2.0 SEC Science Objectives and RFAs Introduction

All of the SEC Division strategic planning and implementation, including the missions, technology, theory and modeling, and education and public outreach, is based on the science objectives in Table 2.1 and 2.2. To accomplish these broadly defined science objectives will require a number of missions spread over several decades and a long-term theory and modeling effort.

In order to articulate near- and intermediate-term missions within the SEC, the science objectives in Tables 2.1 and 2.2 are further divided into Research Focus Areas (RFAs). These Research Focus Areas are further divided into Investigations, which describe specific SEC science goals that map directly into individual missions, theory and modeling, and education and public outreach efforts. Within the next decade, these individual missions are expected to make significant progress on the SEC Investigations.

This section details the SEC RFAs and Investigations. In Section 3, the near- and intermediate-term missions are mapped into the Research Focus Areas and Investigations. Section 4 highlights the technology development needed to complete the missions, Section 5 discusses the near- and intermediate-term theory and modeling, and Section 6 discusses the education and public outreach associated with the missions. Finally, Section 7 introduces the external and internal factors that affect the SEC science program.

Table 2.1 Primary Sun-Earth Connection Science Objectives and Research Focus Areas.

Sun-Earth Connection Science Objectives	Sun-Earth Connection Research Focus Areas	Sun-Earth Connection Investigations
Understand the changing flow of energy and matter throughout the Sun, heliosphere, and planetary environments.	- Understand the structure and dynamics of the Sun and solar wind and the origins of magnetic variability.	(a) Understand the transport of mass, energy, and magnetic fields within the Sun and into the solar atmosphere. (b) Determine through direct and indirect measurements the origins of the solar wind, its magnetic field, and energetic particles.
	- Determine the evolution of the heliosphere and its interaction with the galaxy.	(c) Determine the evolution of the heliosphere on its largest scales. (d) Determine the interaction between the Sun and the galaxy.
	- Understand the response of magnetospheres and atmospheres to external and internal drivers.	(e) Differentiate among the dynamic magnetospheric responses to steady and non-steady drivers. (f) Explore the chain of action/reaction processes that regulate solar energy transfer into and through the coupled magnetosphere-ionosphere-atmosphere system.
Explore the fundamental physical processes of space plasma systems.	- Discover how magnetic fields are created and evolve and how charged particles are accelerated.	(a) Discover the mechanisms for creation, annihilation, and reconnection of magnetic fields. (b) Determine how charged particles are accelerated to enormous energies.
	- Understand coupling across multiple scale lengths and its generality in plasma systems.	(c) Understand how small scale processes couple to large-scale dynamics. (d) Test the generality of processes in diverse plasma environments.
Define the origins and societal impacts of variability in the Sun-Earth connection.	- Develop the capability to predict solar activity and the evolution of solar disturbances as they propagate in the heliosphere and affect the Earth	(a) Develop the capability to predict solar activity and its consequences in space. (b) Develop an understanding of the evolution of solar disturbances, how they propagate through the heliosphere, and affect the Earth.
	- Specify and enable prediction of changes to the Earth's radiation environment, ionosphere, and upper atmosphere.	(c) Develop the capability to specify and predict changes to the radiation environment. (d) Develop an understanding of the upper atmosphere and ionosphere response to solar forcing and coupling from the lower atmosphere.
	- Understand the role of solar variability in driving space climate and global change in the Earth's atmosphere.	(e) Understand the connection between solar variability, the Earth's upper atmosphere, and global change. (f) Develop the capability to predict the long-term climate of space.

Table 2.2 Additional Sun-Earth Connection Science Objectives and Research Focus Areas.

Science Objectives	Research Focus Areas
Understand the structure of the universe, from its earliest beginnings to its ultimate fate.	(a) Develop helioseismological constraints on the structure of the Sun as a star.
Learn how galaxies, stars, and planets form, interact, and evolve.	(a) Determine the roles magnetic dynamos and angular momentum transport play in how stars and planetary systems form and evolve. (b) Delineate the current state of the local interstellar medium and its implications for galactic evolution (c) Determine the interaction between the interstellar medium and the astrospheres of the Sun and other Stars
Understand the formation and evolution of the solar system and Earth within it.	(a) Explore the role of planetary magnetic shielding in establishing diverse atmospheres of Earth, Venus, and Mars
Probe the origin and evolution of life on Earth and determine if life exists elsewhere in our solar system.	(a) Explain the role of varying solar activity as life evolves (b) Search for molecules and the building blocks of life from comets, Kuiper Belt objects, and dust in the heliosphere and the interstellar medium. (c) Understand the effects of energetic particles on the evolution and the persistence of life.
Chart our destiny in the solar system.	(a) Explain the role of varying solar activity in the future of terrestrial climate and habitability.

2.1 Understand the changing flow of energy and matter throughout the Sun, heliosphere, and planetary environments.

The flow of energy and matter within and outward from the Sun, past the planets and into their atmospheres, and finally out past the heliopause and into the interstellar medium encompasses all the major scientific disciplines of the SEC Division. Recent spacecraft missions have focused on individual aspects of this coupled system, identifying processes that may contribute to this flow. In addition, widely separated spacecraft have been able to track the flow of energy and matter through large parts of the coupled system, from its origin in the solar atmosphere through to the Earth's upper atmosphere.

These pioneering investigations have identified and isolated specific areas within the coupled Sun-heliosphere-planet system that require new observational methods to progress from observation to understanding. New missions to unexplored regions of the Sun-heliosphere-planet system coupled with data analysis, modeling, and theory will provide new insights to the processes that allow energy and matter to flow across individual components in the system. New understanding will be developed starting with the processes deep within the Sun, progressing through the solar atmosphere and inner heliosphere, including the magnetospheres, ionospheres, and atmospheres of planets within the solar system, to the outer heliosphere and beyond the limits of the solar system.

(a) Understand the transport of mass, energy, and magnetic fields within the Sun and solar atmosphere.

As illustrated in Figure 2.1, helioseismology has revealed a complex pattern of surface flows within the Sun. Synoptic maps of these Solar Subsurface Weather flow patterns produced from localized helioseismology suggest that solar magnetism strongly modulates the flow speeds and directions of the large-scale horizontal flows just below the surface of the solar convection zone from one day to the next. Turbulent convection, rotation, and large-scale flows interact globally to generate the solar magnetic field and produce the quasi-periodic activity cycle. In conjunction with MHD simulations, multi-instrument observations of these flow patterns place strong constraints on models for the solar dynamo, structure, and 22-year solar cycle. To resolve the most significant discrepancies between observations and MHD model results that occur at high-latitudes, a polar monitor will be essential.

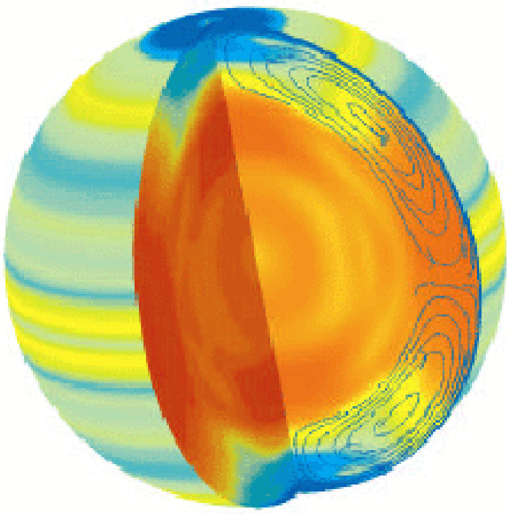


Figure 2.1 Solar rotation and polar flows inside the Sun determined using helioseismology. The colored bands on the surface of the sphere show differences from the average rotation speed of regions on the Sun. Red-yellow is faster than average and blue is slower than average. The light orange bands are zones that are moving faster than their surroundings, extending down approximately 20,000 km into the Sun. Sunspots tend to form at the edge of these bands. The cutaway reveals rotation speed inside the Sun. The large dark orange band is a fast flow beneath the solar equator. Of particular interest are the poorly resolved plasma streams near the poles, which can be seen as a light blue areas embedded in slower moving dark blue regions.

In the absence of fundamental measurements describing the characteristics of the chromosphere and corona, our understanding of the heating that occurs there remains limited. As in all stars with a shell that overlays a strong convection zone, the evolving magnetic field in the atmosphere of the Sun causes heat to be deposited and temperatures to increase to values that are a thousand times higher than that of the solar surface. According to the classical picture, heating occurs within flux tubes reaching from the photosphere into the corona, where coronal energy is conducted downward to form a thin transition region. New measurements indicate that this picture is incomplete. The fact that the chromosphere extends over several thousand kilometers implies that it cannot be hydrostatically stratified, but must be intrinsically dynamic. Moreover, comparison of space and ground-based observations suggests that on small spatial and time scales of 1 arcsec and 15 s, coronal and lower chromospheric heating are not spatially correlated. Either the magnetic topology differs from that expected, or heating within the two domains occurs along different field lines. The classical interface between these two domains, the thin transition region, presents a similar puzzle: it is too extended (as large as 2000km deep). Rather than being a thin layer, it often shows loop-like structures that appear to have no chromospheric counterpart near their footpoints. This leads to the large differences in appearance of images from different layers shown in Figure 2.2.

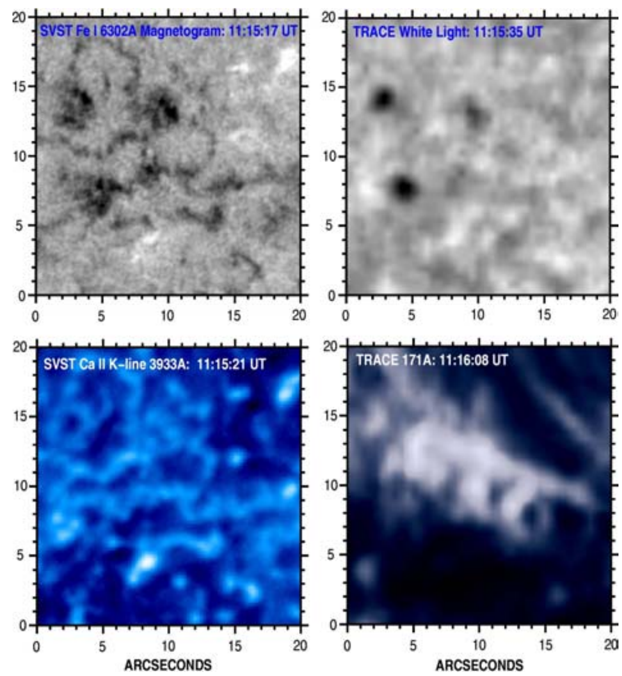


Figure 2.2 Simultaneous views at different altitudes in the solar atmosphere from the chromosphere to the corona of the same solar feature. The lack of correlation between the images shows that, with current time and spatial resolution, chromospheric features do not correspond directly to features in the lower corona.

High spatial and temporal resolution observations of the plasma and magnetic field will be needed to measure the heating processes that occur in the photosphere and atmosphere. The temperature change from a few thousand to a few million degrees across this domain will require multispectral observations from the optical to the extreme ultraviolet.

(b) Determine through direct and indirect measurements the origins of the solar wind, its magnetic field, and energetic particles.

Since the discovery of the solar wind, its true sources and the means by which energy is so rapidly dissipated to heat and accelerate it, have remained elusive. With the remote and *in situ* measurements of solar wind composition from a variety of spacecraft and the determination of large-scale 3-D solar wind structure from the Ulysses spacecraft came the realization that there is not one source, but multiple sources of solar wind plasma. The solar wind is currently classified in terms of fast and slow states that reflect the final speeds of the streams. However, this classification masks a deeper question: is there an inherent bimodality to the solar wind, or is it in truth a continuum of states? *In situ* solar wind measurements made in regions > 0.3 AU from the Sun (>64 Solar Radii) place strong constraints on the means of acceleration of solar wind, the sources of solar wind, and the sources of transient events. Remote UV spectral observations have placed constraints on energy flow and dissipation in the corona near the Sun, wave motion at the base of the corona, and reconnection. Additionally, there exists a critical region below ~20 solar radii where the solar wind is sub-Alfvénic and a region inside ~10 solar radii where the magnetic field often dominates the dynamics. Because of large wave pressures, energetic particles, and non-Gaussian distributions of electrons and ions in this region, the evolution of transient events and shocks is very different than in the region beyond ~0.3 AU.

The region inside ~20 solar radii is also a region where energetic particles are often accelerated. These particle populations are observed at the Sun indirectly by looking at their electromagnetic signatures, or by observing them *in situ* after they escape the Sun's atmosphere and are injected (via open field lines) into the heliospheric magnetic field. The current paradigm is that the electrons are accelerated by shocks in the high corona (>2 solar radii), driven by outgoing CMEs. Since acceleration in most theories is primarily a velocity-dependent process, it is reasonable to expect that very fast ions are also accelerated in the high corona by the same mechanism.

In summary, this region close to the Sun is permeated with plasma in a physical regime that has never been explored, but plays a critical role in controlling the structure, evolution, and variability of the solar wind and particles accelerated in it. Full understanding of this region inside ~20 solar radii requires comprehensive *in situ* and remote sensing measurements of the plasma, fields and energetic particles. Since the corona is essentially 3-D, these near-solar measurements are needed at all latitudes and over as wide a range of solar radii as feasible.

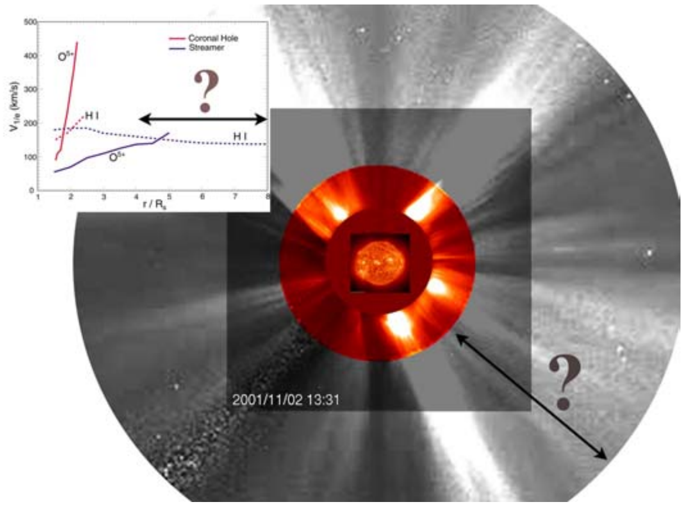


Figure 2.3 In the region inside 20 solar radii, the solar wind is sub-Alfvénic. Because of large pressures and differences in the particle populations, the evolution of solar disturbances is different from that further out. This region remains the last unexplored region of the inner heliosphere that is accessible with today's technology.

(c) Determine the evolution of the heliosphere on its largest scales.

Interplanetary missions in the past three decades have demonstrated that the global configuration of the heliosphere, from the Sun to almost 100AU, is drastically different at the minimum of solar activity compared to the maximum. The large-scale evolution of the solar magnetic field is manifested in the reversals of the Sun's magnetic polarity during each 22-year solar cycle. At solar minimum, fast solar wind from the polar coronal holes maps out into >2/3 of the total volume of the heliosphere. It interacts with the slow solar wind from equatorial latitudes, forming corotating interaction regions (CIRs) of compressed plasma, magnetic field and accelerated energetic particles. On the other hand, at solar maximum the slow solar wind emerges from almost all latitudes. Energetic particle events appear out to 5AU with comparably high intensities from the equator up to polar latitudes >60°. In the distant heliosphere (50-80 AU), during solar minimum there are 26-day recurrent compression regions in the plasma and magnetic field accompanied by increases in energetic particle intensities. These are the remnants of CIRs formed in the inner solar system. In dramatic contrast, during maximum each of the greatest active regions on the Sun produces a series of large solar flares and CMEs that appear a year or more later in the outer heliosphere as huge global merged interaction regions (GMIRs) of plasma, magnetic field, and energetic particles that take months to pass over an individual spacecraft. These great events can even produce episodes of 2-5 kHz radio emissions from the immense region beyond the termination shock.

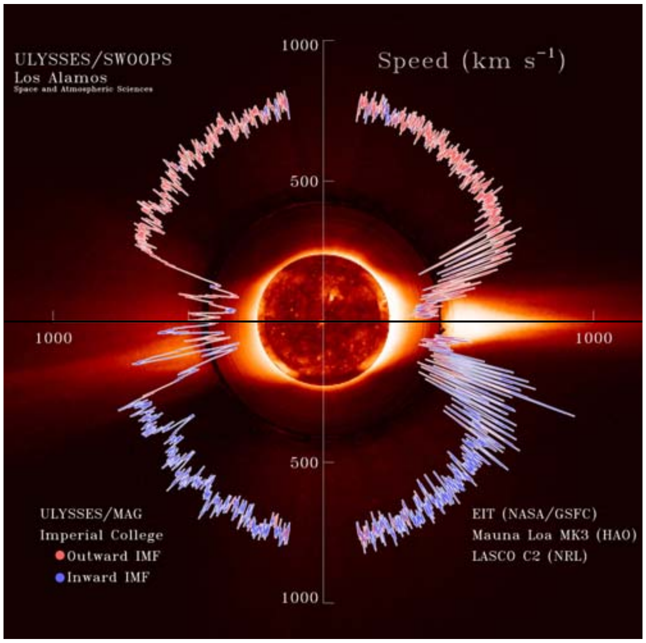


Figure 2.4 The variation in the solar wind speed with latitude observed by the Ulysses spacecraft demonstrates the large-scale structure of the heliosphere. A spatial or temporal interpretation of north-south asymmetries in the heliosphere depends on observations of the high latitude heliosphere over a significant time period.

The critical questions remaining are how the entire heliosphere evolves between the two extreme states and how the solar wind structure/dynamics and energetic particles co-evolve. The basic fact that the solar wind is supersonic (at least for the ions which carry the bulk of the mass and momentum) indicates that physical information describing the plasma propagates outward. Consequently, even though the interactions that form the structure of the outer heliosphere may be quite complex, a great deal of information can be inferred from *in situ* and remote sensing observations of the inner heliosphere (<5AU). Techniques are well established for imaging the solar atmosphere and corona across the electromagnetic spectrum (from the visible through hard x-rays). Outgoing shocks and CMEs can be tracked using "passive sounding" of interplanetary radio bursts. It is now time to extend these efforts to study the phenomena that occur at higher latitudes.

(d) Determine the interaction between the Sun and the galaxy.

Exploration of the interaction of the Sun with the local interstellar medium (LISM) is a voyage into the unknown and holds the promise for a wealth of scientific discovery. This immense structure is truly the outer frontier of the heliosphere in both the sense of matter and of knowledge. The study of the complex boundary itself is a uniquely valuable entry into the basic physics of "astrospheres" (the plasma and magnetic field envelopes of stars having a convective zone like our Sun). The heliospheric termination shock at ~100 AU is where the supersonic solar wind makes its transition to subsonic flow. The immense inner heliosheath region extends another ~100 AU beyond the termination shock. The interstellar plasma then has to deflect its 25 km/s flow around the inner heliosheath (due to the motion of the Sun through the LISM), thus forming the outer heliosheath that is yet another ~100 AU thick in the "upwind" direction. The intervening separatrix between the subsonic solar wind and the interstellar plasma is called the heliopause. Yet further away may be a heliospheric bow shock. Most neutral interstellar matter (thought to be ~10% ionized) flows almost undisturbed throughout this "interface" region. The bulk of the non-ionized gas streams into the inner heliosphere (<3 AU), before photo-ionization by solar extreme-ultraviolet emission or charge exchanged with the solar wind. These ionized atoms and molecules are entrained in the solar wind flow and become "interstellar pickup ions." Definitive numbers for the density, temperature, and bulk flow vector velocity of interstellar hydrogen and helium atoms can be obtained from measurements of the interstellar atoms and interstellar pickup ions in the heliosphere inside 3 AU.

Much about the Sun-galaxy interaction remains unknown. In particular, the properties of the LISM are currently based on estimates and assumptions. Astronomers have learned much about the average properties of the interstellar medium between the Sun and the nearest stars, but because these stars are many light years away, the average properties may differ vastly from those in the LISM. The use of average properties results in large uncertainties in the form and even existence of boundaries such as the heliospheric bow shock. Furthermore, critical factors such as the chemical and isotopic composition, plasma density and temperature, and ionization state, cannot be determined from average properties of the medium.

An entirely new level of information on the interface between the heliosphere and the LISM is required. An extension of the resolution of interstellar pickup ion measurements inside 5 AU to include isotopic ratios of critical astrophysical significance ($^2H/^1H$, $^3He/^4He$, $^{22}Ne/^{20}Ne$, etc.) is needed to make the first real determination of the chemical and isotopic composition of the LISM. These observations may also reveal ionized interstellar molecules. Neutral oxygen is particularly interesting because it provides a diagnostic of the heliospheric interface. Imaging of the outer boundaries ~100 AU away can also be done from orbits near 1 AU. Remote-sensing techniques such as energetic neutral atom (ENA) imaging of energetic protons and EUV (83.4 nm) imaging of oxygen ions near the termination shock and in the heliosheath should provide additional diagnostics of the interfaces with the galaxy. These images may contain information on the magnitude and direction of the local interstellar magnetic field. Surprisingly, the magnitude (and especially the direction) of the local interstellar magnetic field are among the least well known quantities determining the configuration of the Sun's astrosphere. Ultimately, *in situ* measurements at the heliospheric interface with the galaxy will be needed to truly understand the interaction.

(e) Differentiate among the dynamic magnetospheric responses to steady and non-steady drivers.

The magnetosphere responds dynamically to varying solar wind input. The response remains weak and limited to high-latitudes during northward interplanetary magnetic fields (IMF) when the coupling is least efficient, ranges through repeated cycles of substorm activity during intervals of prolonged southward IMF, and increases to giant storms when large-scale, geo-effective interplanetary disturbances, such as CMEs and high-speed streams strike the magnetosphere. Until recently, research centered on "directly-driven" responses to changing solar wind conditions. However, the complex response of the Earth's magnetosphere to relatively steady solar wind conditions is now receiving increasing attention. Even in the absence of solar wind variability, the Earth's magnetosphere exhibits dynamic "loading-unloading" responses that can be triggered by internal processes.

An inability to resolve spatial and temporal variations on a systems level compromises current efforts to differentiate between and understand the responses to time varying and constant solar wind conditions. Understanding the responses will require dedicated solar wind monitors and *in situ* measurements within the magnetotail and other magnetospheric regions at resolutions similar to those provided by current global MHD and kinetic simulations. Indeed, it is the predictions from these simulations that motivate the measurements. The multi-scale measurements at both fine and coarse scales will trace causality, establish linkages, and resolve how mass and energy flows from sources to sinks. Figure 2.5 shows some of the possible structures in the Earth's magnetotail that can be imaged. Magnetospheric imaging techniques, such as stereoscopic energetic neutral atom (ENA) imaging and radio tomography, will provide critical multi-scale measurements of plasma densities without the need to employ large numbers of spacecraft.

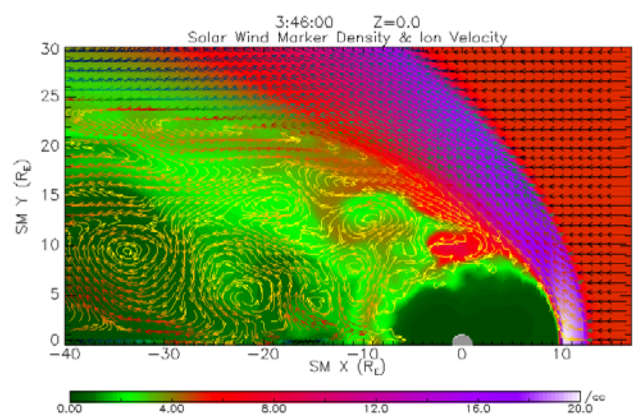


Figure 2.5 Some possible structures in the Earth's magnetotail. Density is color coded and arrows show flows.

In all cases, much of the energy released by magnetospheric activity is dissipated in the ionosphere and thermosphere. The coupling of this energy from high to low altitude is a subject of intense interest, since many of its most dramatic and practically important manifestations involve the low altitude end of the chain. Thus, global imaging of the ionosphere is a critical element in the tracing of energy released by magnetospheric activity.

Finally, differentiating among the dynamic magnetospheric responses to steady and non-steady solar drivers will require extensive monitoring of the solar wind while the multi-scale magnetospheric measurements are being made.

f) Explore the chain of action/reaction processes that regulate solar energy transfer into and through the coupled magnetosphere-ionosphere-atmosphere system.

The electrodynamic interactions between the ionospheric plasma and the thermospheric neutral gas process, redistribute, and dissipate the energy received from the magnetosphere. They also modify the energy exchange process itself through changes in electric conductivity and injection of ionospheric ions into the magnetosphere. While feedback to the magnetosphere is known to exist, its importance is not always clear. For example, it is not known what role the ionosphere may play in facilitating or quenching the development of a magnetic substorm, but the ionospheric plasma plays a critical role in intensifying the ring current during major magnetic storms.

The lower part of the Magnetosphere-Ionosphere-Atmosphere (MIA) system lies at the end of an extended chain of processes. These processes involve the conversion of solar energy that has passed through the MIA system into an upward flow of energy. This conversion is due primarily to wave motions and wave breaking of various types (latitude-seasonal variations, tides, planetary waves, gravity waves, etc.), which occur in the upper mesosphere and lower thermosphere. It is expected that changes in the structure of the neutral atmosphere will affect electrical conductivity in the ionosphere and dynamo action and thus be connected to the ionosphere and magnetosphere. Transport and chemistry of water vapor is another physical process that is dramatically affected by forcing from below. Through a poorly understood dusty plasma interaction, this water vapor is involved in the formation of noctilucent clouds over the summer polar cap. The presence of these clouds has important implications for global climate change, since they modify the Earth's albedo. Finally, tropospheric thunderstorm activity may influence the electrical and chemical properties of the middle and upper atmosphere through upward directed lightning discharges and upward conduction currents. Little is known about the upward flows of mass, energy and momentum but these final links in the coupling likely contain the processes ultimately responsible for modulating weather and climate.

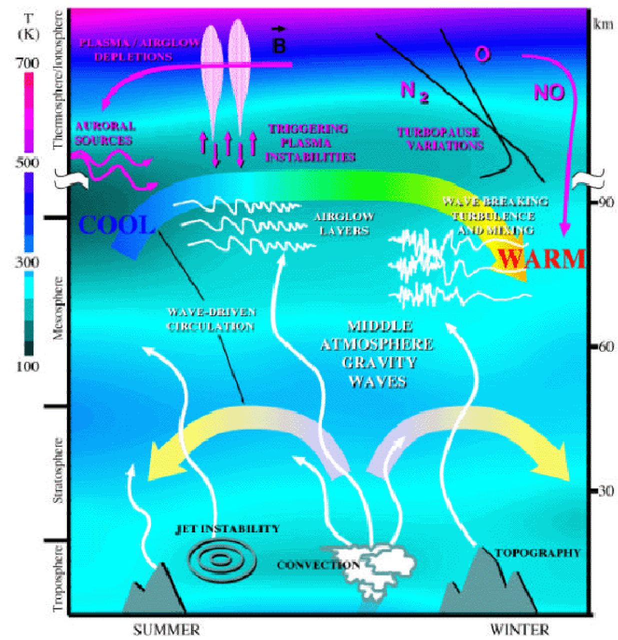


Figure 2.6 Latitude, height cross-section illustrating small-scale wave effects in the ionosphere and atmosphere. Small-scale waves are the fuel driving the summer to winter transport and dynamic coupling between the lower and upper atmosphere. Solar variability influences composition and transport rates. Quantifying the transport between the upper and lower atmosphere requires a better understanding of the relative forcing from above and below the regions.

The steady-state view of the linkage between processes that begin on the Sun and proceed outward/downward through the magnetosphere – ionosphere - atmosphere (MIA) system is slowly being replaced with a dynamical, collective view. There is a two-step path to this better understanding of the system. The first step is a detailed understanding of the dynamical behavior of the individual geospace elements is needed, in particular the ionosphere and atmosphere elements. The second step is an understanding of the interplay between components and feedback processes that dictate the collective global system response. The first step requires detailed observations (either *in situ* or imaging) targeted to a specific region in the

MIA system, and the second step requires combinations of observational techniques and vantage points that target the entire system.

2.2 Explore the fundamental physical processes of space plasma systems.

Most of the universe contains *plasma*, a gas of approximately equal concentrations of positive and negative charged particles. This state of matter behaves in fascinating and complex ways, quite unlike the behavior of neutral gas. Whereas the plasmas in the vastness of the universe are not directly accessible to humans, there are numerous and diverse cosmic "plasma laboratories" within the solar system that are accessible to space missions. Comparison of the near-Earth plasma laboratories with others in the solar system has underscored the amazing diversity among these systems. For example, the magnetic poles of Uranus lie close to the ecliptic plane, presenting a completely different magnetospheric configuration from that at Earth. Mercury's magnetosphere is small and its atmosphere is tenuous, while Jupiter and Saturn's magnetospheres are huge and the presence of a dense atmosphere, moons, rapid planetary rotation, and electrically charged dust and ice play significant, even dominant roles in magnetospheric dynamics.

Although there is a diversity of plasma systems, there is also a commonality of basic physical processes within these systems. Thus, nearby plasma laboratories provide an opportunity to explore the fundamental properties of plasmas that are directly applicable to many other regions of the universe. Three properties of plasmas that are especially important to understanding Sun-heliosphere-Earth connections also, not by coincidence, play a role in many important astrophysical phenomena. These plasma properties are the creation, support, and annihilation of electric and magnetic fields, the acceleration of charged particles to ultra-high energies, and the coupling across physical scales. All of these fundamental properties are related ultimately to the capability of a plasma to maintain electric and magnetic fields. The desire to discover the physics behind these fundamental properties of plasmas leads to three distinct research focus areas.

a) Discover the mechanisms for creation, annihilation, and reconnection of magnetic fields.

Motions in a plasma can generate, modify, and dissipate magnetic fields. Such dynamo action is one of the basic physical processes that determine the nature of the observable universe. This happens in stars, galaxies, the interior of planets, and protostellar clouds of tenuous gas. Most important for the Sun-Earth system are the generation of magnetic fields in the Sun and planets.

Simple models can reproduce cyclic solar magnetic fields in idealized simulations, but they have limited predictive value. Current theories suggest that the solar magnetic field is generated in the "tachocline," a region of strong radial shear between the convection zone of the Sun and the radiative interior, as illustrated in Figure 2.7. However, current models do not reproduce the large-scale structure of the convection zone or the tachocline, much less the detailed generation of magnetic flux and its eruption through the photosphere. The next generation of models requires further developments in computational methods and more complete knowledge of sub-photospheric conditions. Helioseismology holds great promise for detecting strong concentrations of magnetic flux in the convection zone and for measuring the large-scale motions within the convection zone.

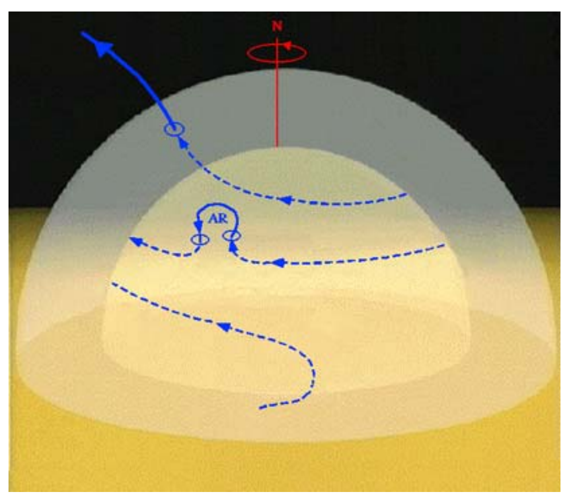


Figure 2.7 The tachocline field line shown in the figure is confined near the bottom of the solar convection zone and has been wrapped up by differential rotation. The deep intense flux tubes may develop instabilities that rise, emerging as active regions in the photosphere. Near the surface, small-scale fields may be generated by intense motions.

Dynamos in some planetary cores, including Earth's, also create magnetic field systems that have an important influence on solar wind interactions. The planetary community investigates these dynamos. However, similarities in the techniques used to study these fields provide cross-fertilization of ideas about the solar and planetary dynamos.

Magnetic field reconfiguration can result in the direct conversion of magnetic energy into plasma kinetic energy (i.e., particle acceleration, heating, and bulk flow) through a process known as magnetic reconnection or magnetic merging. This process is believed to accelerate particles in solar flares, at the Earth's magnetopause, and in the Earth's magnetotail. It is also probably fundamental to the mechanisms by which plasma is energized in the corona and solar wind and is a means for energy transfer in a variety of astrophysical systems. Despite the importance of this process, there is still no fundamental answer to the question: "how do magnetic fields undergo reconnection?" Clues to the answer to this question are found in understanding the stability of the reconnection process and which types of magnetic field configurations are conducive to the process.

Understanding the topology and stability of the magnetic field configuration is crucial for understanding reconnection. In the photosphere, for example, the process has been inferred from the merging and disappearance of myriads of small magnetic flux elements with opposite polarities. These small elements vary little with the 22-year solar cycle and likely result from a distributed near-surface dynamo sustained by small-scale convective motions. This small-scale dynamic field may be in part responsible for driving the solar wind. By imaging the Sun in higher spatial, temporal, and spectral resolution, the topology, physical conditions, and perhaps inferences of micro-instabilities in the reconnection regions will be determined. These improved observations will also advance the understanding of the regions of particle acceleration in solar flares, the photosphere in general, and processes operating in the solar corona.

Single-point and even multi-point measurements of reconnection occurring near Earth lack the global information on reconnection topology available with solar imaging. However, the ability to sample the reconnection region *in situ* provides the opportunity to directly determine micro-instabilities responsible for magnetic reconnection. These micro-instabilities act in the relatively small volume in space called the diffusion region, where electrons decouple from the magnetic field. The electron diffusion region has not been investigated because past *in situ* missions have lacked the proper instrumentation and the multi-point perspective to determine its properties. Furthermore, theory and particle simulations had not progressed to the point where predictions could be made about the nature of this region. Armed with specific predictions from theory, new, more detailed observations will be made to reveal the nature of this critical region in the magnetosphere.

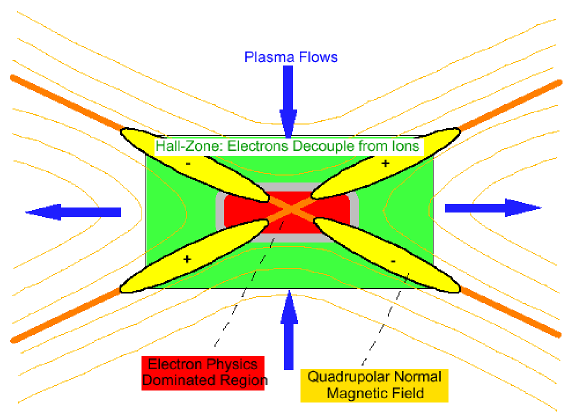


Figure 2.8 The structure and topology of magnetic fields that undergo magnetic reconnection. Magnetic reconnection is dominated by the physics in the very small (~100 km thick) region called the electron diffusion region. High-resolution, *in situ* measurements from a multi-spacecraft mission will reveal the nature of this energy transfer process.

b) Determine how charged particles are accelerated to enormous energies

Charged particles are accelerated in the solar atmosphere, interplanetary space, planetary magnetospheres, the interstellar medium and more-distant astrophysical objects. They exhibit energies ranging from just above the modest thermal energies of the ambient plasma to cosmic rays that enter the heliosphere with enormous energies exceeding 10^{20} eV. It is quite remarkable that energetic particle populations in diverse regions within the heliosphere display the same kind of distribution of number versus energy---a "power law" energy spectrum.

This remarkable property raises the question of how to accelerate a small fraction of the charged particles from the ambient plasma while creating a power law energy distribution that is not very sensitive to the parameters of the particular acceleration process or region of acceleration. Candidates include impulsive particle acceleration by inductive electric fields, resonant acceleration associated with the myriad periodic particle and field modes of a magnetized plasma, and Fermi acceleration during multiple reflection between two magnetic structures approaching each other. One important acceleration site is at collisionless shocks. They are generated throughout the heliosphere, from the solar corona to the termination shock at about 100 AU that marks the outermost limits of the solar wind. Coronal mass ejections are particularly effective drivers of transient shocks in the heliosphere, while the pervasive overtaking of low-latitude slow-speed solar wind by high-speed solar wind from coronal holes produces "corotating" forward/reverse shock pairs. *In situ* observations at all shocks reveal populations of energetic particles. Conversely, understanding the propagation of energetic particles from remote regions along magnetic fields allows inference of the acceleration processes within the region. An example of great interest currently is the population of "anomalous" cosmic rays (ACRs with energies ~100 MeV). Current theories ascribe these to "pickup ions" accelerated at the heliospheric termination shock, i.e., galactic gas atoms that are ionized while drifting though the heliosphere and then swept out to the boundary by the solar wind. This process is diagramed in Figure 2.9. Right now, the strongest observation-based information on the nature of the termination shock comes from inferences drawn from the comparison of acceleration and propagation theory with ACR observations. However energetic particles are also widely found where there are no shocks, e.g., in the trapped radiation of planetary magnetospheres, so other non-shock-associated acceleration mechanisms are required. There is considerable debate about how and even where particles are accelerated in solar flares, even though there are remote sensing of the acceleration process through the tell-tale radio, x-ray and gamma-ray emissions.

Both *in situ* and remote sensing measurements from spacecraft are necessary to determine the modes of acceleration and the transport of energetic particles. In planetary magnetospheres, the solar corona, and throughout the heliosphere, it is essential to go to unexplored regions and novel vantage points to gain this information. Coverage of the energetic particle and electromagnetic spectrum should be as wide as possible. Examples of important vantage points are the polar-regions of the Sun (at distances much closer than the 2.4 AU passes of Ulysses) and of a giant planet like Jupiter, out to the heliospheric boundary, and within the Earth's magnetosphere at high and low altitudes.

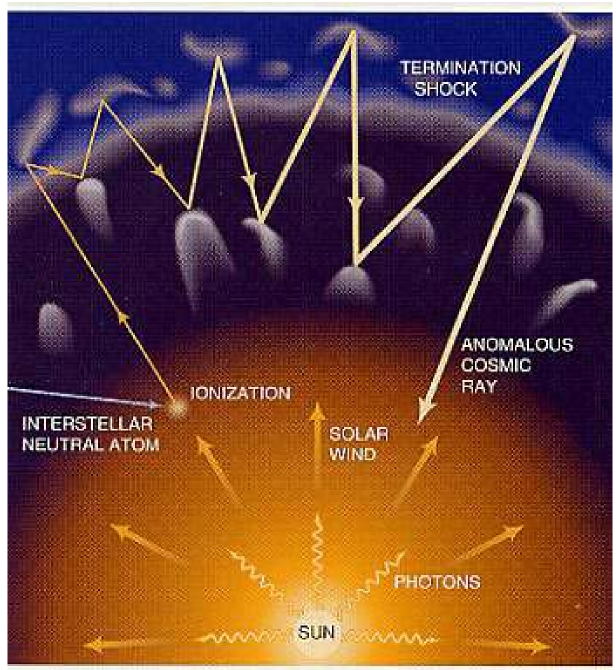


Figure 2.9 Anomalous cosmic rays are one of many examples of energetic particle production and transport. These energetic particles may be accelerated by ambient wave turbulence and by acceleration at the termination shock. By studying these and other energetic particle populations in other regions, the primary acceleration mechanism(s) will be revealed.

c) Understand how small-scale processes couple to large-scale dynamics.

Magnetic reconnection (discussed above in 2.2a) is an important example of small-scale processes that couple to large scales. The small-scale reconnection of magnetic fields in the solar corona can destabilize the magnetic field in an entire hemisphere, resulting in a CME. Similar small-scale to large-scale coupling attributed to the reconnection process occurs in the Earth's magnetotail. Progress on the understanding of annihilation of magnetic fields requires an understanding of relevant small-scale processes.

Another prime example of how small-scale processes affect global dynamics is the formation of narrow auroral arcs in the Earth's auroral zone (at latitudes above ~65°). Observations show that particles are accelerated by moving through localized parallel electric fields in these filamented structures. These narrow regions may be fast flow channels that extend deep into the Earth's magnetotail. The difficulty with current observations is that the highest time resolution measurements have been made from single spacecraft and have not resolved spatial and temporal processes. Furthermore, the detailed, *in situ* measurements have not been accompanied by similar resolution (~km scale length) imaging of the aurora. Significant progress in unraveling the contributions of these structures to large-scale dynamics and the time dependence of these structures requires high time resolution, multi-point *in situ* and imaging measurements in the auroral zone.

Turbulence is another very important multi-scale process. Turbulent processes transport particles and fields effectively, but are not well understood. Numerical simulations and laboratory experiments demonstrate that, in the presence of rotation or magnetic fields, turbulent motions create both small-scale and large-scale dissipative structures via "inverse cascades". Examples of this dissipative process are found in the small-scale magnetic interactions in the chromosphere and transition regions of the solar atmosphere and their possible coupling to the corona and solar wind. Further, the solar wind itself has been observed to evolve toward a fully MHD turbulent state as it propagates toward the magnetosphere. In the magnetosphere, thin boundary layers such as the magnetopause may be unstable to turbulent plasma processes (see Figure 2.10). Non-linear growth and saturation of these processes may lead to enhanced particle transport and heating. Finally, turbulence probably plays an important role in the Earth's magnetotail. Recent observations of current disruptions and bursty bulk flows and their correlation with large magnetic field fluctuations appear to have the stochastic (i.e., turbulent) properties.

Fundamental questions concerning turbulence remain unanswered. They include the role of fluid turbulence in transport across plasma boundary layers, the control of the onset of turbulence in thin current sheets, and the processes that drive micro-turbulence and its coupling to large-scale disturbances. Finally, it is not known how turbulence affects the mapping of magnetic fields and the predictability of plasma systems.

By its very nature, turbulence is a multi-scale process and therefore requires multi-point measurements. Multi-point *in situ* observations on a variety of spatial scales within the solar wind and magnetosphere can provide the information needed to characterize turbulence in these regions. Some questions, such as small-scale turbulence in the solar wind or local, turbulent transport across thin boundaries, may be answered using a small number of multi-point measurements, while others, such as small-scale to large-scale coupling in the solar wind or magnetosphere, will require a larger number of multipoint measurements with a variety of separation scales. While solar observations of turbulence cannot be made *in situ*, the multi-scale nature of solar turbulence can be investigated using multi-spectral imaging.

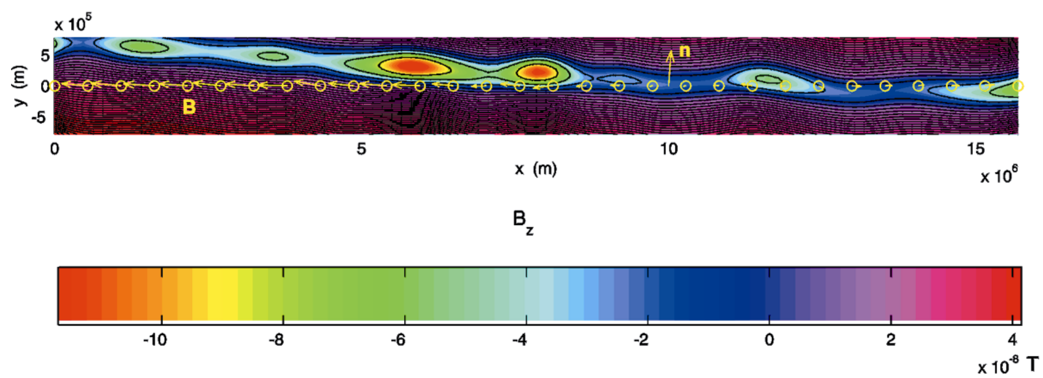


Figure 2.10 Turbulent magnetic field structures, or "islands" inferred from a single spacecraft trajectory through the Earth's magnetopause. The contours and colors represent the strength of the magnetic field. These island structures may grow to saturation and lead to enhanced particle transport.

d) Test the generality of processes in diverse plasma environments.

While the diversity of plasma systems often leads to new and different plasma discoveries, it also affords an opportunity to test the generality of processes that are thought to be fundamental. Tests can be performed using observations in the heliosphere, and the magnetospheres and ionospheres of the Earth and other planets.

The ubiquity of energetic particles or cosmic rays in the rarefied space plasmas provides an opportunity to test a fundamental plasma process in diverse plasma environments. The importance of energetic particles in plasma environments ranges from minimal to significant. The properties of a plasma where energetic particles play a significant role are poorly understood. Detailed energetic particle measurements proceeding from the inner heliosphere and near strong and weak shock waves, to remote observations of regions such as the solar wind termination and interstellar shocks provide a range of measurements that allow determination of the progressive importance of energetic particles in plasmas.

The magnetospheric substorm appears to be a fundamental dynamical mode of Earth's magnetosphere. If this is the case, then this process of magnetic flux conversion probably occurs in other planetary magnetospheres. Scant evidence of substorm-like phenomena exist from observations in the magnetospheres of Mercury and Jupiter. While the substorm (or something very much like it) appears likely to be a feature of planetary magnetospheres in general, it is not understood how this systematic behavior operates so similarly in such vastly different planetary conditions and environments. *In situ* measurements at other planets similar to those available at Earth are needed to answer this question.

An important test of cross-scale coupling and its effects in diverse plasma environments is the electric forcing between different plasmas connected by a magnetic field. At Earth, this process occurs in several regions including the auroral zone. An ideal environment for testing the theories about magnetic field forcing and M-I coupling is Jupiter. Jupiter has the most powerful aurora in the solar system. It is driven by the breakdown of magnetospheric rotation associated with the shedding of angular momentum from the central body to the surrounding nebula by means of magnetic fields and field-aligned currents. It is not known why these field-aligned currents are so highly structured in space and time, whether this structuring is a fundamental aspect of momentum transfer, how the field-aligned impedance is established, and what effect this impedance has on the forcing that occurs between the magnetosphere and ionosphere. Answers to these questions require a detailed comparison of the auroral generation mechanisms at Earth and Jupiter obtained from a combination of imaging and *in situ* measurements at both planets.

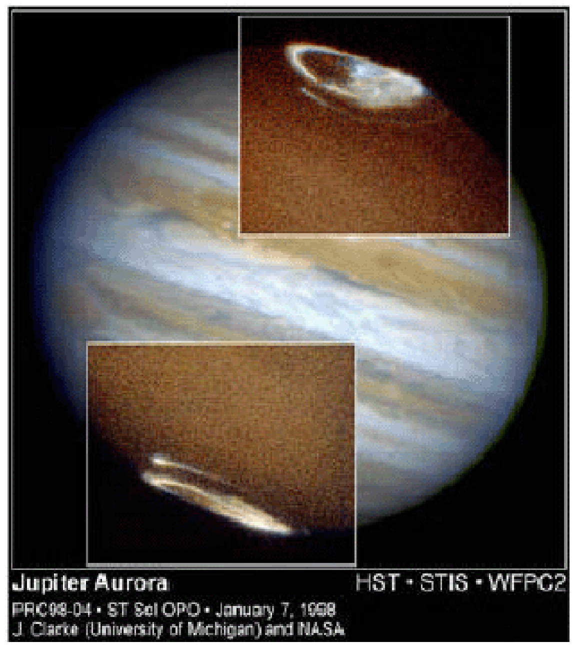


Figure 2.11 The Jovian aurora is the most powerful in the solar system. The aurora results from the shedding of very small amounts of the planet's angular momentum. Study of the fundamental process of auroral formation will be possible when combined imaging and *in situ* measurements of the Jupiter aurora become available.

Finally, the phenomena associated with dusty plasmas are important in a variety of plasma environments such as in the Earth's upper mesosphere, comets, planetary rings, and in interstellar space. The dust provides a sink for plasma and, when charged, affects instabilities in the plasma. One place where the effects of dust in a plasma can be studied is in the mesosphere. Here, the physics of dusty plasmas involves plasma instabilities and polar mesospheric clouds, with dust provided by oblation of meteoroids. The nature of the chemical, dynamical, and plasma processes leading to a variety of phenomena in this region are not well established. Answers require *in situ* measurements of the plasma and neutral gas (including composition) in the region.

2.3 Define the origins and societal impacts of solar variability in the Sun-Earth connection.

Human exploration of space and increasing reliance on space technology drive the need to determine the origins and societal impacts of solar variability. Societal impacts of interest include radiation doses on spacecraft, astronauts, and passengers and crew in high-altitude aircraft; solar induced effects on navigation systems, communication signal paths, and ground electrical currents; and changes in atmospheric structure, chemistry, and climate.

The understanding of the Sun-Earth system has progressed to the point where solar phenomena can often be observationally linked to changes in the heliosphere and in the Earth's magnetosphere, ionosphere, and atmosphere. It is now necessary to develop a more complete understanding of the processes in the Sun-Earth system that lead to specific societal impacts. This system-wide understanding will start with the underlying causes of solar variability, progress through the heliospheric changes to this solar variability, and end with the coupling of this variability to the Earth's magnetosphere, ionosphere, and atmosphere. The next step will be to forecast solar variability and the corresponding magnetospheric and ionospheric responses in the same manner as current weather and long-term climate forecasts.

(a) Develop the capability to predict solar activity and its consequences in space.

Solar activity, as reflected in the sunspot cycle, flares, coronal mass ejections, and changing solar emissions over the solar cycle, is the starting point for defining societal impacts. Phenomena that affect the Earth originate beneath the solar surface, but are often triggered in the solar atmosphere. Advances in helioseismology and simultaneous imaging of the 3-D evolving solar atmosphere make it possible to develop a firm understanding of the solar activity that most influences geospace.

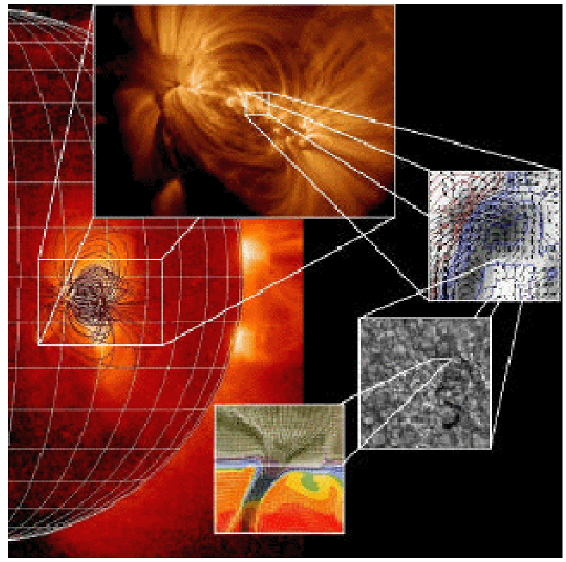
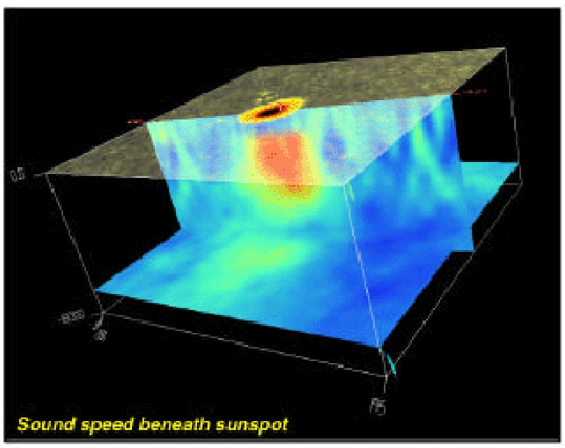


Figure 2.12 A series of images showing an active region, overlying loops, a vector magnetogram of an active region, granulation, and finally a high-resolution model of the magnetic transition region overlying a granule. Evolution on various scales leads to reconnection. Reconnection leads to heating seen in the magnetic loops and also leads to flares.

The birth and evolution of sunspots, active regions, and complexes of activity can be studied using helioseismology. Figure 2.12 reveals a snapshot of the evolving subsurface structure associated with the magnetic field beneath a sunspot. Observations have shown that once an active region emerges, there is a high probability that additional flux eruptions will occur nearby. Current thinking suggests that the flux emerging in active regions originates in the tachocline, but that the photospheric distribution of flux requires no long-term connection to flux below the surface. The critical element in this process is the timing of the weakening, or complete disconnection, of the link between the active region flux and the deeper ropes. This timing needs to be studied by developing uninterrupted sound-speed and flow maps under the visible surface of the Sun combined with surface magnetograms. High-resolution helioseismology will determine if active regions disconnect rapidly from sub-surface flux and will establish the feasibility of using helioseismology to predict active region emergence and the formation of long-lived complexes of solar activity.

Magnetic field reconfigurations release stored solar energy into the heliosphere as radiation, plasma flows, and energetic particles that some-times impact Earth. Establishing the details of processes like the reconfiguration of magnetic flux in supergranules or the eruption of sigmoid magnetic field configurations within the corona into CMEs will help identify the sources of the solar wind, predict the occurrence of transient events such coronal mass ejections, flares, and solar energetic particle events, and determine how they propagate through the heliosphere. To address these questions, accurate high-resolution measurements of the vector magnetic field at the Sun over the wide range of spatial scales illustrated in Figure 2.13 will be needed in conjunction with other remote observations and *in situ* plasma, fields, energetic particle, and composition observations of the solar wind.



Sunspot data from MDI High Resolution, 18 June 1998

Figure 2.13 The magnetic structure below a sunspot is revealed in this plot of sound speed derived using helioseismology. Three planes are presented. The top shows and image of the sunspot with a dark central umbra. The second is a vertical cut (to a depth of 24,000 km) showing areas of faster sound speed as red and slower as blue. The third (bottom) is a horizontal cut showing the horizontal variation of sound speed over a 150,000 square km.

(b) Develop an understanding of the evolution of solar disturbances, how they propagate through the heliosphere, and affect the Earth.

Following their creation at the Sun, heliospheric structures, such as CMEs and co-rotating interacting regions (CIRs) evolve and are modified as they travel outward through the heliosphere. Propagating heliospheric structures can generate strong shock waves that strike the magnetosphere and accelerate energetic particles that enter the Earth's radiation environment. As they move beyond 1 AU some of these disturbances merge to create massive global merged interaction regions (GMIRs) that shield the Earth by modulating galactic and anomalous cosmic rays.

Given the sparseness of heliospheric spacecraft, most *in situ* observations of heliospheric disturbances have been 1-D. However "passive sounding" of CMEs from a distance, using naturally occurring radio emission generated by co-traveling populations of suprathermal electrons (~1-10 keV), has been used to trace the motion of the CMEs out into the heliosphere. Observations of the propagation and development of shocks near 1 AU have been shown to agree reasonably well with simple shock theory, but substantial research is needed to understand how the large-scale magnetic field, plasma flows, and turbulence properties vary throughout the three-dimensional heliosphere.

To unravel the evolution of inherently three-dimensional structures, multiple spacecraft observations will be necessary. Multi-point observations inside Earth's orbit, in conjunction with global imaging of the corona, and remote sensing of the inner heliosphere will help the understanding of the 3-D changes in CMEs as they propagate toward Earth and will help relate these changes back to their solar origin.

(c) Develop the capability to specify and predict changes to the radiation environment.

The understanding of magnetospheric dynamics underwent a revolution following March 1991 CRRES satellite observations. These observations indicating that an entirely new belt of >25 MeV electrons was produced in the magnetosphere in a matter of minutes. Additional observations have now shown that the radiation belts are highly structured and highly dynamic, exhibiting variability on time scales of minutes, days, season, and solar cycle (see Figure 2.14). At the same time, their impact on a technology-based society was increasing with every new satellite system placed into Earth orbit.

Although it is known that solar particle events provide a source population for the radiation belt enhancements and that these enhancements occur in association with high-speed solar wind streams and shock compressions of the magnetosphere, much work is still required to clearly define the physical processes that link these phenomena. Because radiation belt enhancements can occur on time scales shorter than spacecraft orbital periods (typically > 12 h), multiple spacecraft measurements are needed to define the dynamics of the belts.

Radiation belt losses occur through a combination of collisions with cold plasma, interaction with plasma waves, (both generated by magnetospheric processes and by tropospheric lightning), by scattering in magnetospheric current sheets and by drifts out of the dayside magnetopause during compressions (called magnetopause shadowing) and by adiabatic energy changes in response to large-scale magnetic field disturbances. The coupling to the ring current (whose strength and dynamics are influenced by outflows and electrodynamic coupling with the ionosphere-atmosphere system) and to propagating plasma waves from tropospheric lightning means that a complete understanding of the radiation belt structure and dynamics requires treating the ionosphere-thermosphere, radiation belts, inner magnetosphere, and the heliospheric input as an integrated system. This in turn requires simultaneous multi-point measurements of radiation belt and ring current particles as well as the electric and magnetic fields in the various regions. Time dependent radial profiles of relevant quantities must be obtained in order to differentiate among various physical mechanisms such as radial diffusion vs. localized acceleration. All of these measurements must be placed into context with input from the Sun (both photon and particle input) and the response of the ionosphere/atmosphere.

The true test of physical understanding lies in the development of new computational and empirical models of the radiation belts. These next generation models must be based on improved physics and assimilate observations in order to enable a future space weather capability, these models will need to be time-dependent and data-driven. They must also apply over sufficiently long time scales to enable reliable and cost-effective spacecraft design.

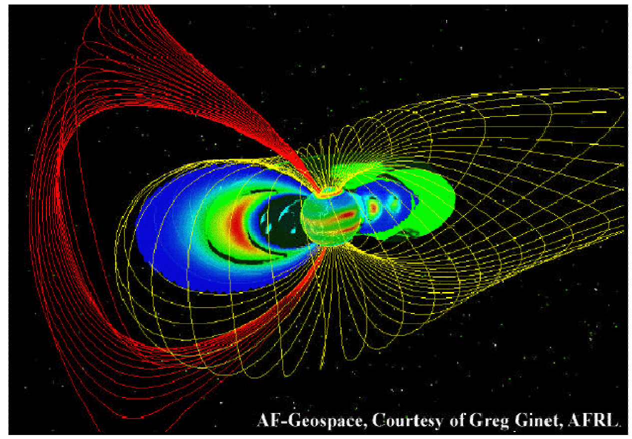


Figure 2.14 The energetic electron radiation environment around the Earth is generally structured into an inner and outer belt clearly seen in cross-section on the left-hand side of the figure. During high-speed streams and/or solar wind disturbances the belts can intensify and move closer to Earth, and the slot region separating the belts can temporarily fill. Occasionally transient, shorter-lived belts (depicted by the small enhancement between the inner and outer belts on the right-hand side) are generated during impulsive events like solar wind shocks or solar particle events. The auroral oval appears at high latitudes on the Earth and maps along field lines to larger radial distances than the radiation belts. Also depicted in two-bands on either side of the equator are equatorial arcs.

(d) Develop an understanding of the upper atmosphere and ionosphere response to solar forcing and coupling from the lower atmosphere.

The highly variable space weather phenomena that occur within the ionosphere and atmosphere represent the most significant threats to human endeavor. This variability is driven by changing solar x-ray, far ultraviolet and extreme ultraviolet (FUV and EUV) radiation as well as by electrodynamic coupling, heat and particle fluxes from the magnetosphere. Gravity waves and tides that propagate into the upper atmosphere can also appreciably modify the basic state and structure of the ionosphere-atmosphere system on a variety of scales. This influence from below is the least understood of all the effects on the ionosphere. Ionospheric structures that develop in response to these inputs span a range of scales from centimeters to 1000 kilometers.

Over the next decade, a revolutionary new view is expected of the dynamical behavior of the magnetosphere-ionosphere-atmosphere (MIA) system as it responds to energy inputs from the Sun and lower atmosphere. This understanding of the "systems" response will follow directly from new multi-point and remote-sensing observational approaches that create snapshots of large regions of the system. The new view will be instrumental in enabling detailed understanding of space weather disturbances in the ionosphere-upper atmosphere system that interrupt communications and cause errors in surveying and navigation systems, introduce background variability within various space-based surveillance systems, generate ground induced currents that interrupt power grids and erode pipelines, and impact satellite lifetimes and satellite tracking capabilities by increasing drag on low Earth orbiting spacecraft.

Of these effects, a highly-focused investigation into the generation of ionospheric irregularities that degrade communications and navigation systems is considered to be a top priority, both because of the severity of the societal consequences and the high probability that major scientific advances in physical understanding and predictive capabilities will result from this effort. At high-latitudes, investigations of irregularities continue through polar-orbiting missions and, at low latitudes, global characterization begins through new observations from space. At mid latitudes, irregularities have a dramatic impact on navigation systems. Global characterization of the ionospheric density and structure by GPS receivers and simultaneous imaging of the inner magnetospheric plasma populations have recently established the association of these irregularities with stormtime changes in the inner magnetosphere. As plasma density fronts move through the mid-latitude ionosphere during these times, steep plasma density gradients play havoc with technologies like the global positioning system (GPS), and ionospheric irregularities disrupt communication systems. As much as 120 m errors in single-frequency GPS positions have been observed in association with this dramatic redistribution of ionospheric plasma.

High levels of geomagnetic activity are known to produce large time-dependent disturbances in the ionosphere and thermosphere that extend to mid-latitudes from both equatorward and poleward sources. Enhancements linked to the poleward transport of equatorial plasma occur during the growth phases of geomagnetic storms. These intrusions of high-density plasma (shown in Figure 2.15) are associated with distorted stormtime electric field patterns resulting from coupling between the partial ring current and ionosphere. During the recovery phase, depletions in the ionospheric plasma appear at mid latitudes. Strong and persistent heating in the auroral oval results in upwelling of the neutral atmosphere in this region with increased density at high altitude, then advection to lower latitudes and downwelling. Associated composition changes in the neutral atmosphere combined with enhanced flows result in ionospheric density depletions that can persist for more than a day. Modeling these mid latitude phenomena is extremely difficult.

The role of upward propagating gravity waves and tides in triggering and generating equatorial irregularities is as important as the more familiar "downward" propagating effects. Large-scale bubbles with dimensions of several thousand km along the magnetic field direction and east-west dimensions of several hundred kilometers appear in the equatorial ionosphere at local sunset. Smaller scale irregularity structures (tens of km to tens of meters) develop through a hierarchy of instability mechanisms. The trigger mechanisms that cause the extreme day-to-day variability have yet to be isolated and may involve gravity waves in the neutral atmosphere, disturbed stormtime electric fields and/or neutral wind patterns. The mechanisms producing day-to-day variability and the longitudinal extent of equatorial irregularities on any given evening remain outstanding problems.

Identifying the mechanisms driving instabilities in MIA coupling will not only require *in situ* measurements of the irregularity spectra, wave characteristics, electric fields, neutral winds, and plasma density gradients in the ionosphere and thermosphere, but also imaging and wave climatology. Airglow imaging of mid-latitude regions will provide the dayside O/N_2 ratios and nightside line-of-sight electron densities needed for the global context. ENA imaging of the ring current structure and dynamics from a high altitude polar platform will provide further details concerning coupling with the mid-latitude ionosphere. Simultaneous FUV images will define is the global magnetospheric dynamics and indicate high-latitude composition disturbances spreading to mid latitudes.

Finally, because the role of upper atmospheric waves in triggering instabilities is unknown but thought to be significant, a global climatology of their occurrence patterns will prove essential. The successful development of global circulation models that specify the state and evolution of physical quantities over large regions will prove the ultimate test of our understanding.

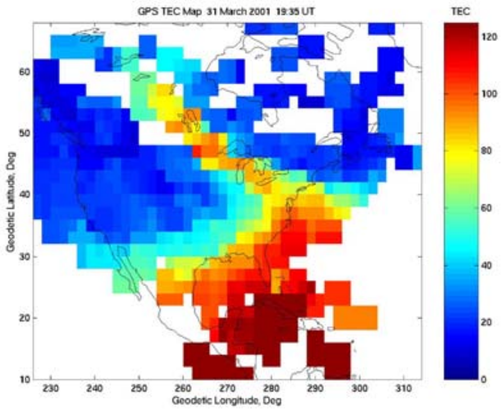


Figure 2.15 GPS map of total electron content in the ionosphere over the United States showing the intrusion of high density low-latitude plasma into midlatitudes as it moves along distorted magnetic stormtime convection patterns toward the polar regions. Steep plasma density gradients cause large position errors in single-frequency GPS navigation systems. The complexity of coupling and feedbacks between geospace regions involved in producing this phenomena makes a true predictive capability difficult to achieve.

(e) Understand the connection between solar variability, the Earth's upper atmosphere, and global change.

The upper atmosphere (altitudes above ~90 km) is important for climate change studies both as an active participant in modulating climate and as a sensitive (and thus early) indicator of possible changes already underway. These changes may not necessarily be induced from anthropogenic sources. For example, a connection between the global mean temperature and the solar activity cycle has been suggested on the basis of statistical evidence. However, observed changes in the solar constant over a solar cycle are too small to be the source of this relationship. Other processes that may provide the basis for such a correlation include: changes in the spectral irradiance which drive the chemistry and dynamics of the middle atmosphere (and have been shown by modeling studies to influence the dynamics of the troposphere), the solar cycle modulation of cloud nucleation through cosmic ray intensity, the impact of energetic particle precipitation on ozone chemistry, and solar cycle variations in the global electric circuit. Figure 2.16 illustrates the dramatic natural variability in the production and transport of Nitric Oxide (NO) at 110 km in quiet times and in response to particle precipitation during a magnetic storm and a solar particle event. NO acts as a thermostat, increasing atmospheric cooling and disrupting mesospheric ozone as it diffuses to low altitudes in the polar regions.

The strong impact of radiative forcing on the atmospheric structure in the upper atmosphere is the basis for its value as an excellent indicator of atmospheric change. Enhancements in anthropogenic trace gases such as carbon dioxide or methane (which increase the radiative forcing) will produce an unambiguous signature – not confused by the complexity of cloud feedbacks and competing effects (due to aerosols) that occur in the troposphere. Possible signs of global climate change include noctilucent clouds formation. These clouds are a high-altitude polar region phenomena and are occurring more frequently with sightings now at lower latitudes. Since these clouds require temperatures below 140^{o} K to form, the equatorward advance of cloud sightings are part of a growing body of information that indicates the upper atmosphere has cooled over the past 20-50 years. This cooling is thought to be the result of a warming trend at lower altitudes, possibly due to anthropogenic influences. The residual circulation (meridional) cell coupling the troposphere to the lower thermosphere upwells in the summer, convects from the summer to the winter pole in the mesosphere, and downwells to the lower atmosphere in the polar night. The fuel for this cell is predominantly small-scale waves coupling from the lower atmosphere into the mesosphere. Variation in the fuel (waves) will affect the rate at which the cell exchanges composition between the regions as well as the cooling rate in the summer mesosphere.

The distinct contrast between Sun-induced and anthropogenic changes in the upper atmosphere and the implications for changes in these regions underscore the importance for separating the possible forcings. Furthermore, the distinct changes in the upper atmosphere that have occurred over the past 20-30 years indicate the need to define the current state of the upper atmosphere structure and dynamics for comparison with measurements made in the near and distant future.

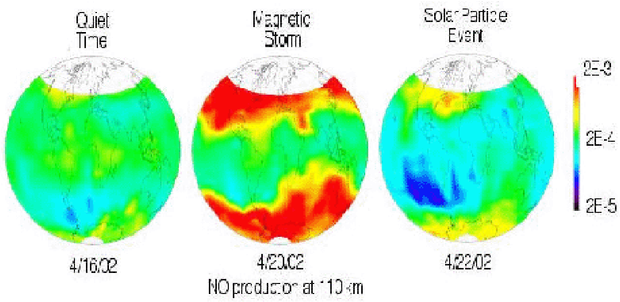


Figure 2.16 Global composite of NO production at 110 km altitude, comparing quiet time to magnetic storm and solar particle event times. The NO is produced by collisions of high-energy charged particles with the neutral atmosphere. There is a significant high latitude source region and diffusion from higher to lower latitudes during disturbed times. NO acts as a thermostat, increasing the cooling of the neutral atmosphere following a magnetic storm. Major questions remain as to how deep magnetic storm effects penetrate into the upper atmosphere. NO diffuses to low altitudes in the polar regions and impacts ozone chemistry in the mesosphere. Only when natural variability is understood can the impact of anthropogenic sources be identified.

(f) Develop the capability to predict the long-term climate of space.

Life on Earth depends on the long-term climate in space as well as the long-term stability of the Earth's atmosphere. The climate of space is the ensemble of transient heliospheric disturbances, modulation of the solar output due to the rotation of the Sun, variations in solar activity associated with the 22-year solar cycle, other, longer and less well understood cycles in solar activity, and very long-term variations in the interstellar medium. A multitude of physical processes couple this space climate to the Earth's climate. This is in addition to those processes within the Earth's atmosphere that contribute to climate variability. Understanding space climate and separating its contribution from internal contributions to the Earth's climate are key elements to determining the destiny of life in the solar system.

To forecast space climate will require an understanding of the behavior of many aspects of the Sun through time, i.e., much broader than simply the Sun's radiation output. Obviously, such understanding will not come directly from study of the Sun over relatively short periods (several solar cycles). However, there is another important way to study the evolution of the Sun and its effect on space climate. Studying other stars like the Sun will provide a context for understanding where the Sun lies in the ensemble of all possible states it might take during its lifetime on the main sequence of evolution. This study holds the best promise for determining the possibility of the Sun becoming radically different from its current state.

A second input to the long-term prediction of space climate is the role of the interstellar medium. Some historic climate variations appear to show twice the frequency of the rotation of the galaxy, which would plausibly be coupled to climate through the interstellar medium. Understanding the implications of the interstellar medium on space climate requires first an understanding of the current properties of the Local Interstellar Medium (i.e., its density, temperature, magnetic field, cosmic ray distribution, dust properties, composition, and ionization state). The next step in understanding the role of the interstellar medium on space climate is to conduct remote sensing observations to compare the properties of the medium directly upstream of the Sun's velocity vector in the galaxy with the properties further away. Such measurements provide a forecast of the impending conditions in the interstellar medium.

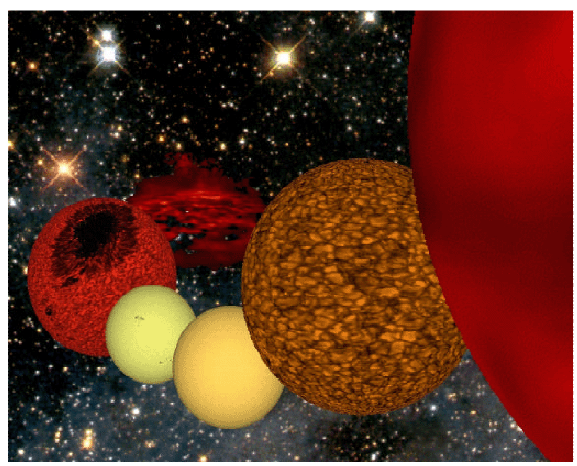


Figure 2.17 The Sun's evolution over ~10 billion years, from a protostellar disk to the magnetically active Sun of today to a red giant. Later in the Sun's life, magnetic activity will likely decrease, well before the red giant stage. This can be studied by observing appropriate solar analogs. Understanding the decay of magnetic activity is important for determining the implications on the Earth's climate.

2.4 Additional Science Objectives

The breadth of the SEC Division science extends beyond the three primary science objectives. SEC missions directly address many other Space Science Enterprise science objectives through connections to astrophysics, astrobiology, and planetary physics. The following sections describe Research Focus Areas that have direct application to other Space Science Enterprise science objectives.

2.4.1 Understand the structure of the universe, from its earliest beginnings to its ultimate fate.

(a) Develop helioseismological constraints on the structure of the Sun as a star.

Solar activity has its origin in a dynamo that generates and processes magnetic field in an interaction of rotation and convection. Much of the magnetic field involved in the dynamo is thought to be generated in, or at least stored in, the layer immediately below the convective envelope; convection pumps magnetic field into these layers as it overshoots the boundary, and that field is subsequently stretched in the rotational shear in this layer (also known as the tachocline). This layer is currently studied by helioseismic techniques from a single, ecliptic vantage point. This vantage point allows only limited depth and latitude resolution and longitudinal structure is poorly tracked. By adding a second vantage point well away from the Earth, the analysis of combined data from the two observatories allows unambiguous derivation of information along all ray paths accessible to the observatories. A mission out of the ecliptic that would reach at least mid-latitudes (or even a polar vantage point) will allow detailed exploration of the seat of the dynamo at a few optimal phases in the orbit and mapping of the flows within the convective envelope. A dedicated mission to a point some 120 degrees trailing the Earth in its orbit would allow exploration of the changes in the tachocline that are induced by the Sun's processes driving solar activity. That mission would also provide imaging of the deep interior to explore the magnitude of the core's magnetic field (likely related to neutrino flavor oscillations), the importance of gravitational settling of relatively heavy elements in the core, and the core's rotation rate.

2.4.2 Learn how galaxies, stars, and planets form, interact, and evolve.

(a) Determine the roles magnetic dynamos and angular momentum loss play in how stars and planetary systems form and evolve.

Stars form (most frequently in pairs) from cool molecular clouds that collapse under their own gravity. The final collapse into a star is hampered by even the smallest amount of rotation within these huge clouds. As the cloud contracts, material spins up more and more rapidly. Unless much of the rotational energy is lost from the center of the cloud, centrifugal forces would prohibit the ultimate formation of a star. Magnetic fields offer one efficient mechanism that may explain the dissipation of rotational energy and the associated transport of angular momentum in this phase. Better understanding is needed of the astrophysical dynamo processes that generate these fields. With this better understanding, models can be produced that provide information on what fraction of rotational energy is destroyed or removed from the cloud, and what fraction remains available for the formation of a planetary system or a stellar binary companion (or both). Knowledge about the dynamo process in stars is obtained by detailed studies of the interior dynamics and dynamo of the Sun and very-high-resolution imaging of other stars and star forming regions. This knowledge is also obtained by studying magnetized planets. Among the magnetized planets, Jupiter is the closest analog to this astrophysical application. Its magnetosphere and intense aurora are powered primarily by planetary rotation, rather than through the energy in the solar wind. Through detailed *in situ* particle and field measurements and imaging of the aurora, a link will be estabilished between magnetosphere-ionosphere coupling processes and astrophysical magnetic torquing processes.

(b) Determine the current state of the local interstellar medium and its implications for galactic evolution.

Until recently, the only direct samples of matter from beyond our solar system came from galactic cosmic rays. However, the properties of the Local Interstellar Medium (LISM) -- its density, temperature, magnetic field, cosmic ray distribution, dust properties, composition, and ionization states – can be determined through *in situ* measurements and by sampling its neutral components that penetrate into the heliosphere. These measurements hold essential clues to galactic history and would help to determine the properties of the Local Interstellar Cloud in which the heliosphere resides. For example, isotopic and elemental abundances in the LISM reflect the current state of matter in the galaxy, whereas the isotopic and elemental abundances from the material in the solar system reflect the state of matter during the solar system's formation. Over time, the interstellar medium becomes increasingly enriched in heavy elements released from stars through stellar winds, supernovae, and novae. The rate of enrichment is a direct measure of star formation and the nucleosynthetic processes occurring in the galaxy. Measurements of this medium follow a logical progression as technology develops. Precursory missions take advantage of the penetration of the interstellar medium into the heliosphere (neutral interstellar atoms penetrate to within ~4 AU, become ionized, picked up by the solar wind, and are observed as so-called pickup ions). High-resolution measurements of these pick up ions and the neutral atoms themselves are needed to determine the composition of the LISM. These precursor missions set the stage for properly instrumented spacecraft to break through heliospheric boundaries and explore the medium that pervades the galaxy.

(c) Determine the interaction between the interstellar medium and the astrospheres of the Sun and other Stars

The interaction between the Sun (or other stars) and the interstellar medium results in well-defined and important boundaries --a bow shock, heliopause and termination shock that bounds the heliosphere (the Sun's astrosphere). Astrospheres are in fact common in the galaxy. However the detailed properties of the Sun's astrosphere are not well understood. For example, the location of the termination shock has been revised considerably in recent years as the search for it continues with the Voyager mission. Measurements of anomalous cosmic rays (and galactic cosmic rays), low frequency radio emissions, high-resolution EUV imaging, and energetic neutral atom imaging are needed to remotely probe heliospheric boundaries. These measurements provide information on the properties of the LISM and the changing interaction of the LISM with the heliosphere. Remote sensing of the heliospheric boundaries provide the precursory preparation for proper instrumentation of a true interstellar explorer spacecraft.

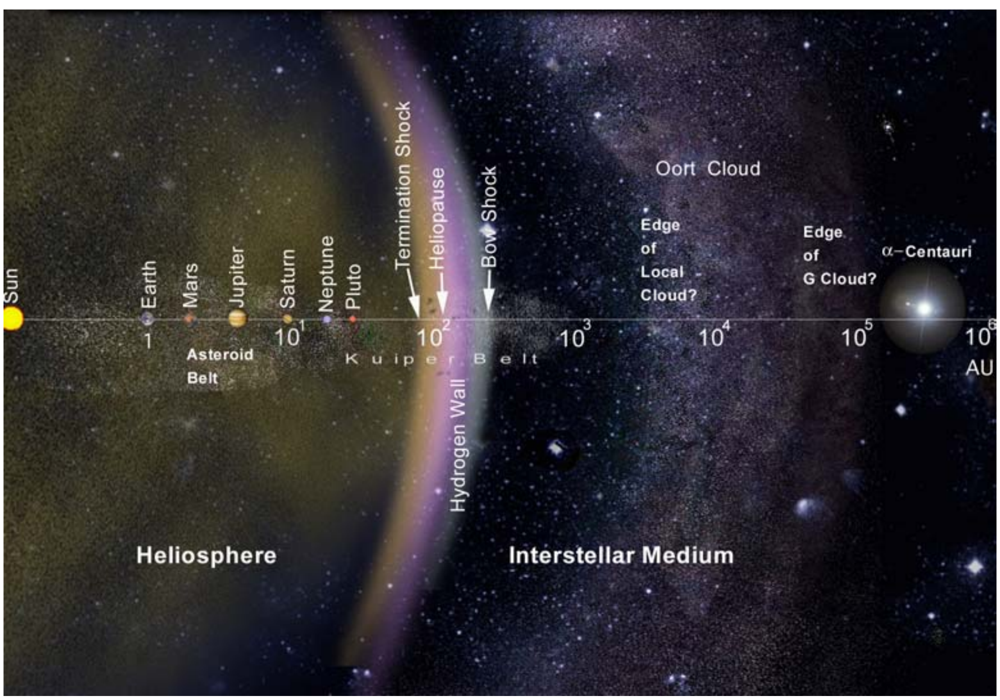


Figure 2.18 Outer boundaries of the heliosphere and LISM on a logarithmic scale. The exact locations of these boundaries are poorly known because conditions in the local interstellar medium are poorly known.

2.4.3 Understand the formation and evolution of the solar system and Earth within it.

a) Explore the role of planetary magnetic shielding in establishing diverse atmospheres of Earth, Venus, and Mars.

Recent observations show that the Earth's atmosphere is not a passive absorber of energized magnetospheric plasma that precipitates into the upper atmosphere during solar storms. In response to power input into the upper atmosphere that can measure in the hundreds of gigawatts during a typical magnetic storm, the Earth loses approximately 100 tons of oxygen from its upper atmosphere at high latitudes. While this oxygen loss is insignificant even on geologic timescales, its mechanism is nevertheless the only effective oxygen loss mechanism at Earth. Conditions at Mars and Venus are significantly different than those at Earth because of their lack of a strong planetary magnetic field to stand off the solar wind. At these planets, there is direct entry of the interplanetary magnetic field and solar wind into the atmosphere. Understanding the influence of the space environment on planetary atmospheres is important for understanding the formation and evolution of atmospheres that support life. Conditions in the Earth's near space environment may have been significantly different during the early stages of the formation of the Earth's atmosphere than they are now. Some of these differences may be reflected in current planetary magnetospheres (such as Venus and Mars). Furthermore, there are periods in the past when the Earth's magnetic field orientation has switched. During the reversal, there may be a period when the magnetic field is nearly zero, producing a solar wind interaction that is more Venus-like. Understanding this type of interaction by investigating the atmosphere-solar wind interactions at Venus and Mars helps understand the influence the space environment has on the formation and evolution of the Earth's atmosphere.

2.4.4 Probe the origin and evolution of life on Earth and determine if life exists elsewhere in our solar system.

(a) Explain the role of varying solar activity in the past, when life developed, and the controlling factors for habitability, climate, and life since then.

When planets form around a young central star, they are subjected to intense X-ray and ultraviolet radiation and are embedded in strong stellar winds and magnetic storms. The X-ray brightness is up to 1000 times higher, and the ultraviolet brightness ten to 100 times higher, than that of the present Sun. Spots in young, active stars cover as much as half a hemisphere (Figure 2.17), resulting in strong modulations of the irradiance of the planets. Studies of other stars, stellar evolution, and dynamo mechanisms are required to understand both the pathways by which these variations affect the origin of life and the continued habitability of the planets and how the solar activity varied and waned over the 5 billion years since the planetary system formed.

(b) Search for molecules and the building blocks of life from comets, Kuiper Belt objects, and dust in the heliosphere.

One of the foremost questions concerning the existence of humankind and the possibility of life elsewhere is the origin of life on Earth. Did the building blocks for life originate on Earth, or were they transported to the planet from another source? Or a combination of both? There is a diverse spectrum of bodies in the heliosphere, from minute grains on the smallest size scales to planets on the largest scales. The chemical, isotopic, and molecular composition of these bodies carry clues about their formation and evolution, clues to the synergistic relationships between these bodies, and finally clues as to whether these bodies carry the very building blocks of life. Better sampling techniques of the bodies within the heliosphere on all size scales must be devised and executed to solve these riddles. Future studies will include direct rendezvous with comets and asteroids by deep space probes; direct and indirect determination of the composition of objects and grains in the Kuiper belt; direct and indirect determination of the composition of interplanetary and interstellar grains; and, finally, the determination of the composition in all of these objects through measurement of pickup particles, created when material is evaporated or sputtered, subsequently ionized, and finally carried out by the solar wind plasma.

(c) Understand the effects of energetic particles on the evolution and the persistence of life.

The ubiquitous presence of very energetic charged particles throughout space has consequences for life. For life on Earth, the heliosphere, magnetosphere and atmosphere provide three different and important shields from the harmful effects of these naturally occurring radiations. Also, one of the major hazards for manned space travel is recognized to be the flux of energetic charged particles and cosmic rays. Intense solar flares produce extremely high transient fluxes (over a period of hours to days) of dangerous radiation which can adversely affect life in the short term, and the always present, higher-energy cosmic rays from the galaxy provide a significant long-time harmful background which is dangerous on longer-duration manned missions. The existence and evolution of life on Earth depends both on the shielding by the heliosphere and magnetosphere and on the constantly occurring mutations which are thought to be significantly influenced by the same radiations. Certainly, a significant increase in the cosmic radiation caused by a dramatic change in the shielding effect of the solar wind and heliosphere could adversely affect life (variations in the local interstellar medium for example could cause major changes in the Sun's astrosphere). A significant change in cosmic rays in the environment could change the rates of evolution, and affect the viability of life. In addition, current speculations on panspermia must deal with the reality that radiation in space can make the transfer of life very difficult. These facts illustrate the importance of quantitative understanding and prediction of the energetic particle environment. This understanding is needed through the entire chain of factors that influence these particles.

2.4.5 Chart our destiny in the solar system.

(a) Explain the role of varying solar activity in the future of terrestrial climate and habitability.

A multitude of physical processes couple the Sun's variability to the Earth's climate. This is in addition to those processes within the Earth's atmosphere that contribute to climate variability. Which processes, if any, are important in changing Earth's climate, and how they magnify or weaken each other's actions in this complex climate system have not been determined. Key to understanding climate variability is the measurement of those solar properties which are either known to affect climate or most likely do; these include the solar spectral irradiance, and the properties of the heliospheric magnetic field and disturbances such as CMEs traveling within it. Also critical is the ability to forecast solar activity on long time scales, from decades to centuries. Theoretical models and efficient knowledge sharing with studies of activity on Sun-like stars of a range of activity levels are essential for the development and validation of long-term forecasting of solar activity. Finally, understanding the role of varying solar activity in the future of terrestrial climate requires a detailed understanding of the coupling between solar activity and the Earth’s magnetosphere. This coupling modifies the effects of solar activity on the Earth’s atmosphere.