline, R.W. Schunk, D.T. Decker, C.E. Valladares, R. Sheehan, D.N. Anderson, and R.A. Heelis, Modeling polar cap F region patches using time varying convection, *Geophys. Res. Lett.*, 20, 1783, 1993.
- Valladares, C.E., and H.C. Carlson, Jr., The electrodynamic, thermal, and energetic character of intense Sun-aligned arcs in the polar cap, *J. Geophys. Res.*, 96, 1379, 1991.
- Valladares, C.E., Su. Basu, J. Buchau, and E. Friis-Christensen, Experimental evidence for the formation and entry of patches into the polar cap, *Radio Sci.*, 29, 167, 1994a.
- Valladares, C.E., H.C. Carlson, Jr., and K. Fukui, Interplanetary magnetic field dependency of stable sun-aligned polar cap arcs, *J. Geophys. Res.*, 99, 6247, 1994b.
- Valladares, C.E., D.T. Decker, R. Sheehan, D.N. Anderson, T. Bullett, and B.W. Reinisch, Formation of polar cap patches associated with north-to-south transitions of the interplanetary magnetic field, *J. Geophys. Res.*, 103, 14657, 1998.
- Valladares C.E., D. Alcayde, J.V. Rodriguez, J.M. Ruohoniemi, A.P. van Eyken, Observations of plasma density structures in association with the passage of traveling convection vortices and the occurrence of large plasma jets, *Ann. Geophysicae*, 17, 1020, 1999.

Effects of Large Amplitude Planetary Waves in the Ionosphere and Thermosphere

R.L. Walterscheid and J.H. Hecht, Space Sciences Laboratory, The Aerospace Corporation, Los Angeles, CA

1. Introduction

Tides and planetary waves exhibit a great deal more variability than can be explained in terms of linear theory. By the late 60s it was believed that tides were a solved problem and that remaining difficulties could be resolved with a better description of solar heating and the background state. However, observations in recent years have shown that linear theory is unable to describe tidal variability on time scales of days and weeks. There is clear evidence that nonlinear coupling between tides and planetary waves modulates tidal amplitude [Haldoupis et al., 2004; Pancheva and Mukhtaro, 2000; Pancheva et al., 2003; Teitelbaum and Vial, 1991; Hecht et al., 2010 and references therein]. Likewise there is evidence that tides strongly affect planetary waves and that tide-planetary wave interactions involve an exchange of energy. Interactions between tides and planetary waves can produce sideband waves that can modulate the tides or induce instabilities that involve an exchange of energy between tides and planetary waves. It is clear that tide-planetary wave interactions are important in explaining tidal variability on sub-seasonal time scales and may be involved in explaining the amplification of traveling planetary waves.

A particularly interesting case is the rapid amplification of the southern hemisphere two-day wave (TDW) in austral summer to amplitudes exceeding 50 m/s important differences in the properties of the wave between the northern and the southern hemispheres. In the northern hemisphere the wave period is 52 hours or longer [e.g., Muller and Nelson, 1978; Harris and Vincent, 1993], wave growth is rather gradual, and the wave often persists near maximum amplitude for 1 or 2 months. In contrast, in the southern hemisphere the January wave event is pulse like. Wave growth is rapid and the period is often indistinguishable from 48 hours, with strong evidence that it is phase locked in local time, i.e., a given phase of the wave (time of maximum northward wind, say) occurs at the same local time for successive cycles [Craig and Elford, 1981; Poole, 1990; Harris and Vincent, 1993; Harris, 1994].

Figure 2 illustrates many of the features of the January TDW event observed with a MF radar at Christmas Island (2°N, 157°W) [Vincent and Lesicar, 1991]. The plots show amplitude and phase of the meridional wind for a 60-day period centered on the peak amplitude. The tic marks represent the middle of the respective month. The amplitude and phase were obtained by complex demodulation of the raw time series with a 48-hour period reference wave [Harris and Vincent, 1993]. The wave amplitudes peak in late January, with wave growth being particularly rapid after mid-January. Wave growth is so rapid that it is difficult to measure accurately, but it appears that amplitude doubling times are ~two cycles or shorter. The phase plot shows that prior to mid-January the

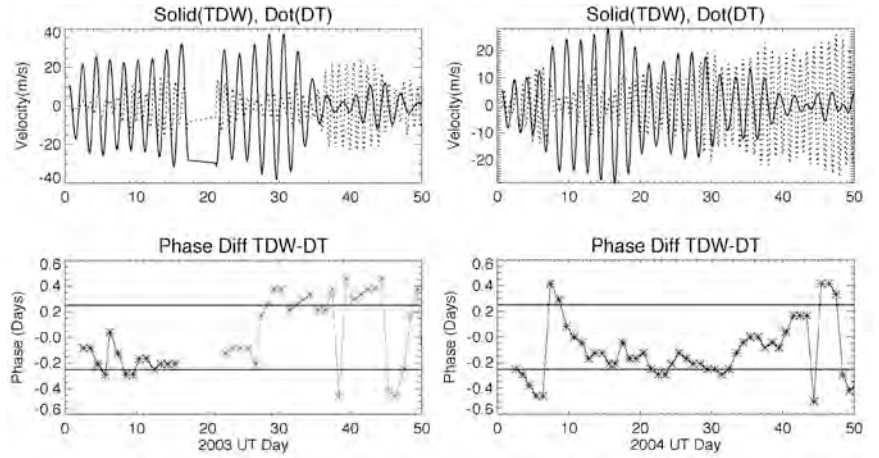


Figure 3. Panels on the left (right) are for 2003 (2004). Upper Panels: Wavelet reconstruction of the wind velocities for the TDW (solid) and diurnal tide (dotted). Lower Panels: Phase difference in days between maxima and minima of the TDW and maxima of the diurnal tide

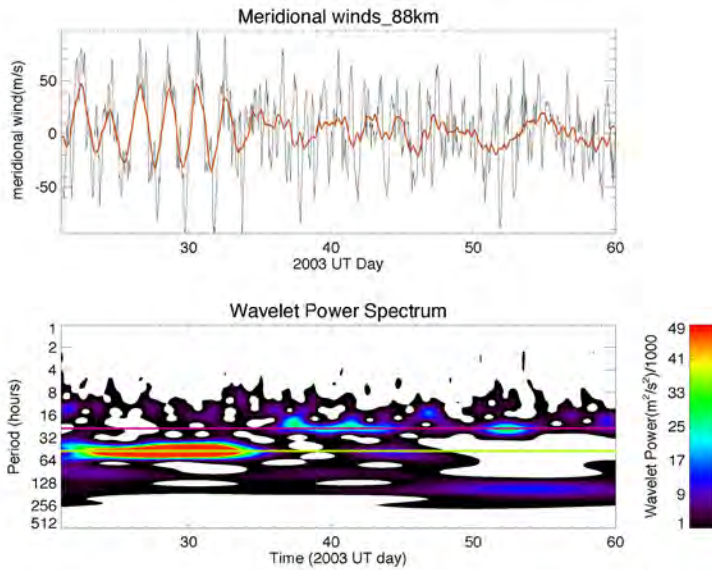


Figure 4. Lower panel: Wavelet analysis of the meridional wind at 88 km over Adelaide 21 January to 20 February 2002. Upper panel: corresponding meridional wind time series. The horizontal green and red lines denote the 48 and 24 hour periods, respectively.

phase gradient is negative, indicating a dominant period longer than 48 hours. When the amplitudes are greatest the period is stabilized close to 48 hours. Similar behavior was seen by Poole [1990] in meteor wind measurements. However, Pancheva et al. [2006] examining low latitude data found that the TDW for the January event was different from 48 hours. The main event and the successive periods of stabilization and phase progression may reflect an instability related to the successive tuning and detuning of

the normal mode resonance to periods near 48 hours.

Figure 3 shows the phase locked nature of the wave. The lower two panels show the phase difference in days between the maxima and minima of the TDW and the maxima of the diurnal tide velocities. Note that there are basically two states ($\pm \frac{1}{4}$ diurnal cycle). This suggests a linkage to the diurnal tide.

We have examined wind and airglow data from a mid-latitude site. Figure 4 shows the results of a wavelet analysis of the wind obtained from the University of Adelaide MF radar located at Buckland Park (34.6°S) in February 2003. Before day 35 there is a

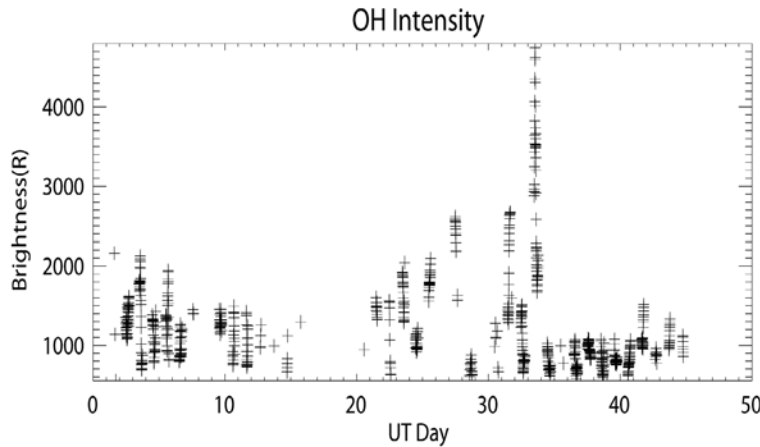


Figure 5. OH Meinel airglow intensity time series corresponding to Figure 4.

strong TDW with the strongest wavelet power by far seen very near 48 hours. The only other periodicity evident is at 16 hours. There is no power evident at tidal frequencies. After about day 35, the TDW is essentially nil except for a brief period near day 44 when a moderate TDW reoccurs. In the absence of the TDW a significant diurnal tide

appears.

Figure 5 shows a time series of nightly averaged OH Meinel airglow intensity measured with the Aerospace Nightglow Imager (ANI) sited at Adelaide. The remarkable feature of the time series is the exceptionally large value of intensity on day 34. This coincided with the sudden demise of the TDW. This was the brightest night that we have observed in five years of observation at Adelaide. The period revealed by complex demodulation of the airglow data is also 48 hours. Enhancements are seen in some other years commensurate with the end of phase-locked event, but are less dramatic. The enhancement in airglow temperature (not shown) toward the end of the 2003 event seems to be larger than can be accounted for by a steady-state TDW and may be induced by wave transience as the TDW rapidly decays [Andrews and McIntyre, 1976].

The observed phase relation between wind and temperature on days when the phase locked TDW amplitude was large indicated a meridional energy flow. This is not consistent with a global steady-state mode and suggests that the TDW is forced regionally and is exporting energy and sensible heat away from where the wave is forced [Hecht et al., 2010]. This means that the wave may force changes in the mean state.

(2) Theory

The theoretical problem for the summertime southern hemisphere TDW is to account for its sudden commencement, stability, and phase locking. Neither a simple normal mode response nor baroclinic instability offers a plausible explanation for the observed phase locking and sudden amplification. First, it is unlikely that a normal mode would

spontaneously amplify as a consequence only of its period becoming 48 hours. Similarly, it is unlikely that a baroclinic instability would manifest the observed regularity and phase stability. Phase locking to local time would be explicable if the TDW were dynamically linked to a diurnally repeatable variation such as the atmospheric thermal tide. The sudden onset and large amplitudes suggest instability. It is also significant that the phase stabilizes at large amplitudes because that is a characteristic of PSI.

A well-known observation is that an oscillating system begins to oscillate at a natural frequency if a parameter of the system is made to oscillate at a multiple of the natural frequency [Minorsky, 1982]. Walterscheid and Vincent [1996] proposed that the phase-locked TDW is the result of a parametric resonance driven by tides. The sudden onset is the result of the tuning of the normal mode TDW to a period close to a subharmonic of the diurnal tide [Hagan et al., 1993].

Walterscheid and Vincent [1996] found schemes that produce waves that may interact subharmonically to generate the phase-locked TDW. Schemes were found that involve

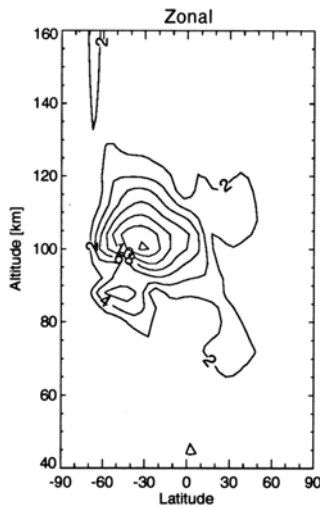


Figure 7. The diurnal $s=6$ (DW6) predicted by the TIME-GCM [Palo et al., 1999]. This is a key interaction products for the TDW tidal interaction .

zonally symmetric and migrating diurnal and semidiurnal tides. In the schemes involving the migrating tides the interaction of the TDW and the tides produces a westward propagation $s = 6$ diurnal tide (DW6). The DW6 tide propagates zonally with the same phase speed as the TDW and a product of its interaction with the TDW is itself the TDW. This is an example of an important subclass of PSI called the self-exciting triad. This scheme is more complex than the one involving the zonally symmetric tides but has the important advantage of providing linkage to the large amplitude main migrating tides.

Simulations [Hagan et al., 1993] suggest that tuning of the frequency of the resonant TDW close to subharmonic frequencies occurs during the southern hemisphere summer. The tuning required to produce the phase-locked TDW may explain why it does not occur in the northern hemisphere summer.

The schemes examined by Walterscheid and Vincent [1996] produce interaction products that include 16-

and 9.6 hour waves and an eastward propagating two-day wave. The 16- and 9.6 hour waves are seen in observations [Manson and Meek, 1990; Harris and Vincent, 1993; Hecht et al., 2010]. The interaction of the tides with the TDW in the upper mesosphere and lower thermosphere has been studied by Palo et al. [1999] using the NCAR TIME-GCM. He found the interaction products predicted by the model are those involved in the Walterscheid and Vincent [1996] scheme leading to the DW6 shown in Figure 7. These authors found a preferred response when the background winds were tuned to resemble observed January winds and when the wave was forced at 48 hours. The strongest TDW response was a mid-latitude peak in the southern hemisphere upper mesosphere. The response in the northern hemisphere was essentially nil at middle and high latitudes. Simulations were also done with the Global Scale Wave Model (GSWM) and these

simulations also showed a preferred southern hemisphere response but showed that the TDW was weaker and peaked much lower down (~75 km versus ~95 km) [Hagan, 1993]. The GSWM also showed a secondary but significant peak in the high-latitude lower thermosphere in the winter hemisphere.

3. Observations

To address these topics one can utilize wind data sets from a number of sites covering a span of latitude from Davis Station in the Antarctic to Maui airglow sets from Adelaide and Alice Springs in Australia and from Cerro Pachon in Chile; data from the TIMED SABER instrument; and the Modern Era Retrospective-analyses for Research and Applications (MERRA) reanalysis data covering the whole atmosphere up to ~80 km. The ionospheric effects can be observed by means of ground based systems such as radar and an a globally extensive array of ionosondes [refs].

Planetary waves are large scale phenomena with times scales of days to weeks. They have been detected by satellite observations in the stratosphere [Walterscheid, 1980, Salby, 1984; Madden, 2007; Hirooka, 2000]. In the ionosphere-thermosphere system these waves can be detected in emissions that respond to changes in composition and electron density, such as HI 121.6 nm, OI 130.4 nm, OI 135.6 nm, and N₂ Lyman-Birge-Hopfield bands 140 to 150 nm and 165 to 180 nm HI 121.6 nm, OI 130.4 nm, OI 135.6 nm, and N₂ Lyman-Birge-Hopfield bands 140 to 150 nm and 165 to 180 nm [Christensen et al., 2003]. In the upper mesosphere and lower thermosphere planetary waves can be detected from ground and space in a variety of emissions (that also give winds and temperature) including OH Meinel, OI 557.7 nm and O₂(0-0) [Sivjee and Walterscheid, Killeen et al, 1999; <http://www.timed.jhuapl.edu/spacecraft/SABER.pdf>]. Planetary waves can be detected from a variety of orbits (LEO, HEO or GEO). Of special interest is the resonantly amplified TDW as well as its effect on tidal components, the global airglow and temperature structure.

4. References

Andrews, D.G. and M.E. McIntyre, Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the mean zonal acceleration, *J. Atmos. Sci.*, **33**, 2031-2048, 1976

Craig, R. L., and W. G. Elford, Observations of the quasi- 2-day wave near 90 km altitude at Adelaide (35°S), *J. Atmos. Terr. Phys.*, **43**,1051-1056, 1981.

Christensen, A. B., L. J. Paxton, S. Avery, J. Craven, G. Crowley, D. C. Humm, H. Kil, R. R. Meier, C.-I Meng, D. Morrison, B. S. Ogorzalek, P. Straus, D. J. Strickland, R. M. Swenson, R. L. Walterscheid, B. Wolven, and Y. Zhang, Initial observations with the Global Ultraviolet Imager (GUVI) in the NASA TIMED satellite mission, *J. Geophys. Res.*, **108** (A12), 1451, doi:10.1029/2003JA009918, 2003.

Hagan, M. E., J. M. Forbes, and F. Vial, A numerical investigation of the propagation of the quasi 2-day wave into the lower atmosphere, *J Geophys. Res.*, **98**, 23,193-23,205, 1993.

Haldoupis C., D. Pancheva, N. J. Mitchell, A study of tidal and planetary wave periodicities present in midlatitude sporadic E layers, *J. Geophys. Res.*, **109**, A02302, doi:10.1029/2003JA010253, 2004

- Harris, T. J., A long term study of the quasi-two-day wave in the middle atmosphere, *J. Atmos. Terr. Phys.*, 56, 569-579, 1994.
- Harris, T. J., and R. A. Vincent, The quasi-two-day wave observed in the equatorial middle atmosphere, *J. Geophys. Res.*, 98, 10,481-10,490, 1993.
- Hecht, J. H., R. L. Walterscheid, L. J. Gelinas, R. A. Vincent, I. M. Reid, and J. M. Woithe (2010), Observations of the Phase-Locked Two day Wave Over the Australian Sector Using Medium Frequency Radar and Airglow Data, *J. Geophys. Res.*, doi:10.1029/2009JD013772.
- Hirooka, Toshihiko, 2000: Normal Mode Rossby Waves as Revealed by UARS/ISAMS Observations. *J. Atmos. Sci.*, 57, 1277–1285.
- Killeen, T.L., W. R. Skinner, R. M. Johnson, C. J. Edmonson, Q. Wu, R. J. Niciejewski, H. J. Grassl, D. A. Gell, P. E. Hansen, J. D. Harvey, J. F. Kafkalidis, TIMED Doppler Interferometer (TIDI), *Proceedings of SPIE*, 1999
- Madden, R. A. (2007), Large-scale, free Rossby waves in the atmosphere—an update. *Tellus A*, 59: 571–590. doi:10.1111/j.1600-0870.2007.00257.x
- Manson, A. H., and C. E. Meek, Long period (~8-20 h) wind oscillations in the upper middle atmosphere at Saskatoon (52°), Evidence for non-linear tidal effects, *Planet. Space Sci.*, 38, 1431-1441, 1990.
- Minorsky, N., *Nonlinear Oscillations*, 714 pp., Krieger, Melbourne, Fla., 1982.
- Muller, H. G., and L. Nelson, A traveling quasi 2-day wave in the meteor region, *J. Atmos. Terr. Phys.*, 40, 761-766, 1978.
- Palo, S. E., R. G. Roble, and M. E. Hagan, Middle atmosphere effects of the quasi-two-day wave determined from a general circulation model, *Earth Planets Space*, 51, 629–647, 1999
- Pancheva, D., and P. Mukhtarov, Wavelet analysis on transient behaviour of tidal amplitude fluctuations observed by meteor radar in the lower thermosphere above Bulgaria *Ann. Geophys.*, 18, 316-331, 2000
- Pancheva D., C. Haldoupis, C. E. Meek, A. H. Manson, N. J. Mitchell, Evidence of a role for modulated atmospheric tides in the dependence of sporadic E layers on planetary waves, *J. Geophys. Res.*, 108 (A5), 1176, doi:10.1029/2002JA009788, 2003
- Pancheva D. V., et al., Two-day wave coupling of the low-latitude atmosphere-ionosphere system, *J. Geophys. Res.*, 111, A07313, doi:10.1029/2005JA011562, 2006
- Poole, L. M. G., The characteristics of the mesospheric two-day wave as observed at Grahamstown (33.3°S, 26SE), *J. Atmos. Terr. Phys.*, 52, 259-268, 1990.
- Salby, M. L. (1984), Survey of planetary-scale traveling waves: The state of theory and observations, *Rev. Geophys.*, 22(2), 209–236, doi:10.1029/RG022i002p00209.
- Sivjee, G. G., R. L. Walterscheid, J. H. Hecht, R. M. Hamwey, G. Schubert and A. B. Christensen, Effects of atmospheric disturbances on polar mesopause airglow OH emissions", *J. Geophys. Res.*, 92, 7651-7656, 1987.

Sivjee, G. G., R. L. Walterscheid, and D. J. McEwen, Planetary wave disturbances in the arctic winter mesopause over Eureka (80°N), *Plan. Space Sci.*, 42, 973-986, 1994

Teitelbaum, H., and F. Vial, On tidal variability induced by nonlinear interaction with planetary waves, *J. Geophys. Res.*, 96, 14169-14178, 1991

Vincent, R. A., and D. Lesicar, Dynamics of the equatorial mesosphere; First results with a new generation partial reflection radar, *Geophys. Res. Lett.*, 18, 825-828, 1991.

Walterscheid, R. L. Traveling planetary waves in the stratosphere, *Pure Appl. Geophys.*, 188, 177, 1980.

Walterscheid, R., and R. Vincent, Tidal generation of the phase-locked 2-day wave in the southern hemisphere summer by wave-wave interactions, *J. Geophys. Res.*, 101, 26567-26576, 1996

Effects of Large Amplitude Planetary Waves in the Ionosphere and Thermosphere

R.L. Walterscheid and J.H. Hecht, Space Sciences Laboratory, The Aerospace Corporation, Los Angeles, CA

1. Introduction

Tides and planetary waves exhibit a great deal more variability than can be explained in terms of linear theory. By the late 60s it was believed that tides were a solved problem and that remaining difficulties could be resolved with a better description of solar heating and the background state. However, observations in recent years have shown that linear theory is unable to describe tidal variability on time scales of days and weeks. There is clear evidence that nonlinear coupling between tides and planetary waves modulates tidal amplitude [Haldoupis et al., 2004; Pancheva and Mukhtaro, 2000; Pancheva et al., 2003; Teitelbaum and Vial, 1991; Hecht et al., 2010 and references therein]. Likewise there is evidence that tides strongly affect planetary waves and that tide-planetary wave interactions involve an exchange of energy. Interactions between tides and planetary waves can produce sideband waves that can modulate the tides or induce instabilities that involve an exchange of energy between tides and planetary waves. It is clear that tide-planetary wave interactions are important in explaining tidal variability on sub-seasonal time scales and may be involved in explaining the amplification of traveling planetary waves.

A particularly interesting case is the rapid amplification of the southern hemisphere two-day wave (TDW) in austral summer to amplitudes exceeding 50 m/s important differences in the properties of the wave between the northern and the southern hemispheres. In the northern hemisphere the wave period is 52 hours or longer [e.g., Muller and Nelson, 1978; Harris and Vincent, 1993], wave growth is rather gradual, and the wave often persists near maximum amplitude for 1 or 2 months. In contrast, in the southern hemisphere the January wave event is pulse like. Wave growth is rapid and the period is often indistinguishable from 48 hours, with strong evidence that it is phase locked in local time, i.e., a given phase of the wave (time of maximum northward wind, say) occurs at the same local time for successive cycles [Craig and Elford, 1981; Poole, 1990; Harris and Vincent, 1993; Harris, 1994].

Figure 2 illustrates many of the features of the January TDW event observed with a MF radar at Christmas Island (2°N, 157°W) [Vincent and Lesicar, 1991]. The plots show amplitude and phase of the meridional wind for a 60-day period centered on the peak amplitude. The tic marks represent the middle of the respective month. The amplitude and phase were obtained by complex demodulation of the raw time series with a 48-hour period reference wave [Harris and Vincent, 1993]. The wave amplitudes peak in late January, with wave growth being particularly rapid after mid-January. Wave growth is so rapid that it is difficult to measure accurately, but it appears that amplitude doubling times are ~two cycles or shorter. The phase plot shows that prior to mid-January the

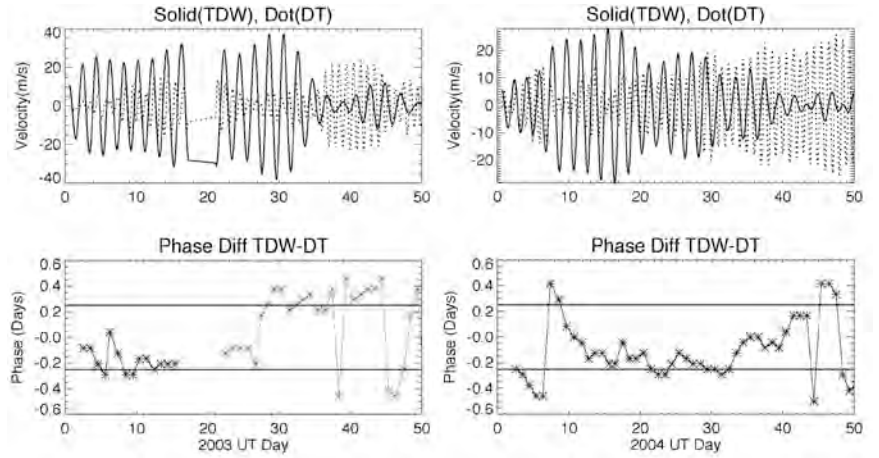


Figure 3. Panels on the left (right) are for 2003 (2004). Upper Panels: Wavelet reconstruction of the wind velocities for the TDW (solid) and diurnal tide (dotted). Lower Panels: Phase difference in days between maxima and minima of the TDW and maxima of the diurnal tide

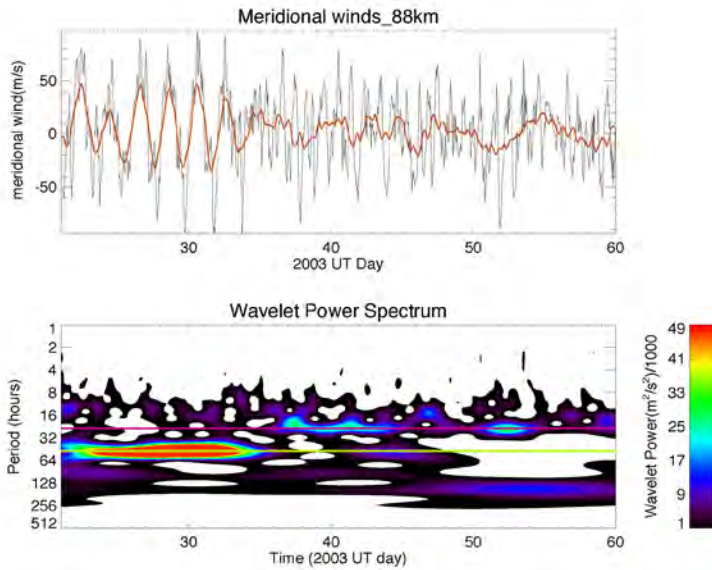


Figure 4. Lower panel: Wavelet analysis of the meridional wind at 88 km over Adelaide 21 January to 20 February 2002. Upper panel: corresponding meridional wind time series. The horizontal green and red lines denote the 48 and 24 hour periods, respectively.

phase gradient is negative, indicating a dominant period longer than 48 hours. When the amplitudes are greatest the period is stabilized close to 48 hours. Similar behavior was seen by Poole [1990] in meteor wind measurements. However, Pancheva et al. [2006] examining low latitude data found that the TDW for the January event was different from 48 hours. The main event and the successive periods of stabilization and phase progression may reflect an instability related to the successive tuning and detuning of

the normal mode resonance to periods near 48 hours.

Figure 3 shows the phase locked nature of the wave. The lower two panels show the phase difference in days between the maxima and minima of the TDW and the maxima of the diurnal tide velocities. Note that there are basically two states ($\pm \frac{1}{4}$ diurnal cycle). This suggests a linkage to the diurnal tide.

We have examined wind and airglow data from a mid-latitude site. Figure 4 shows the results of a wavelet analysis of the wind obtained from the University of Adelaide MF radar located at Buckland Park (34.6°S) in February 2003. Before day 35 there is a

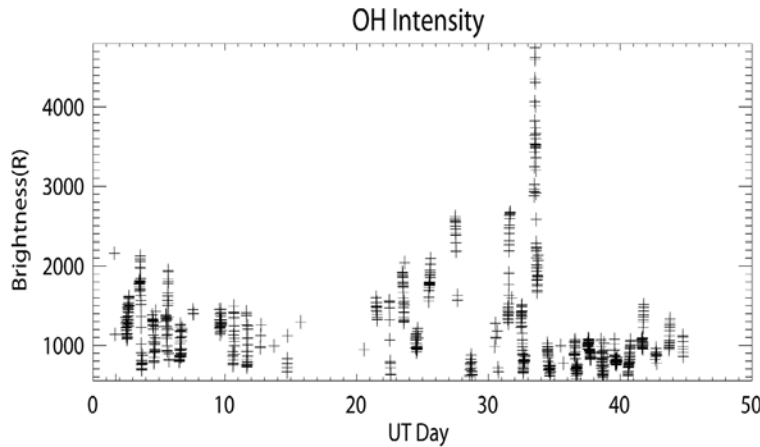


Figure 5. OH Meinel airglow intensity time series corresponding to Figure 4.

strong TDW with the strongest wavelet power by far seen very near 48 hours. The only other periodicity evident is at 16 hours. There is no power evident at tidal frequencies. After about day 35, the TDW is essentially nil except for a brief period near day 44 when a moderate TDW reoccurs. In the absence of the TDW a significant diurnal tide

appears.

Figure 5 shows a time series of nightly averaged OH Meinel airglow intensity measured with the Aerospace Nightglow Imager (ANI) sited at Adelaide. The remarkable feature of the time series is the exceptionally large value of intensity on day 34. This coincided with the sudden demise of the TDW. This was the brightest night that we have observed in five years of observation at Adelaide. The period revealed by complex demodulation of the airglow data is also 48 hours. Enhancements are seen in some other years commensurate with the end of phase-locked event, but are less dramatic. The enhancement in airglow temperature (not shown) toward the end of the 2003 event seems to be larger than can be accounted for by a steady-state TDW and may be induced by wave transience as the TDW rapidly decays [Andrews and McIntyre, 1976].

The observed phase relation between wind and temperature on days when the phase locked TDW amplitude was large indicated a meridional energy flow. This is not consistent with a global steady-state mode and suggests that the TDW is forced regionally and is exporting energy and sensible heat away from where the wave is forced [Hecht et al., 2010]. This means that the wave may force changes in the mean state.

(2) Theory

The theoretical problem for the summertime southern hemisphere TDW is to account for its sudden commencement, stability, and phase locking. Neither a simple normal mode response nor baroclinic instability offers a plausible explanation for the observed phase locking and sudden amplification. First, it is unlikely that a normal mode would

spontaneously amplify as a consequence only of its period becoming 48 hours. Similarly, it is unlikely that a baroclinic instability would manifest the observed regularity and phase stability. Phase locking to local time would be explicable if the TDW were dynamically linked to a diurnally repeatable variation such as the atmospheric thermal tide. The sudden onset and large amplitudes suggest instability. It is also significant that the phase stabilizes at large amplitudes because that is a characteristic of PSI.

A well-known observation is that an oscillating system begins to oscillate at a natural frequency if a parameter of the system is made to oscillate at a multiple of the natural frequency [Minorsky, 1982]. Walterscheid and Vincent [1996] proposed that the phase-locked TDW is the result of a parametric resonance driven by tides. The sudden onset is the result of the tuning of the normal mode TDW to a period close to a subharmonic of the diurnal tide [Hagan et al., 1993].

Walterscheid and Vincent [1996] found schemes that produce waves that may interact subharmonically to generate the phase-locked TDW. Schemes were found that involve

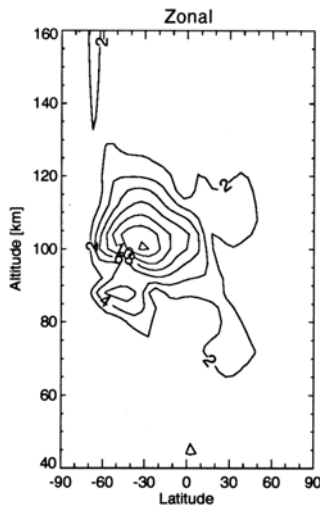


Figure 7. The diurnal $s=6$ (DW6) predicted by the TIME-GCM [Palo et al., 1999]. This is a key interaction products for the TDW tidal interaction .

zonally symmetric and migrating diurnal and semidiurnal tides. In the schemes involving the migrating tides the interaction of the TDW and the tides produces a westward propagation $s = 6$ diurnal tide (DW6). The DW6 tide propagates zonally with the same phase speed as the TDW and a product of its interaction with the TDW is itself the TDW. This is an example of an important subclass of PSI called the self-exciting triad. This scheme is more complex than the one involving the zonally symmetric tides but has the important advantage of providing linkage to the large amplitude main migrating tides.

Simulations [Hagan et al., 1993] suggest that tuning of the frequency of the resonant TDW close to subharmonic frequencies occurs during the southern hemisphere summer. The tuning required to produce the phase-locked TDW may explain why it does not occur in the northern hemisphere summer.

The schemes examined by Walterscheid and Vincent [1996] produce interaction products that include 16-

and 9.6 hour waves and an eastward propagating two-day wave. The 16- and 9.6 hour waves are seen in observations [Manson and Meek, 1990; Harris and Vincent, 1993; Hecht et al., 2010]. The interaction of the tides with the TDW in the upper mesosphere and lower thermosphere has been studied by Palo et al. [1999] using the NCAR TIME-GCM. He found the interaction products predicted by the model are those involved in the Walterscheid and Vincent [1996] scheme leading to the DW6 shown in Figure 7. These authors found a preferred response when the background winds were tuned to resemble observed January winds and when the wave was forced at 48 hours. The strongest TDW response was a mid-latitude peak in the southern hemisphere upper mesosphere. The response in the northern hemisphere was essentially nil at middle and high latitudes. Simulations were also done with the Global Scale Wave Model (GSWM) and these

simulations also showed a preferred southern hemisphere response but showed that the TDW was weaker and peaked much lower down (~75 km versus ~95 km) [Hagan, 1993]. The GSWM also showed a secondary but significant peak in the high-latitude lower thermosphere in the winter hemisphere.

3. Observations

To address these topics one can utilize wind data sets from a number of sites covering a span of latitude from Davis Station in the Antarctic to Maui airglow sets from Adelaide and Alice Springs in Australia and from Cerro Pachon in Chile; data from the TIMED SABER instrument; and the Modern Era Retrospective-analyses for Research and Applications (MERRA) reanalysis data covering the whole atmosphere up to ~80 km. The ionospheric effects can be observed by means of ground based systems such as radar and an a globally extensive array of ionosondes [refs].

Planetary waves are large scale phenomena with times scales of days to weeks. They have been detected by satellite observations in the stratosphere [Walterscheid, 1980, Salby, 1984; Madden, 2007; Hirooka, 2000]. In the ionosphere-thermosphere system these waves can be detected in emissions that respond to changes in composition and electron density, such as HI 121.6 nm, OI 130.4 nm, OI 135.6 nm, and N₂ Lyman-Birge-Hopfield bands 140 to 150 nm and 165 to 180 nm HI 121.6 nm, OI 130.4 nm, OI 135.6 nm, and N₂ Lyman-Birge-Hopfield bands 140 to 150 nm and 165 to 180 nm [Christensen et al., 2003]. In the upper mesosphere and lower thermosphere planetary waves can be detected from ground and space in a variety of emissions (that also give winds and temperature) including OH Meinel, OI 557.7 nm and O₂(0-0) [Sivjee and Walterscheid, Killeen et al, 1999; <http://www.timed.jhuapl.edu/spacecraft/SABER.pdf>]. Planetary waves can be detected from a variety of orbits (LEO, HEO or GEO). Of special interest is the resonantly amplified TDW as well as its effect on tidal components, the global airglow and temperature structure.

4. References

Andrews, D.G. and M.E. McIntyre, Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the mean zonal acceleration, *J. Atmos. Sci.*, **33**, 2031-2048, 1976

Craig, R. L., and W. G. Elford, Observations of the quasi- 2-day wave near 90 km altitude at Adelaide (35°S), *J. Atmos. Terr. Phys.*, **43**,1051-1056, 1981.

Christensen, A. B., L. J. Paxton, S. Avery, J. Craven, G. Crowley, D. C. Humm, H. Kil, R. R. Meier, C.-I Meng, D. Morrison, B. S. Ogorzalek, P. Straus, D. J. Strickland, R. M. Swenson, R. L. Walterscheid, B. Wolven, and Y. Zhang, Initial observations with the Global Ultraviolet Imager (GUVI) in the NASA TIMED satellite mission, *J. Geophys. Res.*, **108** (A12), 1451, doi:10.1029/2003JA009918, 2003.

Hagan, M. E., J. M. Forbes, and F. Vial, A numerical investigation of the propagation of the quasi 2-day wave into the lower atmosphere, *J Geophys. Res.*, **98**, 23,193-23,205, 1993.

Haldoupis C., D. Pancheva, N. J. Mitchell, A study of tidal and planetary wave periodicities present in midlatitude sporadic E layers, *J. Geophys. Res.*, **109**, A02302, doi:10.1029/2003JA010253, 2004

- Harris, T. J., A long term study of the quasi-two-day wave in the middle atmosphere, *J. Atmos. Terr. Phys.*, 56, 569-579, 1994.
- Harris, T. J., and R. A. Vincent, The quasi-two-day wave observed in the equatorial middle atmosphere, *J. Geophys. Res.*, 98, 10,481-10,490, 1993.
- Hecht, J. H., R. L. Walterscheid, L. J. Gelinas, R. A. Vincent, I. M. Reid, and J. M. Woithe (2010), Observations of the Phase-Locked Two day Wave Over the Australian Sector Using Medium Frequency Radar and Airglow Data, *J. Geophys. Res.*, doi:10.1029/2009JD013772.
- Hirooka, Toshihiko, 2000: Normal Mode Rossby Waves as Revealed by UARS/ISAMS Observations. *J. Atmos. Sci.*, 57, 1277–1285.
- Killeen, T.L., W. R. Skinner, R. M. Johnson, C. J. Edmonson, Q. Wu, R. J. Niciejewski, H. J. Grassl, D. A. Gell, P. E. Hansen, J. D. Harvey, J. F. Kafkalidis, TIMED Doppler Interferometer (TIDI), *Proceedings of SPIE*, 1999
- Madden, R. A. (2007), Large-scale, free Rossby waves in the atmosphere—an update. *Tellus A*, 59: 571–590. doi:10.1111/j.1600-0870.2007.00257.x
- Manson, A. H., and C. E. Meek, Long period (~8-20 h) wind oscillations in the upper middle atmosphere at Saskatoon (52°), Evidence for non-linear tidal effects, *Planet. Space Sci.*, 38, 1431-1441, 1990.
- Minorsky, N., *Nonlinear Oscillations*, 714 pp., Krieger, Melbourne, Fla., 1982.
- Muller, H. G., and L. Nelson, A traveling quasi 2-day wave in the meteor region, *J. Atmos. Terr. Phys.*, 40, 761-766, 1978.
- Palo, S. E., R. G. Roble, and M. E. Hagan, Middle atmosphere effects of the quasi-two-day wave determined from a general circulation model, *Earth Planets Space*, 51, 629–647, 1999
- Pancheva, D., and P. Mukhtarov, Wavelet analysis on transient behaviour of tidal amplitude fluctuations observed by meteor radar in the lower thermosphere above Bulgaria *Ann. Geophys.*, 18, 316-331, 2000
- Pancheva D., C. Haldoupis, C. E. Meek, A. H. Manson, N. J. Mitchell, Evidence of a role for modulated atmospheric tides in the dependence of sporadic E layers on planetary waves, *J. Geophys. Res.*, 108 (A5), 1176, doi:10.1029/2002JA009788, 2003
- Pancheva D. V., et al., Two-day wave coupling of the low-latitude atmosphere-ionosphere system, *J. Geophys. Res.*, 111, A07313, doi:10.1029/2005JA011562, 2006
- Poole, L. M. G., The characteristics of the mesospheric two-day wave as observed at Grahamstown (33.3°S, 26SE), *J. Atmos. Terr. Phys.*, 52, 259-268, 1990.
- Salby, M. L. (1984), Survey of planetary-scale traveling waves: The state of theory and observations, *Rev. Geophys.*, 22(2), 209–236, doi:10.1029/RG022i002p00209.
- Sivjee, G. G., R. L. Walterscheid, J. H. Hecht, R. M. Hamwey, G. Schubert and A. B. Christensen, Effects of atmospheric disturbances on polar mesopause airglow OH emissions", *J. Geophys. Res.*, 92, 7651-7656, 1987.

Sivjee, G. G., R. L. Walterscheid, and D. J. McEwen, Planetary wave disturbances in the arctic winter mesopause over Eureka (80°N), *Plan. Space Sci.*, 42, 973-986, 1994

Teitelbaum, H., and F. Vial, On tidal variability induced by nonlinear interaction with planetary waves, *J. Geophys. Res.*, 96, 14169-14178, 1991

Vincent, R. A., and D. Lesicar, Dynamics of the equatorial mesosphere; First results with a new generation partial reflection radar, *Geophys. Res. Lett.*, 18, 825-828, 1991.

Walterscheid, R. L. Traveling planetary waves in the stratosphere, *Pure Appl. Geophys.*, 188, 177, 1980.

Walterscheid, R., and R. Vincent, Tidal generation of the phase-locked 2-day wave in the southern hemisphere summer by wave-wave interactions, *J. Geophys. Res.*, 101, 26567-26576, 1996

The Paired Ionosphere-Thermosphere Orbiters (PITO)

Mission: Multipoint Geospace Science in 3D

R. Walterscheid, D. Nigg, D. Judnick, J. Lang (The Aerospace Corporation)

J. Spann (NASA/Marshall Space Flight Center)

This white paper describes a mission concept for pairing *in situ* measurements with imaging to identify signatures seen in the *in situ* measurements with structures seen in the images. The Paired Ionosphere Thermosphere Orbiters (PITO) concept utilizes two spacecraft in “equal but opposite” eccentric orbits so that when one satellite is at apogee the other is at perigee and located within the field of view of the other. Thus the measurements combine the best features of remote sensing (coverage) with *in situ* measurements (detail).

Orbital Concept:

The PITO mission combines comprehensive in-situ measurements with large field-of-view imaging and

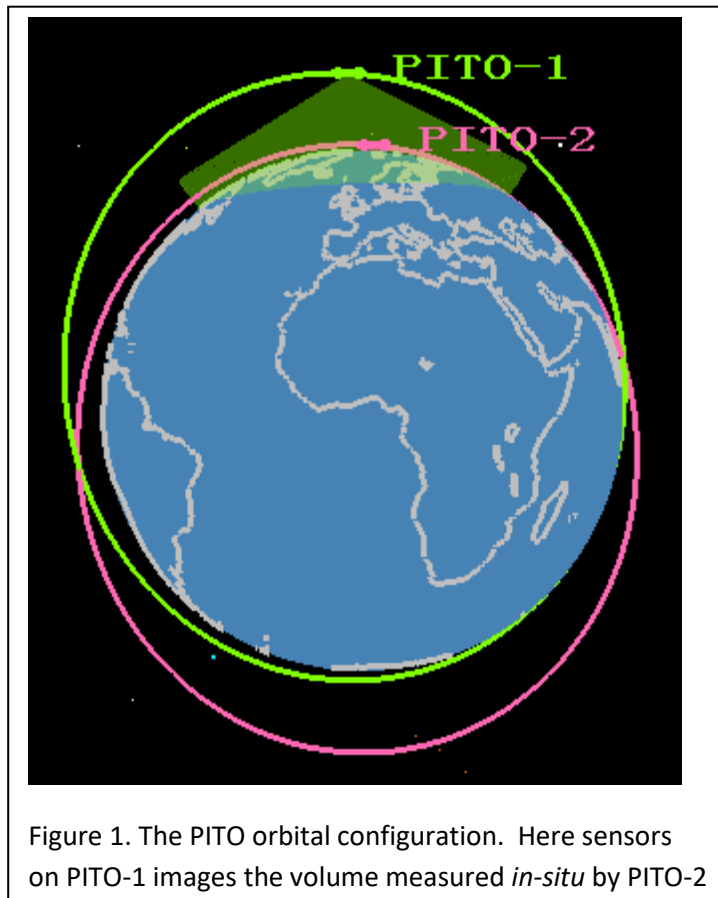


Figure 1. The PITO orbital configuration. Here sensors on PITO-1 images the volume measured *in-situ* by PITO-2

vertical sounding instruments to provide new insights into the physics of the ionosphere-thermosphere system. PITO packs strategic-impact science into a package about the size of a NASA Medium Explorer (MIDEX), or slightly larger mission. The power of PITO is derived through its two-satellite orbital configuration, which provides for in-situ/imaging co-volume measurements twice per orbit. One spacecraft is directly below the other when they are at, respectively, perigee and apogee. In other words, their arguments of perigee differ by 180 degrees. The addition of vertical sounding measurements gives an overall picture that has aspects of a complete three-dimensional description.

The orbital concept is illustrated in Figure 1. In this configuration one satellite (PITO-2) is making *in situ* measurements of features imaged by the other (PITO-1).

Thus PITO is favorably aligned to measure inputs at high latitudes from high altitudes and the responses in the I-T system.

PITO Science

The PITO science concept is to use paired satellites for simultaneous observation of large and small scale ionosphere-thermosphere phenomena in two regions of the atmosphere and the coupling between them to determine on different scales. While the lower satellite makes detailed measurements of the ionosphere and neutral atmosphere in the lower thermosphere the other is providing the context for the measurements (e.g., images of auroral features) and providing information on the energy inputs from the magnetosphere. These measurements can be applied to studies that include how the ionosphere-thermosphere (I-T) system responds to magnetospheric forcing; how solar and geospace forcing causes ionospheric density irregularities below 1000 km; how composition changes in the aurora are driven; how high-latitude electrodynamics exerts control on the global circulation of the atmosphere; and how equatorial depletions are related to atmospheric disturbances. Figure 2 illustrates phenomena that PITO will be able to image while making detailed in situ measurements.

Magnetospheric-Ionospheric (M-I) Coupling

PITO encompasses science objectives from two of the science targets from NASA's 2009 Heliophysics Roadmap, ONEP and INCA pertaining to top-level M-I coupling goals:

1. How does the ionosphere-thermosphere (I-T) system respond to magnetospheric forcing?

PITO will provide simultaneous measurements of the M-I forcing and the I-T response. The great advantage of PITO over single satellites that simultaneously measure inputs and response is that PITO provides the morphology of the structures that are sensed *in situ* and allows one to identify the phenomenon (arc, diffuse aurora, westward traveling surge, etc.) and observe the nature of the response an area much larger than the *in situ* point observation.

2. How is the I-T system dynamically coupled to the magnetosphere?

The I-T system provides feedback to the M-I system (e.g., the flywheel effect). PITO will provide data that will allow one to correlate the energy flowing into the M-I system from the I-T in response to M-I forcing. As a specific example, PITO can address the recently-discovered phenomenon of neutral gas upwelling beneath the magnetospheric cusp. The low-altitude platform will perform in-situ measurements of the associated fields, particles, and winds. The high-altitude platform will provide context measurements of composition changes with the imager as well as primary particle fluxes and fields.

Cross-scale Coupling

PITO will also enable the separation of locally driven changes from those on large spatial scales with different time scales. This determination requires more than single-point measurements, and PITO brings to bear dual *in-situ* measurements and the wide-area measurements of remote imaging on this problem. Questions that become answerable are:

1. What is the role of high-latitude electrodynamics in the global circulation of the atmosphere through heating and wave generation?

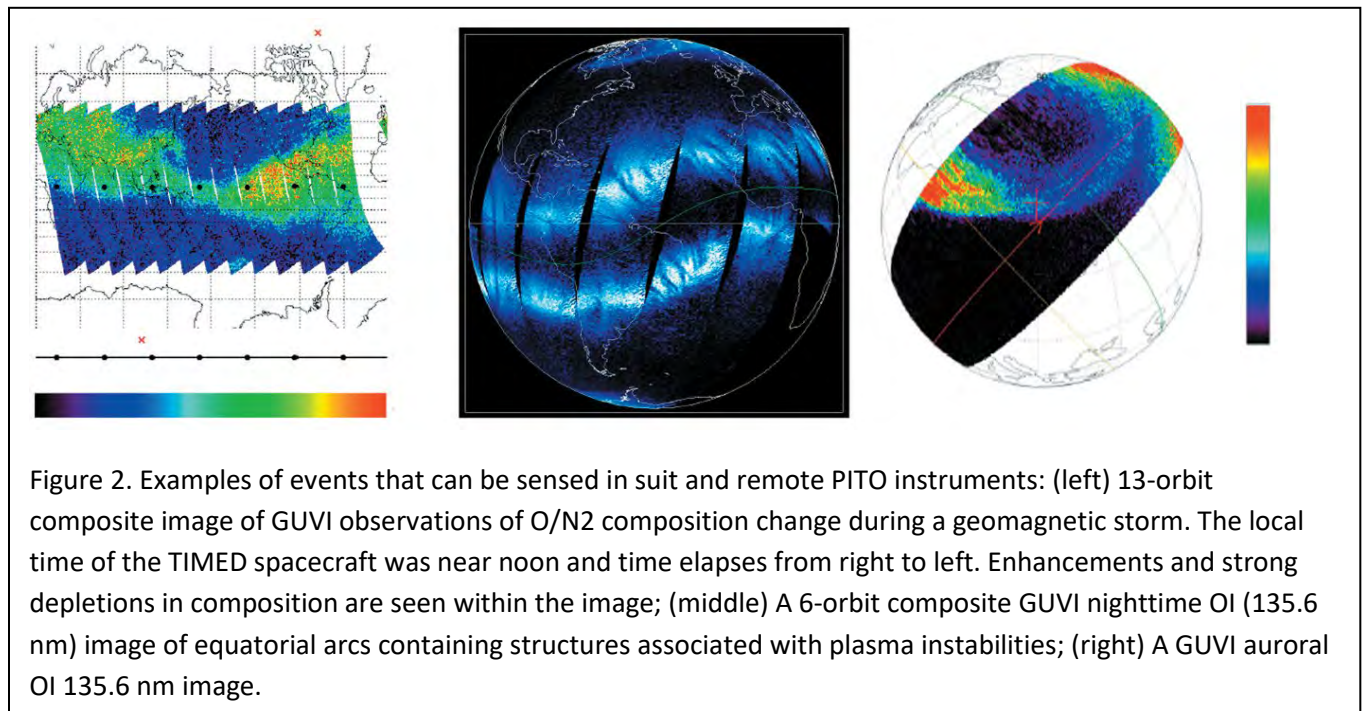
PITO will be able to image waves and infer upwelling generated in response to M-I heat inputs. Heating can be evaluated from PITO measurements.

2. How are equatorial depletions related to small-scale atmospheric disturbances?

PITO can measure *in situ* wave structure below the F-region peak and simultaneously image structure associated with depletions.

3. How are composition changes in the aurora driven?

PITO can image high latitude O depletions during active periods and correlate them with *in situ* measurements of M-I input and the T-I response.



PITO Measurements

PITO's science suite consists of four instrument packages:

- In-situ neutral and ionized gas parameters.
 - Neutral composition, density and winds
- In-situ electrodynamic and charged particle parameters
 - Electron density ion composition, ion temperature, electric and magnetic fields, particle fluxes, profiles of plasma density
- Remote sensing of aeronomic emissions
 - LBH long and LBH short, 135.6 nm giving images of nightside total electron content, dayside O/N₂ ratio, auroral energy input and characteristic energy

- Vertical profiles of plasma densities
- Vertical profiles of neutral density (if spaceborne lidar is available)

The imaging component yields 2-D measurements, while the vertical profiles add an element of 3-D measurement

PITO Instruments

The PITO comprises approximately 80 kg of payload (instruments) that can be used to characterize the I-T system and M-I forcing by *in situ* and remote sensing over a ranges of temporal and spatial scales.

Package 1: In-situ neutral and ionized gas parameters

- Langmuir probe for ionized gas density and temperature
- Ion drift meter and retarding potential analyzer for plasma flow velocity
- Ionization gauge for neutral gas density
- Cross-track and ram wind sensor
- Mass spectrometer for ion and neutral composition

Package 2: In-situ electrodynamic and charged particle parameters

- Three-axis electric field double probe for DC and AC electric fields
- Boom-mounted three-axis fluxgate magnetometer for DC magnetic fields
- Boom-mounted three-axis search coil magnetometer for AC magnetic fields
- Electrostatic analyzer for charged particle distribution functions

Package 3: Remote sensing of aeronomic emissions

PITO will carry a wide field-of-view ultraviolet imager imaging the following emissions

- 135.6 nm
- Lyman-Birge-Hopfield short
- Lyman-Birge-Hopfield long

Package 4: Vertical profiles of plasma densities

- Topside-bottomside sounder for electron density profiles
- (If available) Spaceborne lidar for neutral density profiles

The PITO mission study

PITO has been studied by Aerospace's Concept Design Center and the results of the study are reflected in the description of the mission and concept spacecraft given below. The design fidelity is approximate, but is generally reliable for mission concepts. The process also identifies areas of difficulty or aspects of the mission that require in-depth study. All systems were baselined and evaluated. In all, twelve configurations were evaluated for PITO.

The PITO Mission

Orbit

The PITO orbit is shown in Figure 3.

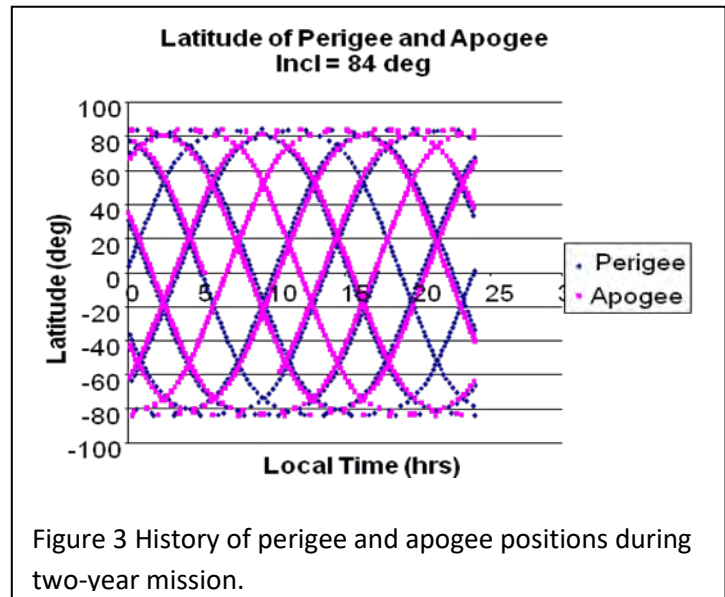
- Perigee Altitude = 200 km
- Apogee Altitude = 2,000 km
- Inclination = 84 deg.
- Period = 107.3 min

Lifetime

- Mission designed to last two years
- *Limited by propellant*
- *Re-enters after 6 years*

Data transmission

- Orbit-average data rate of 500 kbps
- *4 Mbps S-band transmission rate*
- *Ground stations at Wallops Island, Poker Flat, and McMurdo*



The PITO Spacecraft

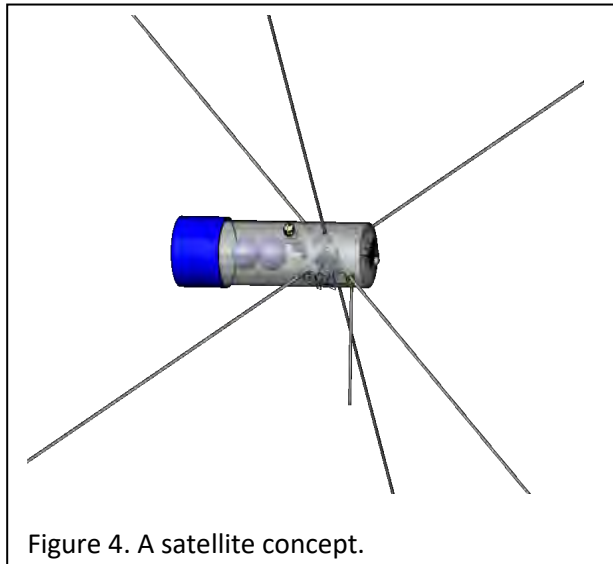


Figure 4. A satellite concept.

The PITO spacecraft is three-axis stabilized to view the nadir. It is modeled on NASA's Geospace Electrodynamic Connections (GEC) strawman concept. The design includes long booms to separate instruments from the body. "Pencil shaped design is similar to C/NOFs. PITO was designed as MIDEX class mission able to fit in a relatively small vehicle to keep costs down. The PITO spacecraft is shown in Figure 4. The stowed configuration is shown in Figure 5.

- Power considerations with 28% triple-junction solar array requires deployable “skirt” with extra array area
 - *Similar to C/NOFS*
- Mass (kg)
 - *Payload: 102*
 - *Spacecraft (dry): 318*
 - *Propellant: 93*
 - *21% margin on Taurus 2110*
- Power (W)
 - *Payload: 119*
 - *Spacecraft: 129*
 - *While transmitting: 226*
 - *Array end-of-life: 600*

Attitude Determination and Control

- Based on the selected sensors and actuators the following pointing budget can be provided.
 - *Attitude Determination $0.05^\circ 3\sigma$*
 - *Attitude Control $0.02^\circ 3\sigma$*
 - *Misalignments $0.5^\circ 3\sigma$*
 - *Total Pointing $0.5^\circ 3\sigma$*

Propulsion Subsystem Summary

The propulsion subsystem selected is a conventional blowdown monopropellant hydrazine system. The number of tanks (one or two) that will be used depends upon propellant needs for the various configurations. A single 5-lbf thruster is used only if the configuration requires drag makeup and/or disposal maneuvers

Summary and Conclusions

PITO has been designed to provide the concurrent, three-dimensional, multipoint measurements needed to advance geospace science while staying within a stringent resource envelope. The mission utilizes a pair of orbiting vehicles in eccentric, high-inclination, coplanar orbits. The orbits have arguments of perigee that differ by 180 degrees and are phased such that one vehicle is at perigee (~200 km) while the second is at apogee (~2000 km). Half an orbit later, the vehicles switch positions. Three complementary types of measurements exploit this scenario: local, *in-situ* measurements on both satellites, two-dimensional imaging from the higher satellite, and vertical sounders.

The main idea is that two-dimensional context information for the low-altitude measurements is obtained by the high-altitude imagers, while information on the third dimension is provided by vertical

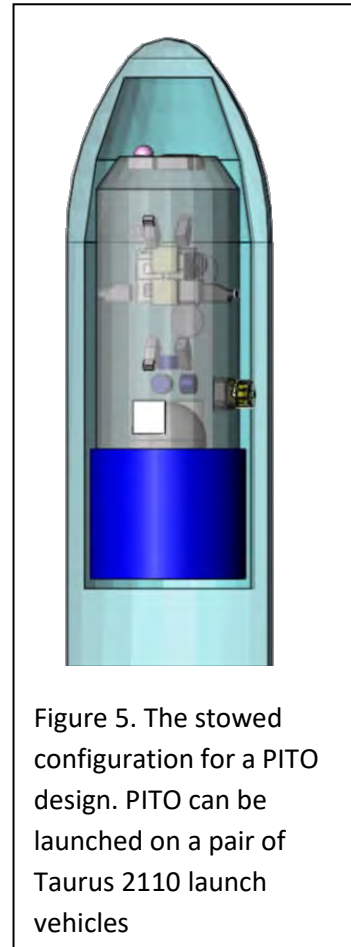


Figure 5. The stowed configuration for a PITO design. PITO can be launched on a pair of Taurus 2110 launch vehicles

profiling. The PITO mission presents an opportunity for performing quasi-3D measurements of the ionosphere-thermosphere system.

PITO's methodology provides the comprehensive, high-duty cycle measurements needed to make progress on several high-priority science targets in I-T science. Such an observation system is capable of providing elements of global coverage, regional coverage, and concurrent coverage in three dimensions. Science goals are presented, as are the results of a detailed implementation plan, including several trade studies on key elements of the mission. The conclusion is that the mission would enable significant new understanding of the ionosphere-thermosphere system within a resource envelope that is consistent with that of NASA's Medium Explorer (MIDEX) line of science missions.

The two-satellite mission can be accomplished within the approximate cost envelope of NASA's Medium-class Explorer mission.

Continuous FUV/EUV Imaging of the Ionosphere from Geosynchronous Orbit

K. S. Wood, K. F. Dymond, S. A. Budzien, S. E. McDonald, C. Coker,
A. C. Nicholas, and M. P. Kowalski
Naval Research Laboratory
Washington, DC

Abstract

Low-Earth orbit spectrometers and spectral imagers, which provide regional observations only at widely-spaced time intervals, need to be supplemented by sensitive instruments in high or even geostationary orbits. These new imaging systems will generate greatly improved measurements, presenting them in 2-dimensional formats at regular time intervals and covering large regions with high spatial resolution. In combination with new retrieval methods, this has led to a concept for a next-generation instrument that will image/monitor, from geostationary orbit, the thermosphere and the ionosphere during the daytime and the ionosphere at night with better than 10 km resolution at approximately 100 second image cadence. Innovations in retrieval methods and the status of instrument will be described, along with how the calculated performance would contribute to the scientific study of mid- and low-latitude ionospheric phenomena.

INTRODUCTION

This white paper describes the Ionospheric Mapping and Geocoronal Experiment (IMAGEr), a payload under development at the Naval Research Laboratory (NRL) for flight on a future United States Air Force Space Test Program launch. It has also been proposed to the National Aeronautics and Space Administration as a component of a Mission of Opportunity payload called TIGRIST, led by Cornell University in partnership with NRL and University of California at Berkeley. Under either realization, the instrument concept is identical, namely a steerable telescope designed to monitor ionospheric weather continuously and at high spatial and temporal resolution, from a geostationary platform, by observing any of four airglow lines in the far- and extreme-ultraviolet (83.4, 130.4, 135.6 and 155.0 nm), making movies in those lines with a frame rate of typically 100 s. Remote sensing methods support such a program for dayside and nightside airglow features, either at Earth limb or downward on the disk. Accumulated knowledge and techniques relating to specific lines and retrievals are heritage upon which IMAGEr draws.

The possibility and desirability – one might say the *dream* – of recording ionospheric weather in a cinematic yet diagnostic mode has been evident to ionospheric scientists for more the three decades, ever since Carruthers' Earth image from a far-UV camera on the moon appeared [Carruthers and Page, 1972]. The review by Meier [1991] summarizes the extensive information, available prior to 1990, on the uses of atmospheric airglow for monitoring the ionosphere and thermosphere and what spectral lines are seen and how they can be used to infer the atmospheric state. In the intervening decades since the first observations, techniques, both for sensors and retrievals, have been refined to bring the dream within reach. Instruments and systems such as Tiny Ionospheric Photometer (TIP) on the *Constellation Observing System for Meteorology, Ionosphere, and Climate* [Anthes et al., 2008; Dymond, Nee, and Thomas, 2000 a,b; Coker et al., 2008] satellites or the Global Ultra-Violet Imager (GUVI) on the NASA *Thermosphere, Ionosphere, Mesosphere Energetics and Dynamics* [Christensen et al., 2003] satellite, and the Ultra-Violet Imager (UVI) on the NASA IMAGE satellite [Immel et al., 2006; Sagawa et al., 2005] have had great scientific success in improving understanding of the nighttime ionosphere. The UV instruments on these satellites all observed the 135.6 nm emission produced at night primarily by radiative recombination of O^+ ions and electrons, a natural decay process of the ionosphere. The 135.6 nm emission is visible during the daytime and is very bright and easily observable; however, during the daytime it is dominated by photoelectron impact excitation and not useful as an ionospheric diagnostic [Meier, 1991]. Daytime ionospheric measurements have been made using the COSMIC satellites each of which carries a GPS occultation sensor (GOX) on-board [Anthes et al., 2008], which has both daytime

and nighttime ionospheric coverage. But, GPS occultations are sporadic, almost random, by nature and the sampling is particularly underrepresented in the equatorial region. Also, the long slant paths through the ionosphere cause inaccuracies due to along line-of-sight electron density gradients that are difficult to remove due to the sporadic occultation coverage. The COSMIC satellites also lack sensors for the neutral atmosphere. Global observations have been made of the response of the thermosphere during the superstorm of October 30-31, 2003 using the GUVI sensor and interpreted using thermospheric global circulation models. Crowley *et al.* [2005] shows good agreement between the GUVI observations of the O/N₂ column density ratio produced by the global circulation model. Daytime measurements of the neutral density using the GUVI sensor on the TIMED satellite have been used in conjunction with ground-based GPS measurements so that some studies of the neutral thermosphere and ionosphere have been made, but these studies have been limited by the local time coverage of the GUVI sensor and the density of the GPS observations in some regions, especially over the oceans.

Two central themes set the IMAGER approach apart from LEO-based satellite systems and ground based GPS systems. One is observing strategy. There is a grave problem achieving the needed timescale coverage in monitoring done from one or more platforms in low earth orbit (LEO) due to re-visit time limitations. A LEO-based sensor will re-visit the same location on the globe only every few days. For LEO satellites, the local time of the equatorial crossing of the orbit drifts slowly so the sensors make essentially fixed-local time measurements. Fixed local time measurements are very useful for studies of the quiescent ionosphere and thermosphere as the insolation of the atmosphere/ionosphere is essentially constant or varying on scales typical of the solar rotation rate, ~28 days. However, when geomagnetic storms or solar flares occur, the externally induced drivers are functions of universal time not local time; so storm-time observations can be handicapped by the fixed local time observations from LEO-based sensors. Furthermore, it is not uncommon for the LEO satellite to be in the “wrong place” over the Earth to make a desired observation. *A sensor on a geosynchronous (GEO) platform overcomes these difficulties* while meeting other goals discussed below. The second theme is that IMAGER *not only provides thermospheric coverage with maps of the ratio of O and N₂ column densities ($\Sigma O/\Sigma N_2$) during the daytime but also gives coincident ionospheric diagnostics*. IMAGER therefore offers most of what one would naturally regard as an ideal comprehensive suite of measurements to cover ionospheric and thermospheric weather, as complete as any single instrument could provide at present.

SCIENCE OVERVIEW

The IMAGER experiment is designed to address several key science questions to improve our understanding of the coupled ionosphere–thermosphere system. The instrument parameters (spatial resolution, focal length, field-of-view, sensitivity, etc.) are all driven by either technical considerations, such as detector size, or by the science requirements. In this section, we provide an overview of the key science issues and discuss the constraints they place on the IMAGER system design.

What is the global response of the ionosphere and thermosphere to solar and geomagnetic forcing? The ionosphere-thermosphere system is a complex, highly coupled system that responds strongly to the influence of external drivers, such as solar extreme-ultraviolet flux variations. Geomagnetic effects, such as a charged particle precipitation in the auroral zones, deposit substantial energy in the lower thermosphere. The electric fields and charged particles associated with geomagnetic activity can substantially alter the global current distribution in the ionosphere on short time-scales, ~10 minutes. Thus, we desire temporal resolution on the order of minutes for the instrument.

How are ionospheric composition changes related to the underlying neutral thermosphere? The ionosphere is produced by photoionization of O. One of the principal recombination loss processes is charge exchange of the O⁺ with either N₂ or O₂ to produce N₂⁺ or O₂⁺. These species rapidly recombine with ambient electrons. The charge exchange process is strongly dependent on the N₂ and O₂ densities. The ratio of O to N₂ column densities has been observed to change during high geomagnetic activity. As this ratio decreases, the rate of photoionization of O decreases and the rate of recombination increases. Thus, the ionospheric electron density is expected to decrease because the production rate decreases and the loss rate increases. This has been observed using the *Dynamics Explorer* (DE) satellite [Strickland *et*

al., 2001]. The DE data however, do not have very high spatial resolution and the temporal resolution is not very high due to the high degree of averaging of the data required to get improve signal-to-noise. The IMAGER instrument will be able to perform studies like the DE studies at temporal resolution of approximately 300 seconds (3 images at 100 seconds each) and at a spatial resolution of 10 km over a 1000 km field-of-view.

What are the characteristics, behavior, and origin of ionospheric irregularities including recently observed mid-latitude depletions and storm enhanced densities? The mid-latitude nighttime ionosphere has long been thought to have little morphological structure with only weak density gradients. However, recent measurements of the ionosphere at mid-latitudes at night have shown that strong mid- to small-scale ionospheric irregularities exist. Recently, mid-latitude storm enhanced density (SED) structures have been observed. These SEDs have primarily been observed in the late afternoon and in the early evening in the American sector [Foster *et al.*, 2002, 2004, 2005]. Although more recent measurements using ground-based GPS measurements show SEDs to be a global phenomenon in conjugate hemispheres [Foster and Coster, 2007]. Recent measurements using the Tiny Ionospheric Photometer (TIP) on the COSMIC satellites have demonstrated the capability to observe mid-scale traveling ionospheric disturbances from space [Dymond *et al.*, 2008b]. The sensitivity of the IMAGER instrument is comparable to that of the TIP sensors and, with the capability to stare at a region, it should be possible to study the wavelengths of the MSTIDs and to also image the wavefronts. This will provide additional information about the global occurrence and behavior of MSTIDs.

How do the ionosphere and thermosphere respond to tropospheric influences? Recent observations have confirmed the influence of tropospheric tides on the ionosphere [Sagawa *et al.*, 2005; Immel *et al.*, 2006; England *et al.*, 2006; Kil *et al.*, 2007; McDonald *et al.*, 2008a; Lin *et al.*, 2007; Coker *et al.*, 2008; Dymond *et al.*, 2008]. These tidal influences are driven by the non-migrating diurnal tides [Immel *et al.*, 2006; Hagan *et al.*, 2007; Forbes *et al.*, 2008], which are forced by latent heat release associated with tropical thunderstorms. The aforementioned studies all relied on measurements made by LEO satellites that were averaged over a month or more. It should be possible to observe the diurnal development of the ionosphere in the equatorial anomaly region from a GEO platform and to study the diurnal effects of the tidal forcing. Presumably, the effects of large tropospheric storms, such as hurricanes, can also be reflected in changes to ionospheric morphology. The effects of large tropospheric storms on the ionosphere and thermosphere cannot be observed from the ground, as the storm precludes gathering data. IMAGER will be able to look for ionospheric and thermospheric perturbations caused by tropospheric weather.

OBSERVABLES

Daytime: The principal ionospheric observable during the daytime will be the O II 83.4 nm lines. The O II 83.4 nm emission is produced principally by photoionization excitation of O at altitudes below 150 km. The 83.4 nm photons emitted downward are absorbed by the atmosphere. Those photons emitted upward enter the F-region ionosphere (altitudes 200–600 km) where they undergo multiple resonant scattering by F-region O^+ ions. During the scattering process, the photons are entrapped and radiance field picks up a signature that is dependent on the altitude distribution of the O^+ ions. Dymond [2008] developed an algorithm using the suggestion of Picone [2005] that, for the first time, enables the inversion and retrieval of ionospheric information from observations of the O II 83.4 nm emission. This is described further below.

The experiment will observe the O I 135.6 nm and O I 130.4 nm emissions from neutral oxygen to infer the O density and observe the emission from N_2 in the Lyman-Birge Hopfield bands near 155 nm. On the dayside, all three emissions are primarily excited by photoelectron impact. However, the O emissions have contributions from radiative recombination and $O^+ - O^-$ mutual neutralization. The O I 130.4 nm emission is also excited by fluorescent scattering of sunlight at 130.4 nm. The N_2 LBH bands

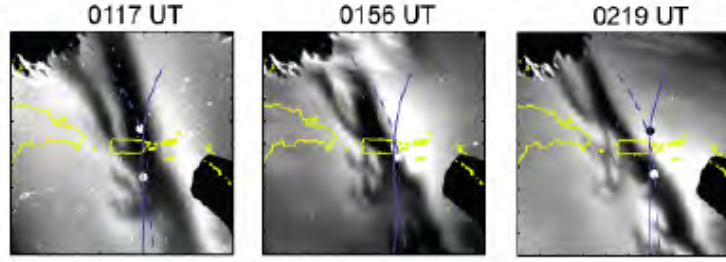


Figure 1: An Ionospheric Irregularity Over Puerto Rico. This figure shows snapshots of an irregularity that formed on Nov. 23, 2002 over Puerto Rico and the Caribbean Sea. The observations were made by an all-sky imager sited at the Arecibo Radio Observatory that observed the forbidden O I 630 nm emission. Cuba and Puerto Rico are indicated by the yellow lines. The dark regions in the upper left and lower right corners of the images are from obstructions near the camera. The images subtend $\sim 1000 \text{ km} \times 1000 \text{ km}$. The irregularity moves in a southwesterly direction across the image.

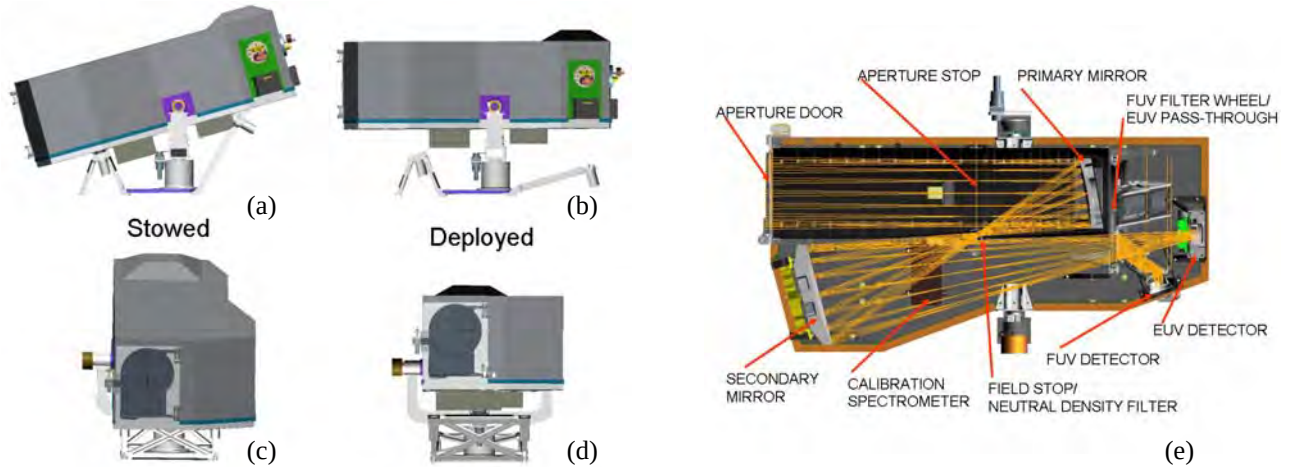


Figure 2: The IMAGER Instrument. External views of the IMAGER instrument are shown in the four views at the left. The upper views show the telescope stowed and then deployed for observations. The bottom two views show a “end-on” views of the instrument stowed (left) and deployed (right). The internal layout of the instrument is shown in the right-most picture. The internal components are labeled. The calibration spectrometer is an optional sensor that might be flown to track the ultraviolet spectrum seen by the instrument to track degradation of the IMAGER optical components with time.

near 155 nm are near the peak of the O₂ Schumann-Runge continuum [Meier, 1991], which will allow the retrieval of the O₂ density by inversion of the limb measurements.

Nighttime: The instrument will observe both the O I 130.4 and O I 135.6 nm emissions at night. Both of these emissions are excited primarily by radiative recombination of O⁺ and electrons. A weaker component comes from O⁺ – O⁻ neutralization. The O⁺ and therefore the electron density (by charge neutrality) can be inferred by inversions of the nighttime radiance.

INSTRUMENT DESCRIPTION

The instrument design was driven by three basic requirements. First, the integration time to gather an image was selected to be approximately 2 minutes. Secondly, the experiment is expected to provide precise ionospheric specification, with 3 σ or better statistical uncertainty, over 100 km areas for the nighttime mid-latitude ionosphere. The minimum radiance assumed for the mid-latitude ionosphere was 0.1 Rayleigh, based on calculations using the IRI-90 [Bilitza, 1990] for the electron densities and the known constants from Melendez-Alvira *et al.* [1999]. The nighttime ionospheric sensing requirement

drove the aperture size and the instrument's sensitivity. Additionally, the instrument is required to have 10 km spatial resolution to be able to resolve the structure of the mid- and low- latitude ionospheric irregularities, shown in Figure 1. These irregularities show spatial structures from the sub-kilometer level upward to the several thousands of kilometers. The equatorial irregularities propagate Eastward and the mid-latitude irregularities propagate Southwesterly; both travel at 10–100 meters per second; images of these phenomena with 100 second exposure time will have <10 km blur. The electron densities, and therefore the radiances, vary from the levels present in the unperturbed ionosphere to nearly zero inside the depleted regions, as shown in Figure 1. So, in summary, the instrument will have better than 10 km spatial resolution over a 1000 km field-of-regard and its field-of-regard can be pointed anywhere within its field-of-view which is the entire visible portion of the globe from geostationary orbit. It can observe with at least 100 second image cadence. The IMAGER instrument hardware has been described in greater detail elsewhere [Dymond *et al.*, 2002; Wood, 2004; Kalmanson *et al.*, 2005].

Imaging Subsystem: The instrument consists of an off-axis Gregorian telescope with a parabolic primary mirror feeding an off-axis ellipsoidal secondary mirror. The Gregorian telescope is recognized as a compact telescope design with excellent off-axis rejection properties, which is important as the instrument is required to image the nighttime ionosphere near the bright sunlit terminator. Light is reflected from the primary mirror and passes to the secondary mirror and then through a three-reflection narrow passband filter and is imaged onto the input face of an imaging microchannel plate detector (Figure 2e).

The filters are three reflection bandpass filters developed for and flown on the Ultraviolet Imager (UVI) on the NASA POLAR satellite [Zuckic *et al.*, 1993]. These filters have approximately 2.5 nm passbands (FWHM) and peak reflectances of 30–50%. The three far ultraviolet filters (130.4, 135.6, and 155 nm) are contained in a filter wheel assembly that allows selection of the wavelength of interest. An opening in the filter wheel allow light to pass directly through an indium filter to a windowless imaging microchannel plate detector that measures the intensity of the 83.4 nm emission. Figure 2 shows the external and internal layout of the imaging system.

We estimated the instrument's sensitivity for the design described above based on previously measured reflectance and quantum efficiency data from Siegmund, Vallerger, and Tremsin [2005]. The 130.4, 135.6, and 155 nm passbands all have a sensitivity of $\sim 300 \text{ ct}^{-1} \text{ s}^{-1} \text{ Rayleigh}^{-1}$. Thus, during typical daytime conditions at mid-solar cycle and local noon the instrument should produce images with $\sim 1\%$ noise due to counting statistics in the 135.6 and 155 nm passbands and with $\sim 0.5\%$ noise at 130.4 nm during a 100 second exposure with 10 km spatial resolution. The 135.6 and 130.4 nm channels will measure 30 counts per 10 km pixel during a 100 s exposure for $n_m F_2 \sim 10^6 \text{ cm}^{-3}$, typical of the densities observed in the Appleton anomaly in the evening.

The EUV channel's sensitivity is much lower than that of the FUV channels. It has been estimated to be $\sim 0.5 \text{ ct}^{-1} \text{ s}^{-1} \text{ Rayleigh}^{-1}$, using values of the MgF_2 over aluminum reflectance [Hunter, Osantowski, and Haas, 1971], the transmittance of a thin-metal indium film [Palik, 1998], and cesium-iodide photocathode quantum efficiency from Siegmund, Vallerger, and Tremsin [2005]. The greatest concern for the EUV channel is contamination by other airglow lines in the instrument's passband. We have derived a non-83.4 nm contamination of $\sim 5\%$, using the estimated instrument sensitivity and the measured airglow intensities taken from Bishop and Feldman [2003]. A study is under way to ascertain the effects of this systematic error on the ionospheric parameters derived from the 83.4 nm inversion algorithm discussed below.

Gimbal: We will use a two-axis gimbal to point the instrument. This provides the minimum volume for the imager, as the number of reflections is minimized. The Earth subtends approximately 17° , when viewed from geostationary orbit. The gimbal is designed to slew from limb to limb, an angular travel of 25° , in 5 minutes; the additional angular coverage allows stellar calibration on-orbit. The gimbal must be able to point center of the instrument's field-of-view to a point on the globe with accuracy of $\sim 20 \text{ km}$ to accomplish the mission's imaging objectives. Studies are underway to assess replacing the gimbal with a gimbaled scan mirror.

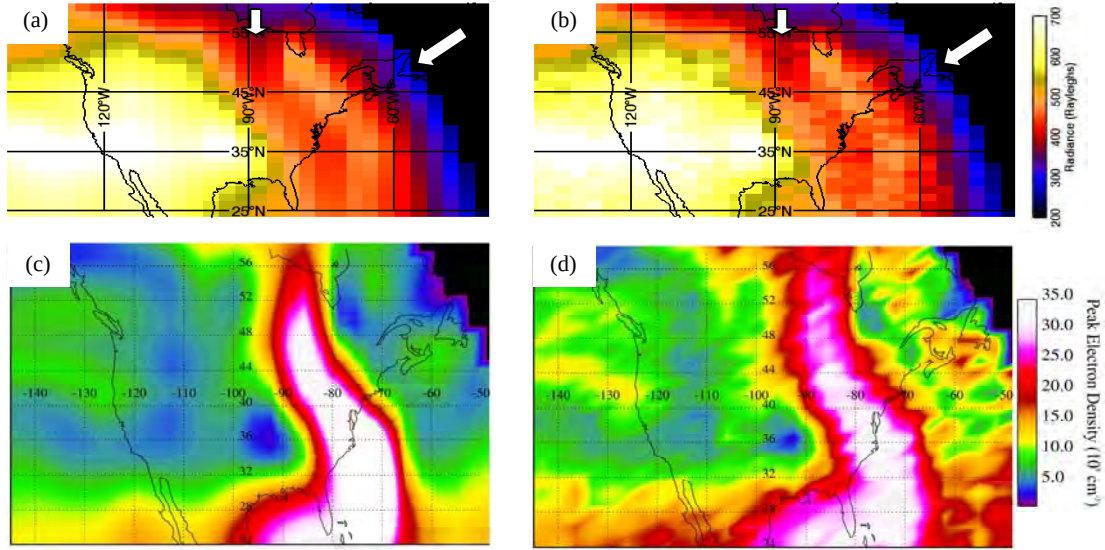


Figure 3: 83.4 nm Algorithm Simulation. Panel (a) shows the 83.4 nm radiance simulated for the SED over the Eastern US on Nov. 20, 2003 at 20 UT. Panel (b) shows the radiance with realistic measurement noise superimposed. The small downward pointing arrows near the top of the panels indicates the SED. The tilted arrows at the upper right of the panels indicated the terminator at 300 km altitude. The local time in the images varies from ~11 Hr at the western extents of the images to ~17 Hr at the right. Notice that the high densities within the SED *reduce* the intensity in the images. Panel (c) shows the peak density, nmF2, in the SED taken from the GAIM model, which had assimilated data over the continental US. Panel (d) shows the nmF2 retrieved by inversion of the image Panel (b). The overall agreement is very good with much of the structure of the underlying ionosphere being retrieved and the SED structure being well-resolved. The effects of the photon shot noise show up as small-scale fluctuations of the density. The retrieval tends to slightly over-estimate the nmF2 except in the SED, where it tends to underestimate slightly.

Ionospheric Retrieval Algorithms

Coincident measurement of the neutral thermosphere and the ionosphere during the daytime on a large, global scale has not yet been achieved. This important set of measurements is not possible even using the dayside measurements of the GUVI instrument on the TIMED satellite and the ground-based GPS network. The ionosphere is not only affected by the photochemical production and loss, as characterized by the O/N_2 column density ratio, but is also driven by transport. This transport is particularly important in the low- and mid-latitude regions and is of paramount importance during storm-time events. The capability to simultaneously measure the neutral thermosphere and the ionosphere will provide exciting opportunities to understand the dynamics of the thermosphere and ionosphere that cannot be accomplished from low-Earth orbiting systems.

Inversion of the 83.4 nm O^+ feature on the day side is a longstanding challenge that has recently radiance, the radiance with realistic instrumental noise superimposed, the n_mF2 from the input storm enhanced density (SED) plume, and the retrieved n_mF2 using the new algorithm. For this simulation, the SED plume was taken from a run of the GAIM model, in the regional, Kalman filter mode, which had assimilated ionospheric data globally [Schunk *et al.*, 2004 a, b]. The NRLMSISE-00 model [Picone *et al.*, 2002] was used to calculate the neutral thermospheric densities that were used to calculate the initial source of excitation for the 83.4 nm emission and the pure absorption by O, O_2 , and N_2 that affects the *al.*, 2002] was used to calculate the neutral thermospheric densities that were used to calculate the initial source of excitation for the 83.4 nm emission and the pure absorption by O, O_2 , and N_2 that affects the radiation transport of the 83.4 nm emission. The agreement between the input SED n_mF2 (Figure 3c) and the n_mF2 retrieved by the algorithm (Figure 3d) is remarkable and demonstrates the validity of the algorithmic approach. The capability to sense the daytime ionosphere coincident with measurements of the neutral density opens up new possibilities to explore the coupled ionosphere-thermosphere system.

Another newly developed algorithm will be used by the IMAGER to characterize the nighttime ionosphere. This technique is based upon earlier work by McCoy and Anderson [1984] who realized that

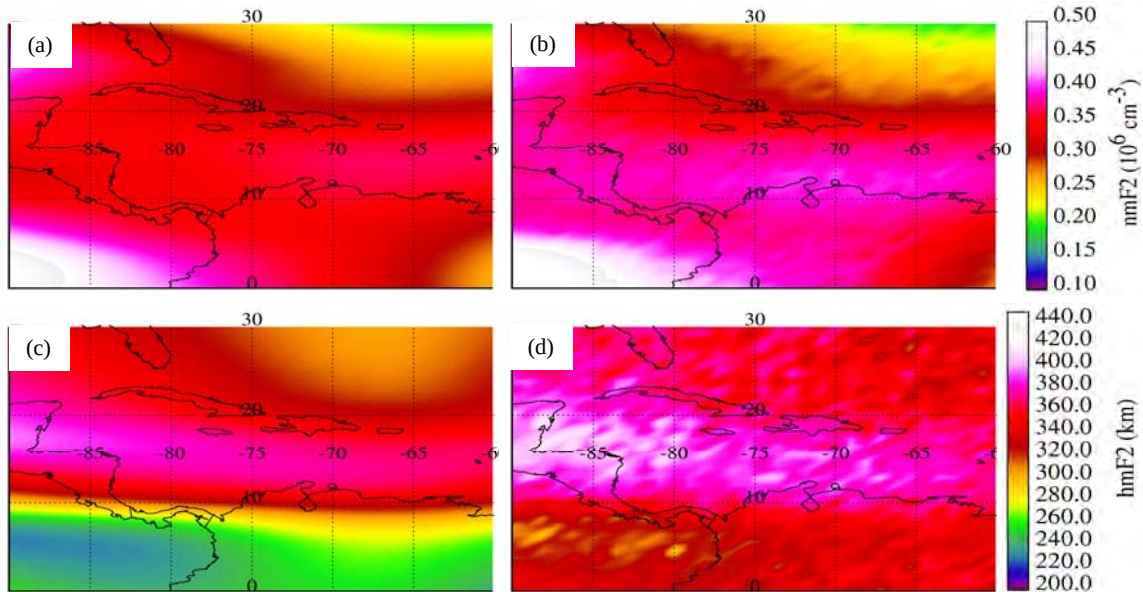


Figure 4: 130.4/135.6 Algorithm GAIM Simulation. Panel (a) shows the GAIM (input) peak density over the Caribbean and panel (b) shows the retrieved peak density. The agreement is excellent with only small fluctuations due to the photon shot noise superimposed on the radiances. Panels (c) and (d) show the GAIM peak height (input) and the retrieved peak height, respectively. Although the agreement between the images is very good, there is a +27 km bias in the retrieved peak heights. The peak height in the lower left corners of the images shows the poorest agreement. The peak height that GAIM produced in this area is ~220-230 km. This low altitude was not reachable by the algorithm as the peak height was constrained to be > 250 km and this may have skewed the inversion.

that the overlap of the ionosphere with the neutral thermosphere, in particular overlap with the neutral oxygen profile, would modulate the intensity ratio of the 130.4 and 135.6 nm emissions rendering it sensitive to the height of the ionosphere. *Dymond* [2009] modernized the *McCoy and Anderson* algorithm to create and demonstrate the new algorithm. Figure 4 summarizes the results of the *Dymond* [2009] work. *Dymond* simulated the radiances of the O I 130.4 and O I 135.6 nm emission using the electron density from a run of the regional version of the GAIM model which had assimilated global ionospheric measurements. The neutral atmosphere was taken from the NRLMSISE-00 [*Picone et al.*, 2002] appropriate to the geomagnetic and solar conditions. Realistic instrumental noise was superimposed upon the radiances using the sensitivities given above and the noisy radiances were inverted using a non-linear least squares technique to retrieve the n_mF2 and the h_mF2 . Figure 4a shows the n_mF2 from the GAIM run used to simulate the “data” and Figure 4b shows the retrieved n_mF2 ; the agreement is excellent. Figure 4c shows the h_mF2 from the GAIM run and Figure 4d shows the retrieved h_mF2 ; the agreement is very good with the retrieved h_mF2 showing a systematic error of +27 km. However, when the bias is subtracted, the agreement is very good. The capability to globally sense the n_mF2 and h_mF2 will enable new scientific studies of the low- and mid-latitude ionosphere, such as coincident measurement of the n_mF2 and h_mF2 in regions containing ionospheric bubbles produced by the Rayleigh-Taylor instability.

Summary

IMAGER is designed to sense the ionosphere during the daytime and at night, disk, and limb. It can also image the neutral thermosphere on the dayside limb and disk. The day side ionospheric imagery relies upon the 83.4 nm line of O^+ using a new algorithm and also utilizes two other FUV lines to determine the O to N_2 ratio. The night side imagery uses the 135.6 nm line to image ionospheric features and disturbances, and combines this with measurements of the radiance in the 130.4 nm line to constrain vertical distributions and extract n_mF2 and h_mF2 . This work, and particularly the design and prototyping effort on IMAGER, was generously supported by the Office of Naval Research.

References

- Anthes, R. A., P. A. Bernhardt, Y. Chen, L. Cucurull, K. F. Dymond, D. Ector, S. B. Healy, S.-P. Ho, D. C. Hunt, Y.-H. Kuo, H. Liu, K. Manning, C. McCormick, T. K. Meehan, W. J. Randel, C. Rocken, W. S. Schreiner, S. V. Sokolovskiy, S. Syndergaard, D. C. Thompson, K. E. Trenberth, T.-K. Wee, N. L. Yen, and Z. Zeng, The COSMIC/FORMOSAT-3 Mission: Early Results, *Bull. Amer. Meteor. Soc.*, 313-333, DOI:10.1175/BAMS-89-3-313, 2008.
- Bilitza, D., *International Reference Ionosphere 1990*, National Space Science Data Center, NSSDC/WDC-A-R&S 90-20, Greenbelt, Maryland, 1990.
- Bishop, J., and P. D. Feldman, Analysis of the Astro-1/Hopkins Ultraviolet Telescope EUV–FUV dayside nadir spectral radiance measurements, *J. Geophys. Res.*, **108**(A6), 1243, doi:10.1029/2001JA000330, 2003.
- Carruthers, G. R., and T. Page, Apollo 16 Far-Ultraviolet Imagery of the Polar Auroras, Tropical Airglow Belts, and General Airglow, *J. Geophys. Res.*, **81**, 483, 1972.
- Christensen, A. B., L. J. Paxton, S. Avery, J. Craven, G. Crowley, D. C. Humm, H. Kil, R. R. Meier, C.-I. Meng, D. Morrison, B. S. Ogorzalek, P. Straus, D. J. Strickland, R. M. Swenson, R. L. Walterscheid, B. Wolven, and Y. Zhang (2003), “Initial Observations with the Global Ultraviolet Imager (GUVI) in the NASA TIMED Satellite Mission”, *J. Geophys. Res.*, **108** (A12), 1451, doi:10.1029/2003JA009918.
- Coker, C., K. F. Dymond, S. A., Budzien, D. H. Chua, J.-Y. Liu, D. N. Anderson, Su. Basu, and T. R. Pederson, Observations of the Ionosphere Using the Tiny Ionospheric Photometer, *Terrestrial, Atmospheric & Oceanic Sciences (TAO)*, In press, 2008.
- Crowley, G., R. Meier, D. Strickland, B. Tapley, S. Bettadpur, M. Cheng, X. Pi, L. Paxton, A. Christensen, Thermosphere-Ionosphere Response During the October-November 2003 Superstorms, *Proceedings of the 2005 Ionospheric Effects Symposium*, Alexandria VA, 2005.
- Dymond, K. F., J. B. Nee, and R. J. Thomas, The Tiny Ionospheric Photometer: An Instrument For Measuring Ionospheric Gradients For The COSMIC Constellation, *Terr., Atmos. Ocean. Sci. (TAO)*, Vol. **11**, No. 1, 273-290, 2000a.
- Dymond, K. F., J. B. Nee, and R. J. Thomas, The Tiny Ionospheric Photometer: An Instrument For Measuring Ionospheric Gradients For The Cosmic Constellation, in *Applications of the Constellation Observing System for Meteorology, Ionosphere, and Climate*, p 273-290, Springer, New York, New York, 2000b.
- Dymond, K. F., R. P. McCoy, K. S. Wood, S. A. Budzien, A. C. Nicholas, S. E. Thonnard, M. C. Kelley, and J. J. Makela, An Experiment for Studying Spatial and Temporal Behavior of the Ionosphere, *Optical Spectroscopic Techniques, Remote Sensing, and Instrumentation for Atmospheric and Space Research IV*, Allen M. Larar, Matrin G. Mlynchak, Editors, *Proceedings of SPIE* Vol. 4485, 266-273, 2002.
- Dymond, K. F., A New Algorithm for Sensing the Global Daytime Ionosphere from Space, *Proceedings of the 2008 Ionospheric Effects Symposium*, Alexandria, VA, May 13-15, 2008.
- Dymond, K. F. (2009), “Remote sensing of nighttime *F* region peak height and peak density using ultraviolet line ratios”, *Radio Sci.*, **44**, RS0A28, doi:10.1029/2008RS004091.
- Dymond, K. F., S. E. McDonald, C. Coker, S. A. Budzien, and D. H. Chua, COSMIC Observations of the Temporal Evolution of the Low-latitude Ionosphere, *Proceedings of the 2008 Ionospheric Effects Symposium*, Alexandria, VA, May 13-15, 2008a.
- Dymond, K. F., C. Watts, C. Coker, P. Bernhardt, N. Kassim, P. Ray, J. Lazio, A. Cohen, K. Weiler, P. Crane, L. J. Rickard, G. Taylor, F. Schinzel, Y. Philstrom, S. Close, P. Colestock, S. Myers, and A. Dhatta, The Combined Radio Interferometry and COSMIC Experiment in Tomography (CRICKET) Campaign, *Proceedings of the 2008 Ionospheric Effects Symposium*, Alexandria, VA, May 13-15, 2008b.
- England, S. L., T. J. Immel, E. Sagawa, S. B. Henderson, M. E. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton (2006), Effect of Atmospheric Tides on the Morphology of the Quiet Time,

- Postsunset Equatorial Ionospheric Anomaly, *J. Geophys. Res.*, **111**, A10S19, doi:10.1029/2006JA011795.
- Forbes, J. M., X. Zhang, S. Palo, J. Russell, C. J. Mertens, and M. Mlynczak (2008), Tidal Variability in the Ionospheric Dynamo Region, *J. Geophys. Res.*, **113**, A02310, doi:10.1029/2007JA012737.
- Foster, J. C., P. J. Erickson, A. J. Coster, J. Goldstein, and F. J. Rich (2002), Ionospheric signatures of plasmaspheric tails, *Geophys. Res. Lett.*, **29**(13), 1623, doi:10.1029/2002GL015067.
- Foster, J. C., A. J. Coster, P. J. Erickson, F. J. Rich, and B. R. Sandel (2004), Stormtime observations of the flux of plasmaspheric ions to the dayside cusp/magnetopause, *Geophys. Res. Lett.*, **31**, L08809, doi:10.1029/2004GL020082.
- Foster, J. C., and A. J. Coster, Conjugate Localized Enhancement of Total Electron Content at Low Latitudes in the American Sector, *J. Atm. Sol. Terr. Phys.*, **69**, 1241–1252, 2007.
- Foster, J. C., et al. (2005), Multiradar observations of the polar tongue of ionization, *J. Geophys. Res.*, **110**, A09S31, doi:10.1029/2004JA010928.
- Hagan, M. E., A. Maute, R. G. Roble, A. D. Richmond, T. J. Immel, and S. L. England (2007), Connections Between Deep Tropical Clouds and the Earth's Ionosphere, *Geophys. Res. Lett.*, **34**, L20109, doi:10.1029/2007GL030142.
- Hunter, W.R., Osantowski, J.F., and Hass, G., Reflectance of Aluminum Overcoated with MgF₂ and LiF in the Wavelength Region from 1600 Å to 300 Å at Various Angles of Incidence, *Appl. Opt.*, **10**, 540-544 (1971).
- Immel, T. J., E. Sagawa, S. L. England, S. B. Henderson, M. E. Hagan, S. B. Mende, H. U. Frey, C. M. Swenson, and L. J. Paxton (2006), Control of Equatorial Ionospheric Morphology by Atmospheric Tides, *Geophys. Res. Lett.*, **33**, L15108, doi:10.1029/2006GL026161.
- Kalmanson, P. C., J. Wilczynski, K. Wood, K. Dymond, S. Thonnard, and J. Spann, The Optomechanical Design and Operation of the Ionospheric Mapping and Geocoronal Experiment, *Proceedings of SPIE Volume: 5901*, Editor(s): Silvano Fineschi, Rodney A. Viereck, 2005.
- Kil, H., S.-J. Oh, M. C. Kelley, L. J. Paxton, S. L. England, E. Talaat, K.-W. Min, and S.-Y. Su (2007), Longitudinal Structure of the Vertical $\mathbf{E} \times \mathbf{B}$ Drift and Ion Density Seen from ROCSAT-1, *Geophys. Res. Lett.*, **34**, L14110, doi:10.1029/2007GL030018.
- Lin, C. H., W. Wang, M. E. Hagan, C. C. Hsiao, T. J. Immel, M. L. Hsu, J. Y. Liu, L. J. Paxton, T. W. Fang, and C. H. Liu (2007), Plausible Effect of Atmospheric Tides on the Equatorial Ionosphere Observed by the FORMOSAT-3/COSMIC: Three-Dimensional Electron Density Structures, *Geophys. Res. Lett.*, **34**, L11112, doi:10.1029/2007GL029265.
- McCoy, R. P., and D. E. Anderson, Jr., Ultraviolet remote sensing of the F2 ionosphere, Effect of the Ionosphere on C3I Systems; Ionospheric Effects Symposium, 295-302, 1984.
- McDonald, S. E., K. F. Dymond and S. E. Summers (2008a), Hemispheric Asymmetries in the Longitudinal Structure of the Low-Latitude Nighttime Ionosphere, *J. Geophys. Res.*, in press.
- Meier, R. R., Ultraviolet Spectroscopy and Remote Sensing of the Upper Atmosphere, *Space Science Reviews*, **58**, No. 1 & 2, 1-186, 1991.
- Melendez-Alvira, D. J., Meier, R. R., Picone, J. M., Feldman, P. D., and McLaughlin, B. M., Analysis of the Oxygen Nightglow Measured by the Hopkins Ultraviolet Telescope: Implications for Ionospheric Partial Radiative Recombination Rate Coefficients, *J. Geophys. Res.*, **104**, 14901, 1999.
- Palik, E.D. (ed.), Handbook of Optical Constants of Solids II, Boston, Academic Press, Inc., Harcourt Brace Jovanovich, Publishers, 1998.
- Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 Empirical Model of the Atmosphere: Statistical Comparisons and Scientific Issues, *J. Geophys. Res.*, **107**(A12), 1468, doi:10.1029/2002JA009430.
- Picone, J. M., Patent Disclosure, Navy Case #97,745, October 27, 2005.
- Sagawa, E., T. J. Immel, H. U. Frey, and S. B. Mende (2005), Longitudinal Structure of the Equatorial Anomaly in the Nighttime Ionosphere Observed by IMAGE/FUV, *J. Geophys. Res.*, **110**, A11302, doi:10.1029/2004JA010848.

- Scherliess, L., D. C. Thompson, and R. W. Schunk (2008), Longitudinal Variability of Low-Latitude Total Electron Content: Tidal Influences, *J. Geophys. Res.*, **113**, A01311, doi:10.1029/2007JA012480.
- Schunk, R.W., L. Scherliess, J. J. Sojka, and D. Thompson, (2004a), Global Assimilation of Ionospheric Measurements (GAIM), *Radio Sci.*, **39**, RS1S02, doi:10.1029/2002RS002794.
- Schunk, R. W., L. Scherliess, J. J. Sojka, and D. C. Thompson, USU Global Ionospheric Data Assimilation Models, Proc. of SPIE, Vol. 5548, 0277- 768x104, doi:10.117/12.562448, 2004b.
- Siegmund, O., J. Vallerger, and A. Tremsin, Characterization of Microchannel Plate Quantum Efficiency, UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XIV edited by Oswald H. W. Siegmund, Proc. of SPIE Vol. **5898**, 58980H, doi: 10.1117/12.621708, 2005.
- Strickland, D. J., Craven, J. D., and Daniell, R. E., Global Enhancements in O/N₂ Recorded by the DE-1 FUV Imager During Geomagnetically Disturbed Periods, *Eos Trans. AGU*, **84** (20), Spring Meeting Suppl., Abstract SA62A-02, 2001.
- Wood, K. S., The Ionospheric Mapping and Geocoronal Experiment (IMAGER): A New System for Monitoring Ionospheric Space Weather, *Proceedings of the 84th American Meteorological Society (AMS) Annual Meeting*, Seattle, WA, Jan. 11-15, 2004. <http://ams.confex.com/ams/84Annual/techprogram/paper_72360.htm> (24 June 2008).
- Zuckic, M., Torr, D. G., Kim, J., Spann, J. F., and Torr, M. R., Filters for the International Solar Terrestrial Physics Mission Far-Ultraviolet Imager, *Opt. Eng.*, **32**, 3069, 1993.

Understanding the unique equatorial electrodynamics in the African Sector

Endawoke Yizengaw, Keith Groves, Tim Fuller-Rowell, Anthea Coster

Science Background

Satellite observations (see [Figure 1](#)) show unique equatorial ionospheric structures only in the African sector [[Hei et al., 2005](#); [Su, 2005](#); [Burke et al., 2006](#); [Yizengaw et al., 2010](#)]. It has been observed that in this region the bubbles are much deeper and occur more frequently than bubbles observed in any other longitudinal sector [[Hei et al., 2005](#) and [Su, 2005](#)]. It was also reported that the depletions in the African region rise to high altitudes (up to 1000+ km) more frequently compared to other longitudes [[Burke et al., 2004](#)]. However, these observations have not been confirmed, validated or studied in detail by observations from the ground due to lack of suitable ground-based instrumentation in the region, and the question of what causes or drives these unique density irregularities in the region is still not yet fully understood.

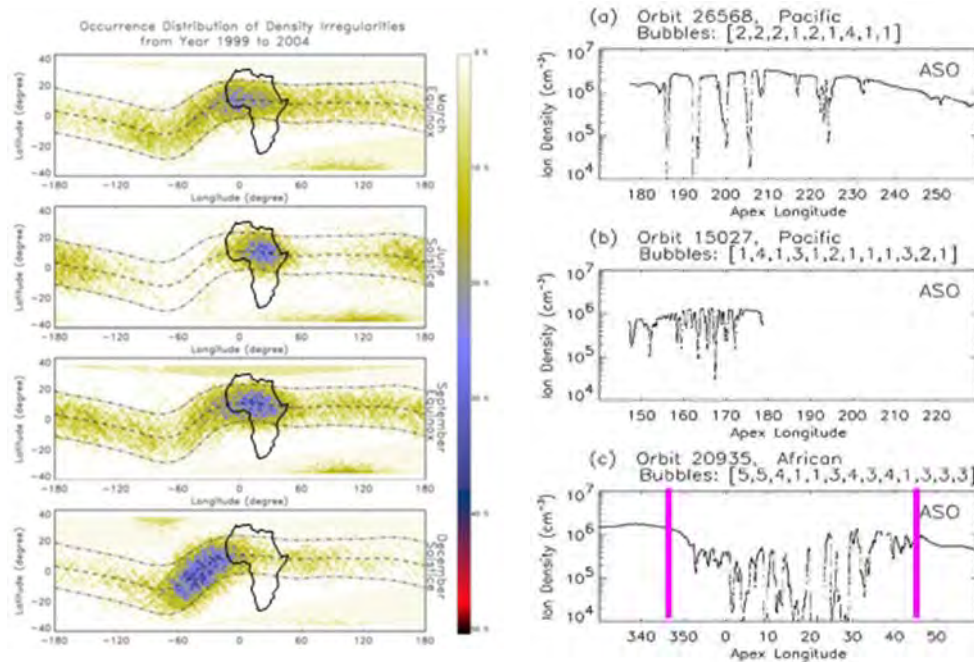


Figure 1. Satellite data suggest both the spatial structure and seasonal behavior of disturbances above Africa are unique. (left panels) seasonal behavior of the density irregularity occurrence probability using ROCSAT in situ density observation [[after Su, 2005](#)]. (right panels) the typical bubbles observation using AEE satellite in situ density, occurring in the fall (August, September, October (ASO)): (a) and (b) in the Pacific sector and (c) in the African sector [[after Hei et al., 2005](#)].

Despite the fact that much progress has been made in the study of ionospheric density structure and dynamics in the last decade, there are many gaps in our global understanding of the fundamental electrodynamics that governs equatorial ionospheric density irregularities. The uneven distribution of ground-based instruments has been the main barrier that hinders our ability to obtain a global understanding of the dynamics and structures of the ionosphere. Although ground-based instrumentations, such as GPS receivers and magnetometers are located in dense regional arrays, these arrays are primarily in North America and Europe, and only recently with the LISN network in South America. However, in regions like Africa, observations

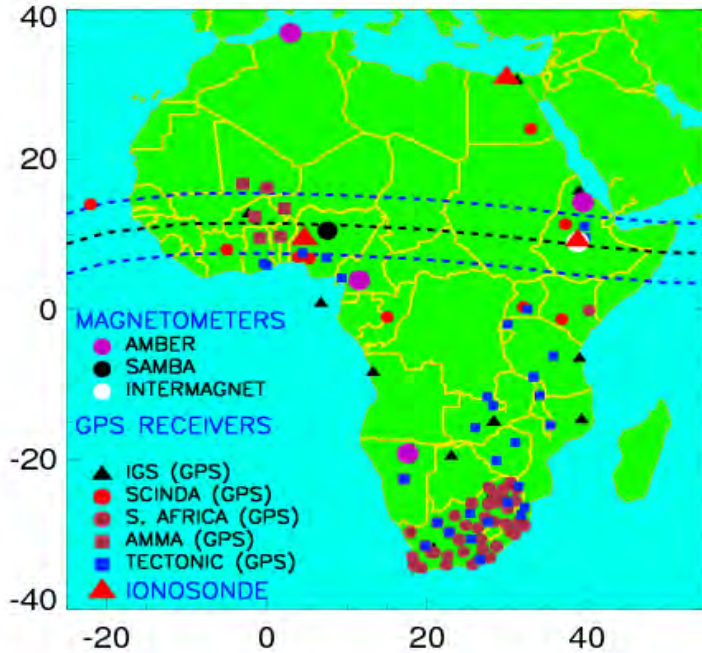


Figure 2 Available ground-based instruments in the African sector.

and navigation technologies depend on understanding, modeling, and mitigating the effects of these irregularities [Doherty *et al.*, 2004]. For example, when scintillation (the rapid amplitude and phase fluctuations of radio signals from space due to turbulence generated by ionospheric irregularities) [Groves *et al.*, 1997] happens the following technological systems will be affected: (1) regional SATCOM outages for extended periods (hours), (2) increased Global Positioning Satellite (GPS) navigation errors, (3) degraded High Frequency (HF) radio communication.

Required instrumentation

To have a complete and global understanding of the electrodynamics of the ionosphere in a cost-effective way, the deployment of small instruments, like ground-based GNSS receivers, is essential. GNSS receivers have a capacity to receive signals from GLONASS, COMPASS, EGNOS, in addition from the more popular GPS satellite constellations, providing higher accuracy and more reliable information in all environments. Therefore, for the upcoming decadal science objectives we propose the expansion of the ground-based instrumentation in the African sector, a region that has been severely lacked ground-based instrumentation for space science (see Figure 2). As can be seen in the figure, a significant number of GPS receivers have been deployed by the Tectonics community, but these receivers can also serve as important data resources for the space science community. However, the most dynamic region in terms of ionospheric irregularities, in the northern and equatorial part of Africa is still devoid of ground-based GPS receivers. Therefore, closing the largest land-based gaps in the northern and equatorial region of Africa will play a vital role in studying the fundamental governing electrodynamics of equatorial ionospheric motion.

In addition, augmenting the proposed GNSS deployment in this region with magnetometer networks will strengthen our investigations. This is because, magnetometer can provide information about the electrodynamics that governs the equatorial ionospheric motion [Anderson *et al.*, 2004]. Moreover, three ionosondes are either deployed or soon to be deployed in the equatorial region of the African continent. The beauty of these ionosondes is that they provide

of the ionosphere are currently not possible due to lack of ground-based instruments. Because of this, detailed investigation and validation of those unique dynamics and structure of the thermosphere and ionosphere in the African region cannot be done. The ionospheric density structure in the African sector has been traditionally estimated by model interpolation over vast geographic areas. This makes it very difficult to observe small-to-medium scale ionospheric irregularities, such as equatorial plasma bubbles, equatorial spread F (ESF), and bottom side spread F (BSSF) at low/mid-latitudes in the region. In fact, our communication

important geophysical parameters, such as winds, electric fields, conductivity, and density profiles, from the E- and F- ionosphere for the first time in the African sector [Valladares *et al.*, 2004; Abdu *et al.*, 2006]. Adding a single incoherent scatter radar (ISR) to this mix of space science instrumentation in the equatorial African sector would be of significant scientific benefit. ISR's deduce height- and time-resolved plasma drift velocities, electron and ion temperatures, electron densities, ion composition, and ion-neutral collision frequencies. These parameters provide further information about the neutral gas, neutral temperatures and winds, and electric fields present in the medium. No other instrument provides all of this ionospheric information. An ISR in Africa would greatly enhance the physics output of these other instruments and would provide information that would greatly enhance our understanding of equatorial electrodynamics. To date, the vast majority of equatorial ISR measurements have been collected at Jicamarca, so that much of what we know about equatorial physics is based on Jicamarca ISR observations. However, Jicamarca is in the American sector where the geomagnetic equator dips, and there is a fairly large excursion between the geomagnetic and geodetic equator. On the other hand, in the African sector the geomagnetic equator is fairly well aligned with the geodetic equator. New data from satellites (e.g. C/NOFS) have also indicated that the equatorial ionosphere in the African sector responds differently than other sectors. Therefore, having ISR in the African sector will significantly enhance our understanding of the physics behind the global equatorial electrodynamics.

Finally, there is a fast growing number of LEO satellites that are equipped with GPS receivers that will also augment the proposed ground-based GNSS network by providing additional information about the structure and dynamics of topside ionospheric and plasmaspheric density [Mannucci *et al.*, 2005; Yizengaw *et al.*, 2006]. In general, within the next decade the combined ground- and space-based observations will allow us, for the first time in the region, to answer the following longstanding specific questions.

Relevant Critical problems to be addressed

With the deployment of the proposed GPS receivers along with other available ground- and space-based multi-instruments, the space science community will be able to address the following main and major critical scientific problems.

1. What is special in Africa that produces unique ionospheric irregularities, such as the large-scale bubble properties (zonal width, depletion level, and spacing)? Is it because of special seeding conditions? Or because of special electrodynamics in the region?
2. Why does the depleted plasma (or bubbles) penetrate to higher altitudes in the African sector than in any other sectors? Does it relate to the ambient ionospheric plasma uplift? How often and under what conditions does vertical bubble penetration happen? What controls the lifetime of the plasma bubbles? What is the role of gravity waves in influencing the upward extension of the bubbles?
3. What are the possible governing mechanisms that create unique equatorial structures in Africa?
4. How intense and widely separated is the equatorial ionospheric anomaly (EIA) in the African sector? How often do such widely separated intense EIA structures occur in the African sector? Is it a year-round phenomenon as the occurrence of bubbles appears to be?
5. What is the role of the neutral atmosphere, including winds, tides and waves, in creating the unique equatorial electrodynamics of the African ionosphere? What are the characteristics of the neutral winds over Africa in the context of the global four-cell non-

migrating tide structure? Do neutral winds provide seeding that may account for the unique characteristics of African bubbles?

These questions are relevant broadly to the global ionospheric community, but given the intensity of local terrestrial weather forcing in Africa and the aforementioned characteristics of the African ionosphere, we believe it may be possible to establish an unambiguous connection between the terrestrial and space weather domains in Africa.

How to address the Critical problems?

These fundamental problems can be addressed by filling the critical ground-based data gaps. While the GPS and ionosondes observe the dynamics and structure of ionospheric density (such as bubbles and scintillation activities), magnetometers and ionosondes will estimate the driving force ($\mathbf{E} \times \mathbf{B}$ drift) simultaneously. More importantly, in addition to ground-based GPS receivers, in situ density observations track the bubbles at different altitudes; such data can be supplied from in situ density observations by LEO satellites, such as C/NOFS (400 x 800 km altitude) and DMSP (~850 km altitude). This will provide information about the altitude penetration of the bubbles which occurs quite often in the African sector [Burke *et al.*, 2004]. Moreover, with the proposed ISR the plasma drift velocities, electron and ion temperatures, electron densities, ion composition, and ion-neutral collision frequencies as function of altitude and time will be measured or estimated. The ISR would also provide the altitude structural images of the equatorial bubbles and ionospheric irregularities. We will also use the opportunity of the combined ground- and space-based instruments in the African region to compare equatorial electrodynamics with similar events at other longitudes and, specifically, the global response to major magnetic storms.

Impact on Research Infrastructure and Education

This proposal has significant broader impacts in creating international research collaborations. In addition to its scientific importance, this proposal will have other imperative impacts for both the U.S. and Africa. It has direct impact in advancing space science research into Africa by establishing and furthering sustainable research/training infrastructure within Africa so that more young scientists will be educated in their own country. This will not only enhance the research/training infrastructure but it will also play a vital role in the future socio-economic development of Africa.

For the United State, this project will serve as a vehicle to create opportunities for graduate and undergraduate students to participate in different phases of the development and deployment process. This will provide US students with high quality international research experiences, which will be quite important for the US scientific community to continue its leading international scientific stature in the future. The experience may provide opportunities for students to gain expertise in instrument calibration, deployment, data analysis, remote-sensing, computer modeling, data assimilation, and grid-based computing and data retrieval.

References

- Abdu, M. A., T. K. Ramkumar, I. S. Batista, C. G. M. Brum, H. Takahashi, B. W. Reinisch, and J. H. A. Sobral (2006), Planetary wave signatures in the equatorial atmosphere ionosphere system, and mesosphere- E- and F-region coupling, *J. Atmos. Solar-Terr. Phys.*, 68(3-5), 509-522.
- Anderson, D., A. Anghel, J. Chau, and O. Veliz (2004), Daytime vertical $E \times B$ drift velocities inferred from ground-based magnetometer observations at low latitudes, *Space Weather*, 2, S11001, doi:10.1029/2004SW000095.
- Burke, W. J., L. C. Gentile, C. Y. Huang, C. E. Valladares, and S. Y. Su (2004), Longitudinal variability of equatorial plasma bubbles observed by DMSP and ROCSAT-1, *J. Geophys. Res.*, 109, A12301, doi:10.1029/2004JA010583.
- Doherty, P., A. J. Coster, and W. Murtagh (2004), Space weather effects of October–November 2003, *GPS Sol.*, 8(3), doi:10.1007/s10291-004-0109-3.
- Groves, K. M., S. Basu, E. J. Weber, M. Smitham, H. Kuenzler, C. E. Valladares, R. Sheehan, E. MacKenzie, J. A. Secan, P. Ning, W. J. McNeill, D. W. Moonan, and M. J. Kendra (1997), Equatorial scintillation and systems support, *Radio Sci.*, 32(5), 2047–2064, doi:10.1029/97RS00836.
- Hei, M. A., R. A. Heelis, and J. P. McClure (2005), Seasonal and longitudinal variation of large-scale topside equatorial plasma depletions, *J. Geophys. Res.*, 110, A12315, doi:10.1029/2005JA011153.
- Mannucci, A. J., B. T. Tsurutani, B. A. Iijima, A. Komjathy, A. Saito, W. D. Gonzalez, F. L. Guarnieri, J. U. Kozyra, and R. Skoug (2005), Dayside global ionospheric response to the major interplanetary events of October 29–30, 2003 “Halloween Storms”, *Geophys. Res. Lett.*, 32, L12S02, doi:10.1029/2004GL021467.
- Valladares, C. E., R. Sheehan, and J. Villalobos (2004), A latitudinal network of GPS receivers dedicated to studies of equatorial spread F , *Radio Sci.*, 39, RS1S23, doi:10.1029/2002RS002853.
- Yizengaw, E., M. B. Moldwin, A. Mebrahtu, B. Damtie, E. Zesta, C. E. Valladares, and P. H. Doherty (2010), Comparison of storm time equatorial ionospheric electrodynamics in the African and American sectors, *J. Atmos. Solar-Terr. Phys.*, *in press*.
- Yizengaw, E., M. B. Moldwin, P. L. Dyson, and T. J. Immel (2005), Southern Hemisphere ionosphere and plasmasphere response to the interplanetary shock event of 29–31 October 2003, *J. Geophys. Res.*, 110, A09S30, doi:10.1029/2004JA010920.
- Yizengaw, E., M. B. Moldwin, P. L. Dyson, B. J. Fraser, and S. K. Morley (2006), First Tomographic Images of Ionospheric Outflows, *Geophys. Res. Lett.*, 33, L20102, doi:10.1029/2006GL027698.

Understanding upper atmospheric climate and change

A science white paper to the Decadal Review Committee

Authors: Shun-Rong Zhang¹ and John M. Holt (MIT Haystack Observatory)

Signatory: William L. Oliver (Center for Space Physics, Boston University)

Abstract

Both theoretical models and observational evidence have revealed that Earth's upper atmosphere is cooling and thermospheric density and ionospheric electron density are experiencing corresponding long-term changes. These changes will have strong influence on human activities. Addressing this topic of upper atmospheric climate and long-term change is timely and appropriate. It seems to us that the research focus for the next decade should include (1) to obtain a very comprehensive global pattern of the long-term change in both the ionosphere and the thermosphere. Neutral temperature at different heights from above the mesosphere through the upper atmosphere is crucial information; (2) to identify physical causes of the derived trend pattern; (3) to understand upper atmospheric climatology within a changing global environment. Continuously monitoring the upper atmosphere and solar-terrestrial environment remains an essential strategy toward a complete understanding of Earth's climate.

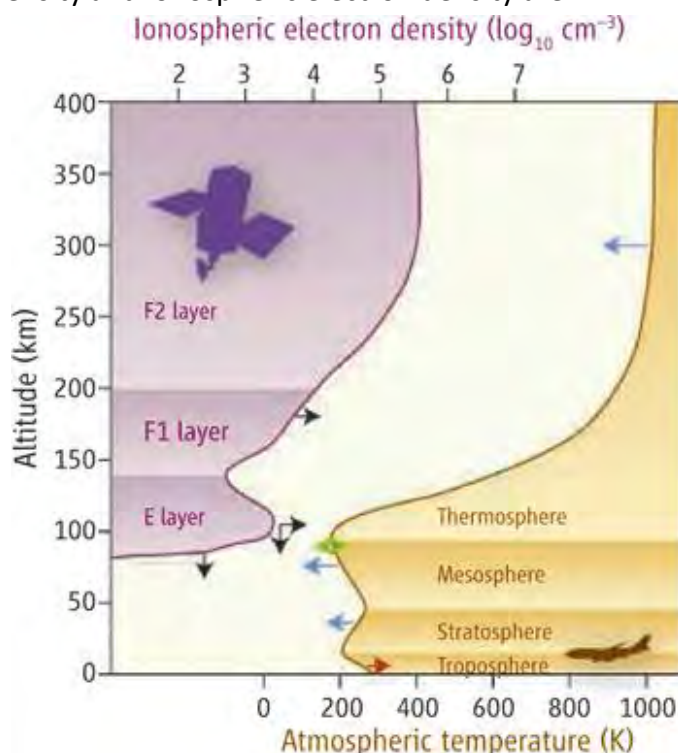


Figure 1 Pattern of ionospheric and thermospheric long-term trend derived from observations. Ionospheric electron density and thermospheric temperature are shown. Arrows to the right indicates "increase" and to the left "decrease". These patterns are qualitatively consistent with the greenhouse gas effect expected from theoretical

¹ Contact Info: email – shunrong@haystack.mit.edu; phone: 781-981-5725

1. Introduction: The falling sky

More and more evidence from recent observations in the ionosphere and thermosphere shows that the earth's upper atmosphere is falling during the recent a few decades. The falling, a clear indication of global climate change, is due to the cooling trend in the upper atmospheric temperature. Climate changes in the upper atmosphere have the potential to affect our increasing activity in space. Upper atmospheric cooling can cause a decrease in neutral density such that satellite orbits and lifetime are modified due to decreased air drag. It can also alter neutral composition such that radio propagation is affected due to changes in the ionospheric electron density. These effects exist whether the climate change is due to natural processes or anthropogenic increases in greenhouse gases.

A number of important topics are relevant to climate change in the upper atmosphere. Some of these are fundamental toward characterization of earth's environment and its direct impact on human activities. These topics include elaborating on the long-term trend in both the ionosphere and thermosphere upper atmospheric climatology, which may have undergone changes at various times scales. In this white paper, we review some of these compelling research topics that need to be addressed for the next decade. A separate white paper by Emmert et al. [2010] addresses a similar topic of geospace climate change, with more emphasis on implication of long-term climate changes on the neutral atmosphere and the need for long-term monitoring a series of important atmospheric parameters.

2. Long-term changes in the upper atmosphere

Greenhouse gases such as CO₂ and CH₄ are increasing in the lower atmosphere due to human activities. The effect of this on the upper atmosphere was first investigated in a theoretical modeling study by Roble and Dickinson [1989], suggesting a major greenhouse cooling in the thermosphere in response to increases in the CO₂ and CH₄ concentration at 60 km. Such a cooling effect causes a global reduction in neutral densities including O, N₂ O₂ and total neutral mass density as the neutral temperature T_n decreases. If greenhouse gas concentrations are doubled, as predicted to happen by the end of the 21st century, Roble and Dickinson indicated that the decrease in thermospheric temperature will be as much as 50 K, and the decrease in thermospheric densities at a fixed height between 200-300km will be 40-50%. The ionospheric consequences, as a result of thermal contraction, include a decrease in the F2 peak height, h_mF2 [Rishbeth, 1990; Rishbeth and Roble, 1992], a decrease in topside ionospheric density and an increase in the F1 region ionospheric density, with less change in the maximum density N_mF2. Some recent modeling studies using TIMEGCM and a more realistic CO₂ concentration (single site measurements) have essentially confirmed prior model findings [Qian et al., 2006; 2008].

The above results, arising from theoretical considerations, imply that ionospheric parameters can be a sensitive indicator of greenhouse gas atmospheric effects. Many subsequent investigations of long-term ionospheric and thermospheric change have occurred and the area is a very active research topic. Most of these studies are based on long-term ionosonde observations which date to the early days of ionospheric research and are readily accessible. Bremer [1992] published the first such ionospheric long-term trend observational results. Some, but not all, of these results concluded that hmF2 indicated a decreasing trend; other studies did not find a long-term trend or found a trend opposite to the sense predicted by Rishbeth and Roble [1992]. See Laštovička et al. [2006; 2008] for reviews of some recent study (see Figure 1). Observations of thermospheric total mass density by satellites revealed a 2-5% decrease per decade [Keating et al., 2000; Emmert et al., 2004; Emmert et al., 2008] (see Figure 2). These have been considered to be evidence of thermospheric cooling, supporting the Roble and Dickson [1989] theoretical calculations of the greenhouse effects.

Some direct evidence of temperature change has been provided by incoherent scatter radar observations. Zhang et al. [2005] provided evidence of the cooling in the ionospheric ion temperature for most altitudes and seasons based on Millstone Hill Incoherent Scatter Radar data. Holt and Zhang [2008] focused on Ti cooling for noontime at the altitude of 375 km, and provided a detailed analysis of confidence level of the estimated cooling trend. Donaldson et al. [2010] conducted an extended study of the ion temperature trend which includes height and time dependency using ISR data from observations made at St. Santin for a two-solar cycle period (1966-1987). Recently, Holt and Zhang [2010] and Zhang and Holt [2010] used Millstone Hill ISR data for the 1964-2010 period to determine the neutral temperature trend (exospheric temperature) and oxygen density variation, and height dependency of ion temperature (see Figure 3). Their direct temperature results revealed unambiguously the cooling in the upper atmosphere and its variability with height and solar activity.

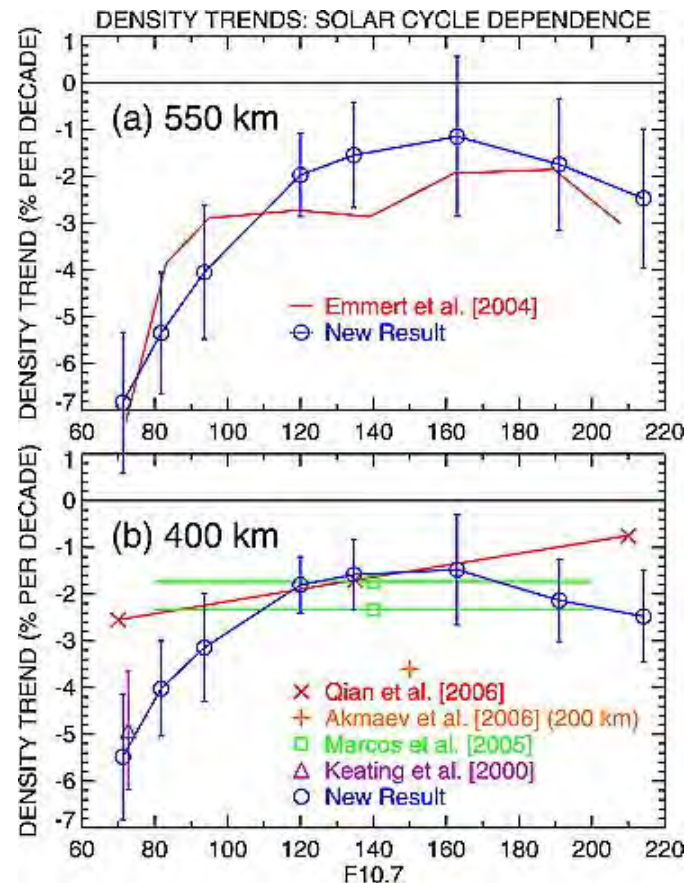


Figure 2 Long term trends in the neutral density at 550km (upper panel) and 400 km (lower panel) as a function of solar activity. (from Emmert et al., 2008)

3. Upper atmospheric climatology

Upper atmospheric climatology represents changes that occur repeatedly, persistently and predictably; therefore it may be considered as the background or reference conditions and the most essential information of the earth atmosphere. In the early years of ionospheric and thermospheric researches, many climatology studies focused on the so-called “anomaly” phenomena. Ionospheric electron density variations over time and space have traditionally been termed an anomaly if they do not follow the solar-controlled photochemistry dominated variation pattern.

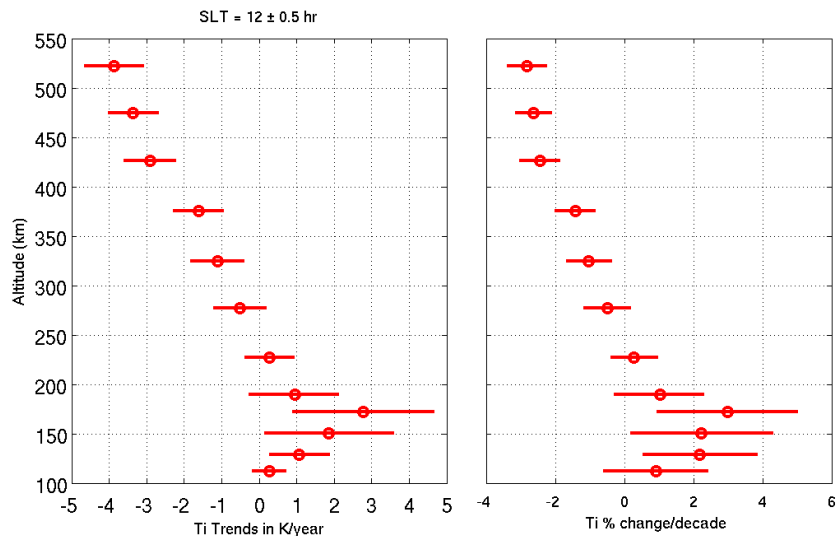


Figure 3 Height profile of long term trend in ion temperature (1964-2007) at noon over Millstone Hill. The cooling trend is significant in the upper F region and increases with height.

These included, for instance, the equatorial ionization anomaly (EIA), the seasonal anomaly or winter anomaly, and the Weddell Sea Anomaly. Ionosonde parameters were extensively studied. With a large number of observations from other instruments becoming available, especially satellite-based thermospheric and ionospheric density observations, many recent interests in our community have been on the subjects such as annual and semiannual variations which change with latitude, and the annual anomaly (north-south hemispheric asymmetry), etc.

Mendillo et al. [2005] noted the hemispheric asymmetry in the GPS TEC and ascribed it to corresponding changes in the neutral composition as a result of global circulation. Liu et al. [2007] reported the enhanced neutral density in the EIA regions. Qian et al. [2009] pointed out the importance of eddy mixing in the mesopause region to account for the observed annual/semiannual variations. Ground-based observations provide insights into local scale but detailed processes of upper atmospheric climatology. Using incoherent radars, Kawamura et al. [2000] investigated meridional winds patterns at mid-lower latitudes in Asia; Zhang et al. [2010] examined the height and latitudinal development of annual and semiannual variations of the ionosphere during the recent International Polar Year.

Empirical models are an important component of climatology studies and they are very useful tools for space weather and variability studies with engineering applications. Over the last decade, a number of models have been updated such as the horizontal wind model (HWM), the

mass spectrum and incoherent scatter (MSIS) model, and the International Reference Ionosphere (IRI) model. A number of new models have come online, such as incoherent scatter radar based ionospheric models (ISRIM) [Zhang and Holt, 2007], and the global average mass density model (GAMDM) [Emmert and Picone, 2010]

Our community still maintains very strong momentum, after many years of studies and numerous publications, in further characterizing and understanding this fundamental property of the upper atmosphere.

4. Important science topics

Under the big picture of global change in the whole earth's atmospheric system, addressing the following science topics in the next decade is timely and appropriate.

(1) Characterization of the long-term change in the upper atmosphere

There are still many unknowns in our understanding of ionospheric and thermospheric long-term trends. It is essential to generate a comprehensive view of long-term change in the thermosphere and ionosphere and to provide clues to understanding causative factors for upper atmospheric climate change. More specifically, we should determine long-term trends in neutral and plasma temperatures as direct evidence of climate change, along with trends in ionospheric and thermospheric densities as a result of climate change. It seems that the cooling trend varies with height from the mesopause to 500 km. The thermal status of the topside ionosphere and the E region ionosphere are different due to different sources of heating and cooling. In the 100-160 km height regime, where there exists significant coupling between the lower and upper thermosphere, neutral density and temperature are the key to understanding thermal response to greenhouse gas influences and the resulting ionospheric electron density changes. Other than height dependency, we also need to understand variability with solar activity, season, local time, and geographic location.

These studies will become more feasible in the next decade due to the much better data availability. Determination of the long-term trend requires removing various other variations in the data. This can only be possible with sufficient data (space-based and ground-based) that can clearly exhibit these variation patterns. Supporting solar-geophysical parameters, such as EUV flux, are also important.

(2) Understanding the causes of the long-term changes in the upper atmosphere

A number of possible causes have been proposed, including effects due to anthropogenic changes in greenhouse gases, the ozone layer, water vapor, and natural secular changes in solar-terrestrial magnetic activity, solar activity, and the earth's magnetic field. Indeed this type

of research provides a best case study that deals with some fundamental aeronomy questions relevant to the impact of the lower atmosphere. In the next decade, more investigations need to be carried out to examine all these drivers and to identify their relative importance using both additional data and modeling. Efforts need to be made to understand observed features inconsistent with the known general trend pattern in the ionosphere and thermosphere. Additional data include neutral and plasma density and temperature for various heights in the 100-500 km height range; neutral winds and plasma drift also provide important information on the dynamics. Measurements of neutral temperature and its height variation are the key to understanding changes in various ionosphere and thermospheric parameters. Global models with self-consistent coupling between the lower and upper thermosphere, and between the neutrals and the ions, are essential. Sophisticated theoretical modeling requires better specification of key reaction rates and the model boundary conditions, with an appropriate magnetic field model that includes secular changes and low thermospheric composition processes.

(3) The changing upper atmospheric climatology

It can be expected that the best known features of upper atmospheric climatology undergo some changes within the changing environment which modifies the general upper atmospheric circulation. This is because some of the “anomaly” climatology mentioned earlier, such as the winter anomaly, annual and semiannual change and the hemispheric asymmetry, is a result of neutral composition change caused by the circulation. An important task is to reassess those known phenomena and compare the early years to recent years and to further investigate future development. It is known that cooling in the upper atmosphere changes with location, height, solar activity, and local time. Therefore climate change in the upper atmosphere will show some variability as well. Studies such as the significant thermospheric density drop during the current extend solar minimum and behavior at upcoming solar maximum are examples that can allow for accurate definition of some basic characteristics of the earth’s atmosphere at extreme solar activity conditions. Long-term changes in the system should be taken into account in sophisticated upper atmospheric empirical models.

5. Summary

Global change in the lower atmosphere has imposed a significant impact on Earth’s upper atmosphere for both neutral and ionized particles. Both theory and observation have suggested a significant cooling trend in the upper atmosphere. This upper atmospheric climate change will affect various human activities: satellite orbits can be altered as a result of the long-term decrease in atmospheric density throughout the entire upper atmosphere; changes in the ionospheric electron density modify the radio wave propagation, in particular, the performance of GPS and other space-based navigational or communication systems. Therefore in the decade

to come, our research should address this major issue of these long-term changes and their societal influence. We suggest the following main research themes:

- (1) To obtain a comprehensive global pattern of the long-term change in both the ionosphere and the thermosphere. It should include dependency on solar activity, season and time. Neutral temperature at different heights from above the mesosphere through the upper atmosphere is crucial information.
- (2) To identify physical causes of the derived trend pattern. Studies on greenhouse gas influences and other processes provide us a unique opportunity to further our understanding of coupling between the lower atmosphere and upper thermosphere and ionosphere. Modeling and data analysis remain the essential ways of deal with the problem.
- (3) To understand upper atmospheric climatology within a changing global environment. Ionospheric and thermospheric climatology, including empirical models, derived from early observations may need to be revisited or revised.

Long-term monitoring the upper atmosphere and solar-terrestrial environment from space and ground is a prerequisite and indispensable to achieving appropriate understanding of the earth's upper atmosphere. The observations should be made in a continuous and consistent manner for precisely detecting long-term changes in the system.

References

- Bilitza D. and B. Reinisch (2008), International Reference Ionosphere 2007: Improvements and new parameters, *Adv. Space Res.*, 42, #4, 599-609, doi:10.1016/j.asr.2007.07.048.
- Bremer, J. (1992), Ionospheric trends in mid-latitudes as a possible indicator of the atmospheric greenhouse effect, *J. Atmos. Terr. Phys.*, 54, 1505–1511.
- Donaldson, J. K., T. J. Wellman, and W. L. Oliver (2010), Long-term change in thermospheric temperature above Saint Santin, *J. Geophys. Res.*, 115, A11305, doi:10.1029/2010JA015346.
- Drob, D. P., et al. (2008), An empirical model of the Earth's horizontal wind fields: HWM07, *J. Geophys. Res.*, 113, A12304, doi:10.1029/2008JA013668
- Emmert, J. T., J. M. Picone, J. L. Lean, and S. H. Knowles (2004), Global change in the thermosphere: Compelling evidence of a secular decrease in density, *J. Geophys. Res.*, 109, A02301, doi:10.1029/2003JA010176.
- Emmert, J. T., J. M. Picone, and R. R. Meier (2008), Thermospheric global average density trends, 1967–2007, derived from orbits of 5000 near-Earth objects, *Geophys Res Lett*, 35, L05101, doi:10.1029/2007GL032809.
- Emmert, J. T., J. L. Lean, and J. M. Picone (2010), Record-low thermospheric density during the 2008 solar minimum, *Geophys. Res. Lett.*, 37, L12102, doi:10.1029/2010GL043671.
- Emmert, J. T., and J. M. Picone (2010), Climatology of globally averaged thermospheric mass density, *J. Geophys. Res.*, 115, A09326, doi:10.1029/2010JA015298.
- Holt, J. M., and S. R. Zhang (2008), Long-term temperature trends in the ionosphere above Millstone Hill, *Geophys. Res. Lett.*, 35, L05813, doi:10.1029/2007GL031148.

- Holt, J. M. and S.-R. Zhang (2010), Millstone Hill ISR observations of upper atmospheric long-term changes: thermospheric temperature and composition, manuscript to be submitted to *J. Geophys. Res.*.
- Kawamura, S., Y. Otsuka, S.-R. Zhang, S. Fukao, and W. L. Oliver (2000), A climatology of middle and upper atmosphere radar observations of thermospheric winds, *J. Geophys. Res.*, 105(A6), 12,777–12,788, doi:10.1029/2000JA900013.
- Keating, G. M., R. H. Tolson, and M. S. Bradford (2000), Evidence of long term global decline in the Earth's thermospheric densities apparently related to anthropogenic effects, *Geophys. Res. Lett.*, 27, 1523–1526.
- Laštovička, J., R.A. Akmaev, G. Beig, J. Bremer, J. Emmert (2006): Global change in the upper atmosphere. *Science*, 314 (5803), 1253-1254.
- Laštovička, J., R.A. Akmaev, G. Beig, J. Bremer, J.T. Emmert, C. Jacobi, M.J. Jarvis, G. Nedoluha, Yu.I. Portnyagin, T. Ulich (2008): Emerging pattern of global change in the upper atmosphere and ionosphere. *Ann. Geophysicae*, 26, 1255-1268.
- Liu, H., H. Lühr, and S. Watanabe (2007), Climatology of the equatorial thermospheric mass density anomaly, *J. Geophys. Res.*, 112, A05305, doi:10.1029/2006JA012199
- Mendillo, M., Huang, C.-L., Pi, X., Rishbeth, H., and Meier, R. (2005), The global ionospheric asymmetry in total electron content, *J. Atmos. Solar-Terr. Phys.*, 67, 1377–1387.
- Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin (2002), NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *J. Geophys. Res.*, 107(A12), 1468, doi:10.1029/2002JA009430.
- Qian, L., S. C. Solomon, and T. J. Kane (2009), Seasonal variation of thermospheric density and composition, *J. Geophys. Res.*, 114, A01312, doi:10.1029/2008JA013643.
- Qian, L., R. G. Roble, S. C. Solomon, and T. J. Kane, Calculated and observed climate change in the thermosphere, and a prediction for solar cycle 24, *Geophys. Res. Lett.*, 33, L23705, doi:10.1029/2006GL027185, 2006
- Qian, Liying, Stanley C. Solomon, Raymond G. Roble, and Timothy J. Kane (2008), Model simulations of global change in the ionosphere, *Geophys Res Lett*, 35, L07811, doi:10.1029/2007GL033156.
- Rishbeth, H. (1990), A greenhouse effect in the ionosphere?, *Planet. Space Sci.*, 38, 945–948.
- Rishbeth, H., and R. G. Roble (1992), Cooling of the upper atmosphere by enhanced greenhouse gases: Modeling of the thermospheric and ionospheric effects, *Planet. Space Sci.*, 40, 1011–1026
- Roble, R. G., and R. E. Dickinson (1989), How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere?, *Geophys. Res. Lett.*, 16, 1441–1444.
- Zhang, S.-R., J. M. Holt, A. P. van Eyken, C. Heinselman, and M. McCready (2010), IPY observations of ionospheric yearly variations from high- to middle-latitude incoherent scatter radars, *J. Geophys. Res.*, 115, A03303, doi:10.1029/2009JA014327.
- Zhang, S.-R., and J. M. Holt (2010) Millstone Hill ISR observations of upper atmospheric long-term changes: ionospheric height dependency, to be submitted to *J. Geophys. Res.*.
- Zhang S-R, J. Holt, A. P. van Eyken, M. McCready, C. Amory-Mazaudier, S. Fukao, M. Sulzer (2005), Ionospheric climatology and model from long-term databases of worldwide incoherent scatter radars, *Eos Trans. AGU*, 86(18), Jt. Assem. Suppl., Abstract SA52A-03.
- Zhang, S.-R. and J. M. Holt (2007), Ionospheric Climatology and Variability from Long-term and Multiple Incoherent Scatter Radar Observations: Climatology in Eastern American Sector, *J. Geophys. Res.*, 12, A06328, doi:10.1029/2006JA012206

Partition and variability of the magnetospheric energy input into the polar ionosphere

Yongliang Zhang, and Larry J. Paxton

The Johns Hopkins University Applied Physics Laboratory, Laurel, MD

1. Primary Science Objective

Energy from the magnetosphere into the polar ionosphere is the major driver of global disturbances in the ionosphere and thermosphere during geomagnetic storms. The energy input originates from the solar wind. However, accurate quantitative specification of the energy input and its partition (Poynting flux versus particle flux and ionosphere dissipation) is still not available. Four specific science questions to answer are:

- (1) Quantitative relationship between the Poynting flux and energy flux of precipitating particles over different aurora forms and various geomagnetic conditions.
- (2) The relationship among the energy input (Poynting flux and particle flux) and ionosphere dissipation (Joule heating) rate under various geomagnetic conditions.
- (3) Ionosphere (enhanced conductance) feedback effect on the energy input and dissipation.
- (4) Variability of the energy input and their relation to the dayside magnetopause merging rate.

2. Why is it important?

The proposed project will provide quantitative specification of the magnetospheric energy input, its partition and dissipation in the polar ionosphere and their relationship to the solar wind conditions (magnetopause merging rate [*Newell, et al.*, 2007]). Such quantitative information will significantly advance our understanding of magnetosphere-ionosphere (M-I) coupling as well as provide tools to further improve space weather forecast accuracy.

3. What implementation is proposed to address that objective?

To address the science questions, electric field (E), magnetic field (B) and particle (electron and ion) precipitation have to be measured in space. Auroral images, thermosphere neutral wind and ionosphere conductance are also needed.

The pointing flux ($S = E \times B / \mu_0$) can be directly calculated from the E and B measurements. The energy flux of precipitating particles can also be directly estimated from the measured particle spectra. The Joule heating (QJ) can be estimated by $\sigma_P (E + \mathbf{V}_N \times \mathbf{B}_0)^2$, where σ_P , $\mathbf{V}_N$, and $\mathbf{B}_0$ are

the height integrated Pedersen conductance, neutral wind velocity and earth's magnetic field at the ionosphere altitude, respectively. The Pedersen conductance depends on solar EUV flux and precipitating particles. It can be estimated based on solar EUV flux or F10.7 flux and the measurement of precipitating particles or the assimilative mapping of ionospheric electrodynamics (AMIE) from ground based measurements (e.g. <http://amie.engin.umich.edu/>). In addition to the satellite based E field measurement, the ground based SuperDarn [Ruohoniemi and Baker, 1998] can provide E field map in the polar ionosphere. Note the strength of electric field is often anti-correlated with auroral brightness (energy flux of precipitating particles) [Baker *et al.*, 2004]. Figure 1 shows an example map of Joule heating (QJ) and particle heating (QP) based on TIMED/GUVI and SuperDarn measurements. One interesting feature in the plot is the low E field associated with the bright auroral spot at the nightside poleward edge. This indicates the feedback effect of ionospheric conductance on the E field, and thus the Poynting flux.

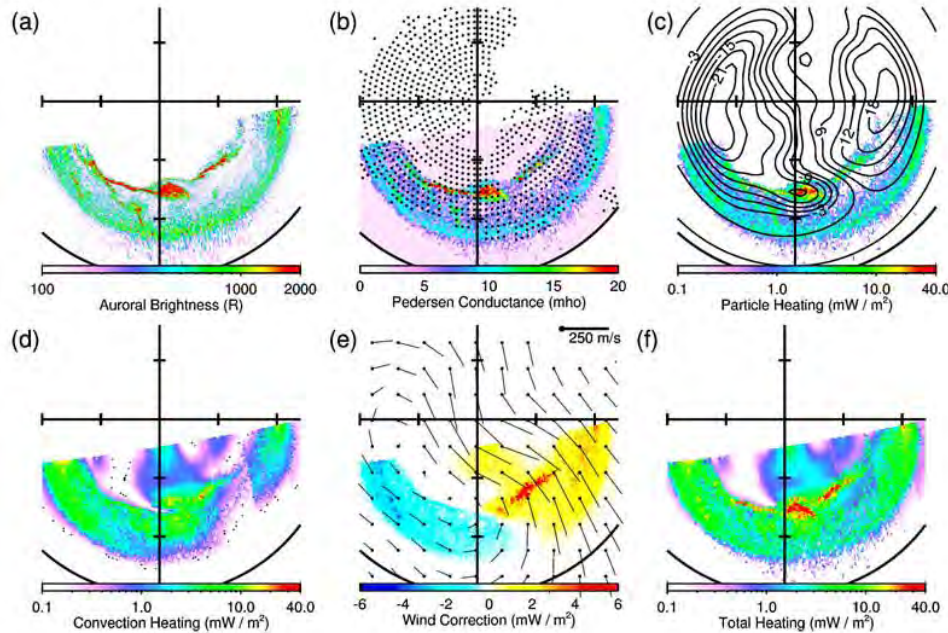


Figure 1. GUVI-SuperDARN measurements obtained between 2143–2201 UT on December 4, 2002 in MLAT-MLT format (noon is upward). Minimum latitude shown is 60 MLAT. Parameters are: (a) GUVI auroral brightness; (b) GUVI SP plus a 2-mho background and overlain SuperDARN measurement locations; (c) QP with overlain contours of SuperDARN “composite” electrostatic potential in kV; (d) Convection heating; (e) Neutral wind correction with overlain HWM93 wind vectors at 160 km; and (f) total heating, $QT = QJ + QP$. (after Baker *et al.*, 2004).

The neutral wind in Figure 1 is from a model [Hedin *et al.*, 1996]. But the neutral wind can also be obtained by ground based All-Sky Doppler Imager (ASDI) [Kosch *et al.*, 2010] and satellite based neutral wind imager (e.g. TIMED/TIDI, Killeen *et al.*, 1999). Figure 2 shows examples of

the neutral wind map in auroral region from an All-sky Doppler imager together with the ion drift from SuperDarn. A chain of ASDI could be used to cover the whole polar region.

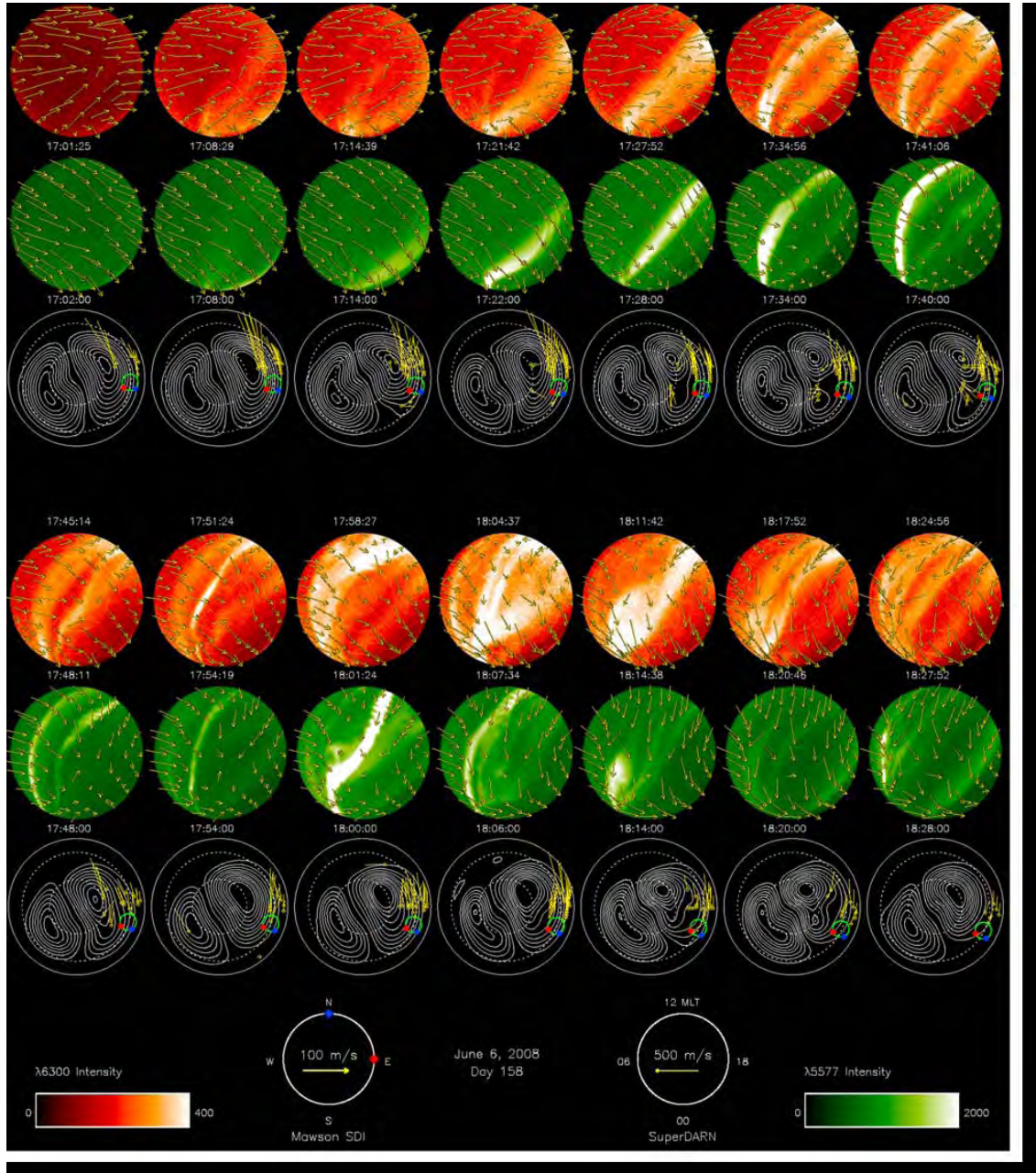


Figure 2. (a) The 144° field of view of the All-Sky Doppler Images. For each group of three rows, the top row show F layer 630 nm Doppler images, the middle row shows E layer 557.7 nm Doppler images, and the bottom row shows SuperDARN ion convection equipotential contours with 3 kV increments (white) and vectors on 6 June 2008 (D158). Each plot is labeled with UT. The optical intensity is not calibrated. The optical data are plotted in geographic coordinates with wind vectors overlaid for each zone (yellow). The wind vectors are centered on their measurement points. The vector for the zenith sector shows the average wind field, not the vertical wind. The ion convection plots are plotted in geomagnetic coordinates with the SuperDARN vectors (yellow) overlaid at reduced

spatial resolution. The large (small) dotted circles represent 70° (80°) geomagnetic south. The ASDI's 557.7 nm field of view (green circle) as well as north (blue dot) and east (red dot) geographic directions are overlaid. The ion vectors start on their measurement points. Note that the ion and wind velocity vectors have different scales and shaped vectors for clarity. (b) The format is identical to that of Figure 2a, except it is for 10 June 2008 (D162). (after *Kosch et al.*, 2010)

The energy flux of precipitating particles are mostly carried by electrons and ions with energy from tens of eV to tens of keV. Particle detectors for such precipitating electrons and ions have been developed and used for satellite based measurements for many decades (e.g on DMSP satellites [*Hardy et al.*, 1984; 1985], NOAA POES satellites [*Evans and Greer*, 2004]).

A simple CMOS based auroral imager with 557.7 (green) and 630.0 (red) nm filters provides auroral morphology (e.g discrete versus diffuse aurora). This determines the type of auroras associated with the energy input and the partitioning. Furthermore, the net auroral intensities of the green and red also provide the energy flux and mean energy of the precipitating electrons.

Solar wind and IMF data from ACE and WIND satellites are available from the NASA's data center for estimation of the dayside magnetopause magnetic merging rate [*Newell et al.*, 2007]. The merging rate is found to be well correlated with the auroral hemispheric power [*Newell et al.*, 2007, *Zhang et al.*, 2010]. Therefore, the merging rate can be used as an indicator of total solar wind energy input to the magnetosphere.

Option 1.

To reduce the cost and risk, the first option is to add the electric field, particle detector and auroral imager on commercial satellites: The Iridium constellation at ~780 km altitude (http://en.wikipedia.org/wiki/Iridium_satellite_constellation) is a good candidate. The existing magnetic field data from Iridium satellites have already been processed and used to extract field aligned current intensity through the AMPERE (Active Magnetosphere and Planetary Electrodynamics Response Experiment) project which is funded by NSF. The Iridium constellation includes tens of polar orbiting satellites at different longitudes which are ideal for the proposed polar energy input and partition studies and their relationship to the dayside magnetic merging rate (solar wind).

Option 2.

The second option is to add the electric field, magnetic field, and particle detector and auroral imager on Russian satellites in Molniya orbit. A Molniya orbit is a type of highly elliptical orbit with an inclination of 63.4 degrees and an orbital period of precisely one half of a sidereal day. A satellite placed in a Molniya orbit spends most of its time over a designated area of the earth as a result of "apogee dwell". This provides a great opportunity to study the energy input partitioning above the auroral electron acceleration altitudes (~ between 10000 and 20000km), yet with high time resolution.

The option 2 can go with option 1 concurrently or after the option 1 has been successfully carried out. If both of the options are selected, one additional important science question can be addressed: How much of the Poynting flux above the auroral electron acceleration altitude is used to speed up auroral electrons?

Furthermore, the electric and magnetic field detectors can also be placed on tens or hundreds cube satellites to ensure relatively even and near continuous coverage in the polar region.

4. How feasible is the implementation?

All of the instruments in this proposal have been developed and used for many decades. They are quite mature with very low risk. The ride on commercial satellites further minimizes the cost of satellite launch and operation. Other supporting data from ground based observations are available at no cost or small cost to expand (e.g. All-Sky Doppler Imager chain for thermosphere neutral wind measurements).

References

Baker, J.B., Y. Zhang, R.A. Greenwald, L.J. Paxton, and D. Morrison, Height-integrated Joule and auroral particle heating in the night side high latitude thermosphere, *Geophys. Res. Lett.*, 31, L09807, doi:10.1029/2004GL019535, 2004.

Evans D.S., and M.S Greer, Polar orbiting environmental satellite space experiment monitor-2: instrument descriptions and archive data documentation, NOAA Technical Memorandum version 1.3. NOAA Space Environment Center, Boulder, Colorado, 2004.

Hardy, D.A., L.K. Schmitt, M.S. Gussenhoven, F.J. Marshall, H.C. Yeh, T.L. Schumaker, A. Huber, and J. Pantazis, Precipitating electrons and ion detectors (SSJ/4) for the block 5D/flight 6-10 DMSP satellites: Calibration and data presentation, Rep. AFGL-TR-84-0314, Air Force Geophys. Lab., Hanscom Air Force Base, Mass., 1984.

Hardy, D.A., M.S. Gussenhoven, and E. Holeman, A statistical model of auroral electron precipitation, *J. Geophys. Res.*, 90, 4229, 1985.

Hedin, A.E., E.L. Fleming, A.H. Manson, F.J. Schmidlin, S.K. Avery, R.R. Clark, S.J. Franke, G.J. Fraser, T. Tsuda, F. Vial, and R.A. Vincent, Empirical wind model for the upper, middle and lower atmosphere, *J. Atmos. Terr. Phys.* 58, 1421-1447, 1996.

Killeen, T.L., et al., TIMED Doppler interferometer (TIDI), Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research III, Allen M. Larar, Editors, pp.289-301, DOI: 10.1117/12.36638, 20 October 1999.

Kosch, M. J., C. Anderson, R. A. Makarevich, B. A. Carter, R. A. D. Fiori, M. Conde, P. L. Dyson, and T. Davies, First E region observations of mesoscale neutral wind interaction with auroral arcs, J. Geophys. Res., 115, A02303, doi:10.1029/2009JA014697, 2010.

Newell P. T., T. Sotirelis, K. Liou, C.-I. Meng, F. J. Rich, A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables, J. Geophys. Res., 112, A01206, doi:10.1029/2006JA012015, 2007.

Zhang, Y., L. J. Paxton, and D. Morrison, Auroral and thermospheric response to the 9 day periodic variations in the dayside reconnection rate in 2005, Space Weather, 8, S07001, doi:10.1029/2009SW000559, 2010.

Dayside Aurora and Auroral Conjugacy

Xiaoyan Zhou¹, Athanasios Boudouridis, John P. Dombek, Eric Donovan, David J. Knudsen, Kan Liou, Dag A. Lorentzen, Anthony Lui, Dirk Lummerzheim, Ching I. Meng, Joran I. Moen, Patrick T. Newell, Nikolai Ostgaard, George K. Parks, Robert Rankin, David G. Sibeck, James F. Spann, Robert J. Strangeway, Clare Watt, Yongliang Zhang

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

Overview

This white paper discusses the scientific significance of dayside and conjugate auroras. The dayside aurora and its variation have less ambiguity in the connection to their cause. Examples of such advantages are: 1) the timing of the shock-aurora trigger (i.e., the touchdown time of an interplanetary shock on the subsolar magnetopause) can be as accurate as from few seconds to few minutes; 2) the shock-aurora speed in the ionosphere can be compared with the shock speed in the solar wind. These features provide a unique opportunity to test the existing theories of auroral dynamics, and to establish one-to-one correlations between the auroral forms and particle precipitation mechanisms. Measurements of dayside and conjugate auroras can be achieved by conducting balloon campaigns in the Antarctica and by coordinating the balloon observations with simultaneous Arctic ground-based auroral imaging.

1. Science Background

The aurora provides a map of magnetospheric structure and dynamics and thus auroral images provide rich information about the global magnetosphere that cannot be obtained from in-situ observations alone. The physics of the dayside aurora is different in many ways from that of the nightside aurora because of dayside-nightside asymmetries of the magnetosphere and the ionosphere due to the presence of the solar wind and solar photon ionization. The insight provided by studying the dayside aurora is critical to determining the coupling between variability in the solar wind and the resulting geospace response.

One salient dayside auroral phenomenon occurs when interplanetary shocks or sudden increases in the solar wind ram pressure impinge on the subsolar magnetopause. As a consequence, the magnetospheric compression and magnetic reconnection are enhanced in the local noon sector. Auroras in the dayside auroral ionosphere from the oval to $\sim 65^\circ$ MLat light up within seconds to few minutes indicating causes of the aurora in the outer magnetosphere and/or on the magnetopause $\sim 10 R_E$ away from the ionosphere (Craven et al., 1986; Sandhalt et al., 1994; Spann et al., 1998; Zhou and Tsurutani 1999; Vorobjev et al., 2001; Liou et al, 2002; Zhang et al., 2002; Hubert et al., 2003; Meurant et al., 2003,2004; Zhou et al., 2003; Fuselier et al., 2004; Zhou et al., 2009). The auroras then propagate antisunward along the oval at very high ionospheric speeds of ~ 6 -11 km/s that match the corresponding shock speed in the solar wind. The aurora is found to be caused by both the electron and proton precipitation. The auroral brightness is related not only to the variation in the solar wind, but also the plasma precondition in the dayside magnetosphere.

There have been speculations as to the causes of the aurora, such as magnetic reconnection, wave-particle interaction and magnetic shearing, etc. However, since there is a lack of observations of the small-scale dayside auroral structure, we have little understanding of a one-to-one correlation between auroral forms and the particle precipitation mechanism. For example, theoretically, a fast release of magnetic shear stresses could establish field-aligned potential drops and convert the differential magnetic energy into kinetic energy of auroral particles (Haerendel 2007; 2008). But the theory is not tested by auroral observations and it is unclear what would be the resultant auroral color (i.e., the electron characteristic energy), auroral form and their variations. Another example is that it has been a long-time belief that dayside red arcs are the manifestation of the magnetic reconnection on the magnetopause (e.g., Moen et al., 1998; Sandholt et al., 1998a,b; Chaston et al., 2005). The red oxygen OI^1D emission layer is mainly at ~ 250 km altitude and caused by soft electrons with the characteristic energy of few hundred eV. However, a 1-D MHD model showed that Alfvén-wave accelerated electrons possess energies up to few keV (Chaston et al., 2002). On the dayside magnetopause the kinetic Alfvén wave is very likely generated by the magnetic reconnection (Chaston et al., 2005).

It has been a long-term challenge to image the aurora from the ground under sunlit conditions (e.g., Ree et al., 2000). This is mainly due to substantial contamination by sunlight and the limited land area available for installing imagers near the auroral zone in both hemispheres. While ultra-violet remote sensing from space can measure dayside aurora, it does not resolve small-scale auroral structures and is contaminated by UV airglow. Consequently, we have little knowledge of dayside auroral small-scale structures (i.e., auroral forms), their variations, and their coupling to conjugate aurora. As a result, there are important unanswered questions regarding the coupling between the variable solar wind and the dayside magnetosphere and ionosphere and also questions about the basic phenomenology of small-scale dayside auroral structures.

2. Scientific Questions

These questions include:

- 1). What are the auroral particle precipitation mechanisms? The advantage of studying dayside auroras is that there is less ambiguity in identifying the solar wind and magnetospheric causes of the auroral particle precipitation. The touchdown time of an interplanetary shock on the subsolar magnetopause can be accurately identified up to few seconds – few minutes; the auroral speed in the ionosphere matches that in the solar wind. These characteristics are unique opportunity to test various kinds of auroral mechanisms.
- 2). How do the dayside magnetosphere and ionosphere respond to solar wind variations in terms of auroral forms? (e.g. diffuse auroral patches and their expansion, auroral arcs/beams and their motion, the auroral response to abrupt solar wind pressure changes, jumps in IMF direction or stable but intense IMF such as in magnetic clouds.)

- 3). Are auroras (e.g., those seen during substorms, geomagnetic storms, and magnetospheric compression) hemispherically symmetric? The question covers the symmetry from small-scale structures (e.g., Sato et al., 1998) to global dynamics (e.g., Ostgaard et al., 2004; Laundal and Ostgaard, 2009). The answer can test many of current models and our understanding of geomagnetic disturbances. For example, do substorm onsets start in the auroral ionosphere (Kan and Sun, 1996) or in the magnetotail (Lui, 1991a,b; Baker et al., 2002)? The onset auroras in the two hemispheres are not necessarily conjugate for onsets in the ionosphere, but should be conjugate for those in the tail.
- 4). What signatures in conjugate dayside aurora can be used to diagnose reconnection and the response of the dayside magnetopause to IMF variations? What does this tell us about particle acceleration, magnetopause reconnection rates, and coupling of energy, mass and momentum from the solar wind?
- 5). What are the differences in the auroral emissions in sunlight and darkness? Are there unique signatures in the sunlight and what causes them? The answers will address the ionospheric conductivity effect and how it couples with the dayside and nightside magnetospheric dynamics.
- 6). What are the differences in auroral forms in dayside (~06-12-18 MLT) and nightside ionosphere (~08-00-16 MLT)? This question addresses the difference of the solar wind-magnetosphere-ionosphere interaction in the dayside and nightside geospace areas.

Identifying dayside auroral forms actually has broader applications beyond the above questions. It has been confirmed by the FAST and Cluster observations that ion conics occur along with the broadband VLF waves where soft electrons (with energy less than 1 keV) are seen to be highly field-aligned with the magnetic field (e.g., Chaston et al., 2005). Red auroral arcs are expected at the ionospheric footprints of the field lines. This connection between the red arcs and the ion conics implies that the red arc can be an indication of the oxygen ion outflow that is considered a very important, if not the most, ionospheric source of the ring current energetic ions. Could this indication only valid in the dayside/cusp or change case by case? Furthermore, since the dayside red arcs are caused by the magnetic reconnection (e.g., Chaston et al., 2005), there is a strong likelihood of the reconnection being the very responsibility of the oxygen ion outflow. This speculation is consistent with Strangeway et al.'s flow chart for the generation of ionospheric outflows (Strangeway et al., 2005) and is consistent with the fact that geomagnetic storms develop only when there is a long lasting southward IMF.

3. Technical Approach and the Feasibility

To answer such questions, the basic requirement is to image the aurora in relatively close distance (so auroral forms can be determined) in sunlight. If such measurements can be achieved, conjugate auroral images can be obtained by combining auroral images taken in sunlight (local summer) from Antarctic balloons with simultaneous Arctic ground-based auroral images taken in

darkness (local winter). Thus, the science questions lead to a technical challenge, i.e., how can we image aurora under sunlight? Seeking a feasible method has been a continuing effort (Chakrabarti, 1998; Rees et al., 2000; Pallamraju et al., 2004), which deserves a solution because of the science significance discussed above.

Atmospheric models predict that sky brightness decreases with increasing altitude and wavelength (Anderson et al., 1999; Liou, 2002). Based on these models and test observations, Zhou et al. (2007, 2008) have confirmed that the N_2^+ Meinel line emissions (~ 1100 nm), which gives the best signal noise ratio (SNR) in near-infrared, can be measured in sunlight at typical balloon altitudes (>35 km) where signal to background ratios are adequate. Using a JPL built near-infrared (NIR) InGaAs camera (with a small focal plane array of 320×256 and a small FOV of 9°), the Meinel auroral emissions were detected during twilight when the sky brightness was similar to that of 35-40 km altitude during the highest solar elevation in the Antarctica.

Some of the images are shown in Figure 1. The aurora was detected on April 13, 2005 during evening twilight when the solar zenith angle, SZA, was $\sim 96^\circ$ (i.e., the sun was $\sim 6^\circ$ below the horizon). Corresponding AL was about -400 nT implying a medium level of auroral intensity. This test result confirmed the feasibility of measuring sunlit auroras from balloon. When the measurements in the Antarctica are coordinated with the existing ground-based all-sky imagers, conjugate auroras can be obtained.

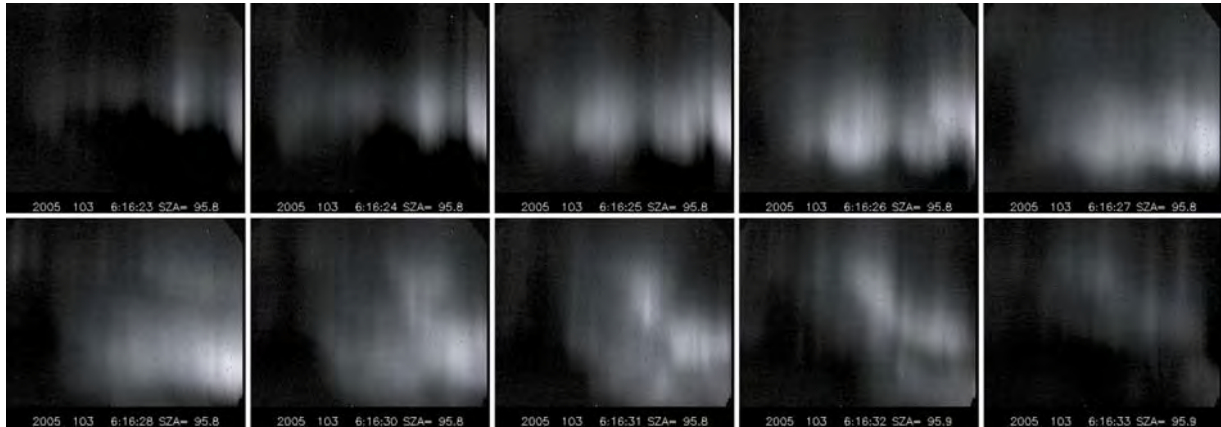


Fig. 1. Auroras detected by the JPL InGaAs NIR camera from Poker Flat, Alaska. Images are shown in one second cadence chronologically from left to right, then down to the second row. Without applying any filter, the images were taken in evening twilight from 2116:23 to 2116:33 LT on April 13, 2005 when SZA was $\sim 96^\circ$. Each image has had a sky background subtracted that was calculated from 40 images around the time of the image (Zhou et al., 2007, 2008). The sky background is mainly contributed by the OH airglow in 1400-1700 nm (Remick et al., 2001). Detailed discussions of the sky intensity can be found in Zhou et al. (2007, 2008).

Comparing to space mission, balloon flight for auroral observations is very cost effective with a total cost only $\sim \$2$ M including the balloon flight and the NIR camera system development. The JPL NIR InGaAs camera technology has a TRL 5-6 and possesses much better quality than commercial NIR cameras that only function well under high frame rate of ~ 100 /s or higher. With current NASA balloon technology, flying at 35-40 km altitudes with a 1000 kg payload is routine.

The NASA balloon altitude record is ~50 km with ~200 kg payload. Driven by the Antarctic circumpolar winds, a balloon can maintain a trajectory around the geographic pole over several weeks. (Such an example of the balloon trajectory relative to the oval and the NIR camera's field-of-view is shown in Zhou et al., 2007.) The NIR camera system requires less than 10 kg weight and consumes ~30 W, which are suitable for the balloon flight and even a balloon piggyback flight.

4. Concluding Remarks

The 24th solar cycle started its ascending phase recently. Increasing solar and solar wind activity will more frequently light up intense auroral bursts. The Antarctic auroral campaign and its leverage with the ground-based all-sky imager arrays and with the Cluster and THEMIS satellite constellation will provide a unique opportunity for dayside and conjugate auroral investigations. (Note there certainly are opportunities to measure the nightside aurora in both hemispheres during a conjugate auroral campaign.) A broader coordination and collaboration with observations of riometer, X-ray, SuperDARN and other ionospheric measurements as well as theory/modeling will be strongly encouraged.

We highly recommend that the communities interested in the solar wind-magnetosphere-ionosphere interaction, high-latitude ionosphere, auroral dynamics, auroral conjugacy, auroral morphology and its magnetospheric causes, and auroral particle acceleration theory be given the opportunity to work jointly on measuring the dayside and conjugate aurora and on the auroral investigation.

The primary benefits of this task include: 1) an insight into the dayside geospace response to the varying solar wind; 2) one-to-one correlations between the auroral form and particle precipitation mechanism. Such correlations are helpful references for the nightside auroral activity; 3) the auroral conjugacy during the dayside reconnection, magnetosphere compression and magnetic shearing as well as during nightside auroral substorms; 4) a test of existing storm and substorm models and speculations; 5) a complementary method (that employs NIR camera for auroral emissions at ~1100 nm) to traditional visible and UV auroral remote sensing. Note that the characteristic energy of precipitated electrons alters with the auroral wavelength; and 6) a more complete knowledge of auroral micro-scale signatures along the global auroral zone.

References

- Anderson, G.P., et al., *Proc.SPIE, Optics in Atmospheric Propagation and Adaptive Systems III*, 3666, 2-10, 1999.
- Baker, D.N., et al., *Geophys. Res. Lett.*, 29, 2190, 10.1029/2002GL015539, 2002.
- Chakrabarti, S., *J. Atmos. And Terr. Phys.*, 60, 1403, 1998.
- Chaston, C. C., et al., *Phys. Rev. Lett.*, 95, 065002, doi:10.1103/PhysRevLett.95.065002, 2005.
- Chaston, C. C., et al., *Geophys.Res.Lett.*, 29, 11, doi:10.1029/2001GL013842, 2002.
- Craven, J. D., et al., in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 367–380, Terra Sci., Tokyo, 1986.

Fuselier, S. A., et al., *J. Geophys. Res.*, 109, A12227, doi:10.1029/2004JA010393, 2004.

Haerendel G., *J. Geophys. Res.*, 112, A09214, doi:10.1029/2007JA012378, 2007

Haerendel G., *J. Geophys. Res.*, 113, A07205, doi:10.1029/2007JA012947, 2008

Hubert, B., et al., *Geophys. Res. Lett.*, 30(3), 1145, doi:10.1029/2002GL016464, 2003.

Kan, J. R., and W. Sun, *J. Geophys. Res.*, 101, 27,271, 1996.

Laundal, K. M., and N. Ostgaard, *Nature*, 460, 491-493, doi:10.1038/nature08154, 2009.

Liou, K., et al., *Geophys. Res. Lett.*, 29(16), 1771, doi:10.1029/2001GL014182, 2002.

Liou, K.N., Academic Press, San Diego, 2002.

Lui, A.T.Y., *J. Geophys. Res.*, 96, 1849, 1991a.

Lui, A.T.Y., In *Magnetospheric Substorms*, Geophysical Monograph 64, Ed. J. Kan, et al., pp. 43, Washington, DC: AGU, 1991b.

Meurant, M., et al., *Geophys. Res. Lett.*, 30(20), 2032, doi:10.1029/2003GL018017, 2003.

Meurant, M., et al., *J. Geophys. Res.*, 109, A10210, doi:10.1029/2004JA010453, 2004.

Moen, J., et al., *J. Geophys. Res.*, 103(A7), 14,855 – 14,863, doi:10.1029/97JA02877, 1998.

Ostgaard, N., et al., *J. Geophys. Res.*, 109, A07204, doi:10.1029/2003JA010370, 2004.

Pallamraju D., et al., *Geophys. Res. Lett.*, 31, L08807, doi:10.1029/2003GL019173, 2004.

Rees, D., et al., *Geophys. Res. Lett.*, 27, 313–316, doi:10.1029/1999GL003696, 2000.

Remick, K. J., et al., *J. Atmos. Sol.-Terr.Phys.*, 63, 295, 2001.

Sandholt, P. E., et al., *J. Geophys. Res.*, 99, 17,323, 1994.

Sandholt, P. E., et al., *J. Geophys. Res.*, 103(A10), 23,325, doi:10.1029/98JA02156, 1998a.

Sandholt, P., et al., *J. Geophys. Res.*, 103(A9), 20,279– 20,295, doi:10.1029/98JA01541, 1998b.

Sato, N., et al., *J. Geophys. Res.*, 103, NO. A6, PAGES 11,641-11,652, JUNE 1, 1998.

Spann, J. F., et al., *Geophys. Res. Lett.*, 25, 2577, 1998.

Strangeway, R.J., et al., *J. Geophys. Res.*, 110, A03221, doi:10.1029/2004JA010829, 2005.

Vorobjev, V. G., et al., *J. Geophys. Res.*, 106, 28,897, 2001.

Zhang, Y., et al., *J. Geophys. Res.*, 107, 8001, doi:10.1029/2002JA009355, 2002.

Zhou, X.-Y., and B. T. Tsurutani, *Geophys. Res. Lett.*, 26, 1097, 1999.

Zhou, X.-Y., et al., *J. Geophys. Res.*, 108(A4), 8019, doi:10.1029/2002JA009701, 2003.

Zhou, X.-Y., et al., *Geophys. Res. Lett.*, 34, L03105, doi:10.1029/2006GL028611, 2007.

Zhou, X.-Y., et al., *J. Adv. Space Res.*, 42, 1676-1682, doi:10.1016/j.asr.2007.02.034, 2008.

Zhou, X-Y., et al., *J. Geophys. Res.*, 114, A12216, doi:10.1029/2009JA014186, 2009.

Meso Scale Transients in Magnetotail and Their Roles in Substorm Dynamics

Ping Zhu

University of Wisconsin-Madison

November 12, 2010

Recent substorm studies, including the THEMIS mission and the ongoing *GEM Focus Group 12: Substorm Expansion Onset: The First 10 Minutes*, suggest that the meso scale transients in magnetotail may play key roles in mediating and regulating the transition process from the late substorm growth phase to the beginning of onset expansion. Examples of the meso scale transients include X-lines and plasmoids from magnetic reconnection, bubbles and Rayleigh-Taylor fingers from interchange-ballooning instability, as well as dipolarization fronts and bursty bulk flows. The meso scales here refer to the spatial scales ranging from the ion skin depth or ion gyroradius ($\sim 0.1 - 1 R_E$), to the length scale of the local inhomogeneities such as the gradient scale length of the pressure profile ($\sim 1 - 10 R_E$). The meso scales are also considered as inertial scales, since they overlap with the inertial regime of energy transfer spectrum. We propose to systematically investigate the meso scale transient structures and couplings in magnetotail, in order to explore emerging scenarios of substorm onset.

One of the major challenges of substorm modeling is the adequate resolution of the multiple spatial-temporal scales intrinsically involved in the expansion onset process. In a typical substorm cycle, the cross-tail current density at the near-Earth magnetotail (“near-Earth” or “near-tail”) region of plasma sheet ($\sim 7 - 11 R_E$) first evolves through a period of slow growth ($\sim 0.5 - 1.5$ hr) and then enters a near-explosive intensification over a short interval (~ 1 min) [Ohtani *et al.*, 1992]. The near-explosive growth phase is then followed by a sudden reduction (usually referred to as “disruption”) of the current sheet in the near-tail region on a time scale of seconds [Lui, 1978; Lopez and Lui, 1990; Jacquoy *et al.*, 1991]. As a result of the cross-tail current intensification, the magnetic configuration of the night-side magnetosphere evolves from dipole-like to highly-stretched and tail-like at near-Earth distance ($\sim 7 - 9 R_E$) [Kaufmann, 1987]. Observations indicate that the thickness of the extended plasma sheet can be reduced to less than $1 R_E$ in the near-tail region before the onset of the expansion phase [Sergeev *et al.*, 1990; Lui *et al.*, 1992; Pulkkinen *et al.*, 1992]. Whereas the longest phase of the current sheet evolution takes place in a slow, macroscopic transport scale, the most dramatic and critical phase of the evolution, the substorm onset, involves processes mostly in the meso, inertial scale.

Two major meso-scale MHD processes in the magnetotail are commonly believed to directly serve, either independently or jointly, as trigger mechanisms for the onset of a substorm: the ballooning-interchange instabilities of the plasma sheet in the near-Earth magnetotail (e.g. [Lui, 1991]), and magnetic reconnection in the plasma sheet of the middle magnetotail (“mid-tail”) re-

gion ($\sim 20 - 50 R_E$) (e.g. [Baker *et al.*, 1996]). The ballooning instability is driven by the pressure gradient in the unfavorable curvature region of magnetic field lines, and its growth is strongest in the near-tail plasma sheet region within the equatorial plane. This has often been observed to be the site of substorm onsets (e.g. [Erickson *et al.*, 2000]). The mid-tail reconnection activities are an inevitable consequence of current sheet intensification due to prolonged solar wind pressing at the dayside and the nightside magnetopause. They are capable of generating fast Earthward bursty bulk flows (BBFs) that are often used to associate with a substorm onset [Sergeev *et al.*, 1995; Shiokawa *et al.*, 1998; Nagai *et al.*, 1998]. However, the exact roles of those processes prior to expansion onsets remain unclear, and quantifying the dynamics of the reconnection process and the ballooning instability, as well as their association with the spatial-temporal scales of the plasma sheet evolution leading to a substorm onset has been the focus of recent substorm modeling efforts.

Recent (THEMIS) observations and analyses have renewed the long time interests in meso scale transient structures in magnetotail (e.g. [Runov *et al.*, 2009; Sitnov and Schindler, 2010]). Our study of the first well documented THEMIS substorm event [Angelopoulos *et al.*, 2008] indicates that ballooning instabilities of the near-tail plasma sheet can be initiated and enhanced by the current sheet thinning and stretching induced by the tearing-like process in the near-tail [Zhu *et al.*, 2009; Raeder *et al.*, 2010]. Such a tearing-like instability on closed field lines in the near-tail, previously thought as being stabilized by finite B_z , could be enhanced by Earthward BBFs and dipolarization fronts from mid-tail [Runov *et al.*, 2009], as shown in recent theory analysis [Sitnov and Schindler, 2010] and meso-scale MHD simulations [Zhu, 2010]. More remarkably, recent aurora observations from the THEMIS All Sky Imagers during the 2008 tail season discovered a predominant equatorward intrusion of the poleward boundary intensification (PBI) prior to the onset of a majority of substorm events [Nishimura *et al.*, 2009], a strong evidence of near-tail and mid-tail interaction [Zhu *et al.*, 2010]. We propose to further investigate these emerging scenarios, in which the couplings of meso scale dynamics between ballooning-like and tear-like processes, and between mid-tail region and near-tail region are essential, using a combined approach of global and meso-scale modelings.

References

- Angelopoulos, V., D. Sibeck, C. W. Carlson, J. P. McFadden, D. Larson, R. P. Lin, J. W. Bonnell, F. S. Mozer, R. Ergun, C. Cully, K. H. Glassmeier, U. Auster, A. Roux, O. LeContel, S. Frey, T. Phan, S. Mende, H. Frey, E. Donovan, C. T. Russell, R. Strangeway, J. Liu, I. Mann, J. Rae, J. Raeder, X. Li, W. Liu, H. J. Singer, V. A. Sergeev, S. Apatenkov, G. Parks, M. Fillingim, and J. Sigwarth, First results from the THEMIS mission, *Space Sci. Rev.*, *141*, 453–476, doi: 10.1007/s11214-008-9378-4, 2008.
- Baker, D. N., T. I. Pulkkinen, V. Angelopoulos, W. Baumjohann, and R. L. McPherron, Neutral line model of substorms: Past results and present view, *J. Geophys. Res.*, *101*, 12,975–13,010, 1996.
- Erickson, G. M., N. C. Maynard, W. J. Burke, G. R. Wilson, and M. A. Heinemann, Electromagnetics of substorm onsets in the near-geosynchronous plasma sheet, *J. Geophys. Res.*, *105*, 25,265–25,290, doi:10.1029/1999JA000424, 2000.
- Jacquey, C., J. A. Sauvaud, and J. Dandouras, Location and propagation of the magnetotail cur-

- rent disruption during substorm expansion: Analysis and simulation of an ISEE multi-onset event, *Geophys. Res. Lett.*, *18*, 389, 1991.
- Kaufmann, R. L., Substorm currents: growth phase and onset, *J. Geophys. Res.*, *92*, 7471, 1987.
- Lopez, R. E., and A. T. Y. Lui, A multisatellite case study of the expansion of a substorm current wedge in the near Earth magnetotail, *J. Geophys. Res.*, *95*, 8009, 1990.
- Lui, A. T. Y., Estimates of current changes in the geomagnetotail associated with a substorm, *Geophys. Res. Lett.*, *5*, 835, 1978.
- Lui, A. T. Y., Extended consideration of a synthesis model for magnetic substorms, in *Magnetospheric Substorms*, vol. 64, edited by J. R. Kan, T. A. Potemra, S. Kokubun, and T. Ijima, p. 43, AGU Monogr. Ser., American Geophysical Union, 1991.
- Lui, A. T. Y., R. E. Lopez, B. J. Anderson, K. Takahashi, L. J. Zanetti, R. W. McEntire, T. Potemra, D. M. Klumpar, E. M. Greene, and R. Strangeway, Current disruptions in the near-Earth neutral sheet region, *J. Geophys. Res.*, *97*, 1461, 1992.
- Nagai, T., M. Fujimoto, Y. Saito, S. Machida, T. Terasawa, R. Nakamura, T. Yamamoto, T. Mukai, A. Nishida, and S. Kokubun, Structure and dynamics of magnetic reconnection for substorm onsets with Geotail observations, *J. Geophys. Res.*, *103*, 4419–4440, doi:10.1029/97JA02190, 1998.
- Nishimura, Y., L. R. Lyons, S. Zou, V. Angelopoulos, and S. B. Mende, Substorm triggering by new plasma intrusion 1: THEMIS all-sky imager observations, in *Eos Trans. AGU*, *90*(52), *Fall Meet. Suppl.*, abstract SM41B-1708, 2009.
- Ohtani, S., K. Takahashi, L. J. Zanetti, T. A. Potemra, R. W. McEntire, and T. Iijima, Initial signatures of magnetic field and energetic particle fluxes at tail reconfiguration: Explosive growth phase, *J. Geophys. Res.*, *97*, 19,311–19,324, 1992.
- Pulkkinen, T. I., D. N. Baker, R. J. Pellinen, J. Bchner, H. E. J. Koskinen, R. E. Lopez, R. L. Dyson, and L. A. Frank, Particle scattering and current sheet stability in the geomagnetic tail during the substorm growth phase, *J. Geophys. Res.*, *97*, 19,283, 1992.
- Raeder, J., P. Zhu, Y. Ge, and G. Siscoe, OpenGGCM simulation of a substorm: Axial tail instability and ballooning mode preceding substorm onset, submitted to *J. Geophys. Res.*, 2010.
- Runov, A., V. Angelopoulos, M. I. Sitnov, V. A. Sergeev, J. Bonnell, J. P. McFadden, D. Larson, K. H. Glassmeier, and U. Auster, Themis observations of an earthward-propagating dipolarization front, *Geophys. Res. Lett.*, *36*, L14,106, doi:10.1029/2009GL038980, 2009.
- Sergeev, V. A., P. Tanskanen, K. Mursula, A. Korth, and R. C. Elphic, Current sheet thickness in the near-earth plasma sheet during substorm growth phase, *J. Geophys. Res.*, *95*, 3819, 1990.
- Sergeev, V. A., V. Angelopoulos, D. G. Mitchell, and C. T. Russell, In situ observations of magnetotail reconnection prior to the onset of a small substorm, *J. Geophys. Res.*, *100*, 19,121–19,134, 1995.

- Shiokawa, K., W. Baumjohann, G. Haerendel, G. Paschmann, J. F. Fennell, E. Friis-Christensen, H. Lühr, G. D. Reeves, C. T. Russell, P. R. Sutcliffe, and K. Takahashi, High-speed ion flow, substorm current wedge, and multiple Pi 2 pulsations, *J. Geophys. Res.*, *103*, 4491–4508, 1998.
- Sitnov, M. I., and K. Schindler, Tearing stability of a multiscale magnetotail current sheet, *Geophys. Res. Lett.*, *37*, L08,102, doi:10.1029/2010GL042961, 2010.
- Zhu, P., Tearing instability of Harris sheet in presence of finite normal component of magnetic field, Center for Magnetic Self-Organization in Laboratory and Astrophysical Plasmas (CMSO) General Meeting, Swarthmore College, PA, June 9-11, 2010 (Presentation available at <http://homepages.cae.wisc.edu/~pzhu/pub/pub10/cmso/gha.pdf>), 2010.
- Zhu, P., J. Raeder, K. Germaschewski, and C. C. Hegna, Initiation of ballooning instability in the near-Earth plasma sheet prior to the 23 March 2007 THEMIS substorm expansion onset, *Ann. Geophys.*, *27*, 1129–1138, 2009.
- Zhu, P., J. Raeder, and X. Y. Xing, Global MHD modeling of pre-onset PBI events, 10th International Conference on Substorms, San Luis Obispo, CA, March 22 - 26, 2010, Paper P118, (Presentation available at <http://homepages.cae.wisc.edu/~pzhu/pub/pub10/ics10/pbi.pdf>), 2010.

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.

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° 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 both must 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.

Local Response of the Ionosphere/Thermosphere to High-Latitude Energy Deposition

Gordon Wilson and Dan Ober

The magnetosphere delivers energy to the high latitude atmosphere in the form of precipitating energetic ions and electrons as well as precipitating electromagnetic energy (Poynting flux). The absorption of this energy can lead to local heating and expansion of the atmosphere (as demonstrated by density measurements from GRACE and Champ) as well as the launching of traveling atmospheric disturbances that can circle the globe. The energy absorption process can be complicated owing to the fact that the precipitating ions and electrons can modify the altitude profile of the conductivity which in turn can determine where the electromagnetic energy is absorbed.

We propose a mission designed to monitor the atmospheric response in a local high-latitude region while simultaneously measuring the magnetospheric energy input to that region. This would be accomplished by flying three, identically instrumented small satellites in a string-of-pearl configuration in a high inclination circular orbit at about 450 to 550 km altitude. Each s/c would carry remote sensing instruments to monitor the state of the atmosphere below the satellite and in-situ instruments to measure the energy deposition along the orbit. The satellites would be spaced about 7-15 min apart so that the atmospheric region of interest remains under continuous observation for about 20 to 35 min. This is an appropriate time scale for the altitudinal redistribution of mass in response to a change in the energy deposition rate.

Measurement Requirements:

The measurements that need to be made to support this project include:

1. Auroral emissions from either side of the satellite orbit, ± 600 km.
2. The altitude profile of O^+ below the s/c and within ± 300 km on either side of the satellite orbit.
3. The altitude profile of the main neutral species (O , O_2 , N_2) below the s/c and within ± 300 km on either side of the satellite orbit.
4. The energy spectrum of precipitating ions and electrons along the track of each s/c.
5. The density, temperature and drift velocity of O^+ along the track of each s/c.
6. The neutral density along the track of each s/c.
7. The perturbations to the local magnetic field along the track of each s/c.

Required instrumentation:

1. A scanning UV spectrometer that could image up going UV radiation from the disk below the satellite (cross track scanning) as well as scan through the earth's limb ahead of the s/c (in track scanning). If such an instrument had the limb scanning capabilities of the DMSP SSULI instrument it would be able to return altitude profiles of O , O_2 , N_2 and O^+ (every 90 s) in regions where the solar zenith angle is $\leq 85^\circ$ (or only the O^+ profile for $SZA \geq 108^\circ$).

2. Accelerometer to measure thermospheric density variations at the s/c altitude, as demonstrated on the GRACE satellites.
3. A particle spectrometer to measure precipitating electrons (30 eV – 30 keV) and ions (30 eV to > 100 keV) along the satellite track. Most auroral electron precipitation falls in this range (30 eV to 30 keV) but often the precipitating ion's energy spectrum extends above 30 keV requiring the means to measure ions up to several 100 keV.
4. Retarding potential analyzer and ion drift meter to measure the O⁺ density, temperature and drifts at the satellite altitude. The O⁺ density and temperature data could be used with the UV limb scan data to produce a more robust O⁺ density profile or alone to track the ionospheric response at the s/c altitude. The drift measurements can be combined with magnetic field perturbation measurements to calculate the down going Poynting flux.
5. Magnetometer to measure $\Delta \mathbf{B}$ produced by field-aligned currents.

By combining the disk UV emissions with the measured flux of precipitating ions and electrons it would be possible to determine the time history of the particle precipitation along the s/c track (and on either side) for the 20-35 min interval of the over flights. The electromagnetic energy input would be sampled three times as each satellite over flies the region. The altitude profiles of the neutral species and O⁺ can be combined to determine the altitude profile of Pedersen conductivity in the altitude range where O⁺ is the dominant ion species. This project would require a modeling component (driven by the precipitating particle data) to determine the Pederson conductivity below the F2 peak (in the E region).

It is important to know what the thermosphere and ionosphere are doing on either side of the satellite tracks so as to know when the main energy deposition is occurring to the side of the ground track. Such energy deposition could move mass across the satellite orbits when no energy deposition occurs along the satellite track. This need drives the requirement for measuring auroral emissions from either side of the satellite ground track as well as measuring the O⁺ and neutral density profiles on either side of the satellite orbits.

The final satellite in the string would observe an atmosphere for which the precipitation history is known. Comparison of the state of the underlying atmosphere with that observed by the previous two satellites would indicate how rapidly the atmosphere responds as a function of altitude for a given type and magnitude of magnetospheric energy input.