<table>
<thead>
<tr>
<th>Manuscript Number:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Title:</td>
<td>Understanding Space Weather I: The Sun as a Variable Star</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Article</td>
</tr>
</tbody>
</table>
| Corresponding Author: | Keith Temple Strong, Ph.D.  
|                    | University of Maryland  
|                    | Bowie, Maryland UNITED STATES |
| Corresponding Author's Institution: | University of Maryland |
| First Author:      | Keith Temple Strong, Ph.D. |
| All Authors:       | Julia Lee Robinson Saba, Ph.D.  
|                    | Therese Kucera, Ph.D. |
| Abstract:          | The AMS has recently adopted Space Weather as a new core competency. This is the first in a series of papers discussing the multidisciplinary aspects of space weather. This paper concerns the physics behind solar variability, the driver of space weather. We follow the tortuous journey of the energy from is production in the solar core until it escapes into interplanetary space, showing how the internal dynamics and structure of the Sun change its nature. We show how the production and dissipation of magnetic fields are a key clue to untangling the riddle of the sunspot cycle and how that, in turn, affects the amount of radiation that the Earth receives from the Sun, the total solar irradiance. |
| Suggested Reviewers: | John Leibacher, phd  
|                    | GONG director, National Solar Observatory  
|                    | jleibacher@nso.edu  
|                    | Knowledge of solar interior |
|                    | Holly Gilbert, phd  
|                    | Lab Chief, NASA Goddard Space Flight Center  
|                    | holly.gilbert@nasa.gov  
|                    | Head of solar physics lab and EPO expert |
Understanding Space Weather: The Sun as a Variable

Star

Keith Strong\textsuperscript{1,2}, Julia Saba\textsuperscript{1,2}, and Therese Kucera\textsuperscript{3}

\textsuperscript{1}SP Systems at NASA GSFC
\textsuperscript{2}University of Maryland, College Park, at NASA GSFC
\textsuperscript{3}NASA Goddard Space Flight Center

\textbf{Corresponding Author:}

Dr. Keith T. Strong

12310 Longwater Drive

Bowie, MD 20721

Tel. 301-249-0638

Email: drkts@aol.com and keith.strong@nasa.gov
Introduction

We live on a planet that is immersed in the outer atmosphere of a variable star. Fortunately for us, our star, the Sun (Fig. 1), is relatively benign compared to some of its overly energetic stellar cousins; otherwise we would likely not be here to appreciate it. The Solar System has a relatively narrow band of habitability for complex life forms that is defined by the distance from the Sun where water remains liquid most of the time. If the Sun were to become significantly more active, that zone would move outward, leaving the Earth inside the hot inner boundary, and our oceans would eventually boil away. If the Sun were to dim appreciably, the Earth would turn into a ball of ice. A quiescent star and relatively stable planetary orbit are therefore prerequisites for the development of complex life forms and the eventual establishment of civilizations.

Space weather researches the effects of solar variability throughout the Solar System. Its ultimate goal is to be able to predict solar variability and how it will affect life and society. Our civilization is increasingly reliant on space-based technologies and so is becoming more vulnerable to space weather effects. Thus space weather is important not only for its scientific value but also because it can help mitigate damage or loss of susceptible technological systems.

This is the first in a series of papers that describe space weather, its impacts, and our fledgling abilities to predict it. In this paper we describe the wide range of phenomena that create long-term solar variability. We follow the energy production in the unimaginably hot and dense fusion reactor at the Sun’s core to the point where it finally starts its long journey into interplanetary space and beyond. In subsequent papers, we describe short-term variability such as flares and coronal mass ejections, the arduous
journey of the solar wind throughout the heliosphere, its interaction with the planets (especially the Earth), and the specific threats that space weather poses to our technology based society. Finally we review of the status of current solar cycle and assess our ability to predict its evolution. We start this journey in the solar core where the Sun produces most of its energy.

The Solar Interior

The Sun is 90% hydrogen; most of the rest is helium, with a trace of heavier elements, originating primarily in the gaseous nebula from which the Sun was formed. The Sun generates vast amounts of energy by fusing protons into helium ions, and in doing so converts some of its mass to energy. Every second the Sun destroys about $4 \times 10^6$ tonnes of mass in producing nearly $4 \times 10^{26}$ J. That is equivalent to about a million times the world’s annual energy production. The transport of this energy outward from the solar core, combined with the internal dynamics of the Sun, creates variations in solar activity that ultimately drive the ambient conditions in interplanetary space and, most importantly to us, in geospace. The different zones in the solar interior play unique roles in the energy transport process, combining to produce many different types of space weather phenomena on a variety of timescales.

The solar interior can be divided into three layers – the core, the radiation zone, and the convection zone – each characterized by the dominant physical processes that take place there (Fig. 2). In general, there are three main mechanisms for transporting solar energy: conduction, radiation, and convection. Conduction is not a very efficient process in the dense solar interior but plays a significant role in the tenuous solar
atmosphere. The conditions favorable for convection occur only nearer the solar surface. Deep inside the Sun, radiation is the dominant transport mechanism.

**The Solar Core:** The core is the zone where nuclear fusion occurs; about half the solar mass and 98% of the energy generation occur within about the first 25% of the solar radius ($R_s$). The peak energy production rate occurs at the center of the Sun, where the density and temperature are highest. Densities can reach 150 T/m$^3$ with temperatures of up to 15 MK. At increasing distance from the center of the Sun, the temperature and density decrease to match the gravity gradient, so the energy production rate drops (Stromgrew, 1965).

The Sun has been producing energy like this for over 4 billion years. How is this possible? The answer lies in the huge size of the fusion region within the Sun and the nature of the fusion reaction. The balance between gravitational and pressure forces contains the nuclear furnace in hydrostatic equilibrium. The rate of a fusion reaction is proportional to both density and temperature. An increase in the rate of fusion will cause the temperature to rise and the core to expand. An expansion of the core reduces its density and damps the fusion reaction rate. Thus the Sun is effectively a naturally occurring, self-regulating fusion reactor. Although the Sun generates only a few watts in a cubic meter in the core, the generator is vast, over 30,000 times the volume of the Earth.

The energetic by-products of the fusion process are neutrinos and electromagnetic radiation ($\gamma$-ray photons). The neutrinos have such a low probability of interacting with anything that they pass through the body of the Sun at nearly the speed of light, carrying away a tiny fraction of excess energy produced by fusion, and play no further role in the
space weather story. Thus it is left to the photons to carry the bulk of the energy away from the core region. The γ-ray photons created in this very high density environment interact repeatedly with the particles of the solar plasma. The photons are scattered, absorbed, and reemitted as altered photons in random directions, but generally able to travel a little farther in the direction of decreasing density before their next interactions, so that the energy is gradually, preferentially, transported outward. As the photons, on average, transfer energy to the ambient plasma, they drop in energy and so increase in wavelength, until, as visible light, they reach the solar surface, the photosphere, from whence the energy escapes into interplanetary space.

How long does this tortuous transport of energy from the core to the surface take? An astonishing 30 million years (Stix, 2003)! That is, the sunlight we see today was generated during the early Oligocene, a period when the continents were nearing their present positions.

**The Radiation Zone:** Immediately above the core, out to a distance of about 0.7 Rs, the dominant mode of energy transport remains radiation. Here the pressure of the solar plasma drops steadily, maintaining hydrostatic equilibrium with the gravitational force throughout the body of the Sun. That requires both the temperature and density of the solar plasma to fall. At the base of the radiation zone, temperatures are about 10 MK, but they decrease to 2 MK at its upper boundary.

**The Convection Zone:** At the top of the radiation zone, the temperature is low enough that some of the heavier trace elements are no longer fully ionized. Since multi-electron ions can more readily absorb photons, the opacity of the plasma increases, causing the temperature to rise and the plasma to expand. Thus it becomes buoyant and rises
relatively quickly to the photosphere. In this upper layer, the optical depth of the plasma drops as the density falls to the point where the visible-light photons can escape unhindered into space. This cools the plasma to about 5800 K, so that it becomes denser and slowly sinks back down, only to be reheated. This sets up a convective pattern, and so convection becomes the dominant energy transport mechanism in the outer layer of the Sun (0.7–1.0 R$_S$).

The smallest-scale convection cells, approximately 1000 km in diameter, are called granules, and can be seen clearly as a pattern covering the entire surface of the Sun (Fig. 3). An individual granule may last only a few minutes. There are indications of larger-scale patterns in the photospheric features, suggesting that there may be a hierarchy of convection cell scales including mesogranules (10,000 km), supergranules (30,000 km), and giant cells (100,000 km) that lie below the solar surface (Hathaway et al., 2000).

**The Rotating Sun**

One theory of the Sun’s formation postulates that, as the huge gaseous nebula that was destined to become the Sun started to collapse, the inner edge had a different velocity from the outer edge because their orbits around the galaxy’s center were at different distances. These velocity differences between the extremities of the gas cloud imparted spin to the steadily collapsing nebula, and, rather like an ice skater pulling in her arms when pirouetting, the gradually forming solar nebula started to spin faster and faster. Most of the angular momentum was eventually shed into the planets and other Solar System debris.
The Sun is not a solid body, and so it would not be expected to rotate rigidly. It rotates differentially: at the surface, the polar regions rotate more slowly than the equator, causing a shearing motion between solar plasma at different latitudes. At the equator, the solar synodic rotation period (i.e., the rotation period with respect to the Earth) is about 26 days, but at the poles it is nearly 36 days.

But what is the rotation rate deep inside the Sun? To look inside the Sun to determine the internal dynamics, we use the tools of helioseismology, akin to Earth-based seismology in the use of sound waves to probe beneath the surface. The buffeting of the convection zone sends out a continuum of sound-wave frequencies, some of which resonate with the structure of the Sun and produce global standing waves. Since pressure is the restoring force for these waves, they are called p-waves. There are over a million different p-wave modes, each having a frequency of about 3.33 MHz, which is a natural resonant frequency for the convection zone cavity. This frequency corresponds to a period of about 5 minutes. The resonant waves are reflected at the photosphere and refracted deep inside the Sun. The longer the wavelength of the waves, the deeper they penetrate into the solar interior.

Although the amplitudes at the solar surface can be tens or hundreds of meters, the p-waves are most readily detected via minute Doppler shifts (a few cm/s) in some of the Fraunhofer absorption lines in the visible solar spectrum, which occur in cooler plasma overlying the emission region; instruments like the Michelson Doppler Imager (Scherrer et al., 1995) on the Solar and Heliospheric Observatory (Domingo et al., 1995) or ground-based systems like the Global Oscillation Network Group (Harvey, 1996) are used to make these Doppler measurements. Thus, the signatures of the p-waves visible on
the surface of the Sun can be used to image the density, temperature, and dynamics of the solar interior. Global helioseismology techniques can be applied to the standing acoustic waves to obtain the large-scale structure and dynamics of the solar interior, to the base of the convection zone; local helioseismology techniques can be used to probe the detailed 3D structure of the subsurface structure and flows and to image large magnetic structures on the far side of the Sun (see, e.g., Hill, 1988; Duvall et al., 1993; Lindsey and Braun, 1999).

Global helioseismic studies of the solar interior have found that the solar rotation is differential in radius as well as in latitude (Fig. 4). The radiation zone and core seem to rotate as a solid body at a frequency of about 430 nHz (27 days), but the deeper the interior is probed, the larger the uncertainties become, so little is currently known of the dynamics of the core region. Differential rotation is maintained throughout the convection zone.

**The Magnetic Sun**

Between the rapidly but uniformly rotating radiation zone and the differentially rotating convection zone, there is an interface layer called the tachocline (Spiegel and Zahn, 1992) where the maximum radial and latitudinal shearing occurs, generating large magnetic fields. It is generally thought that the tachocline dynamo is the source of the solar activity cycle. The magnetic fields generated by this dynamo are caught up in the differential rotation process and start to be stretched and twisted. The poloidal magnetic field (i.e., the north-south dipolar field) is gradually turned into a toroidal field (i.e., an east-west field) through this rotational process (Fig. 5). This effectively increases the strength of the field by converting rotational energy into magnetic energy. Over a period
of years, the magnetic pressure near the tachocline builds up in twisted magnetic “ropes” until it is greater than the local plasma pressure, and the magnetic flux ropes expand and become buoyant (Fan, 2008) and rise through the convection zone until they break through the visible surface of the Sun, the photosphere, to be viewed as sunspots (Fig. 6).

Why does a sunspot appear dark? The presence of strong magnetic fields (up to 5000 G, about 25,000 times stronger than the Earth’s magnetic field) inhibits the convective-energy flow from below. A lower energy flux means lower temperatures (~4800 K or less), so the areas where these magnetic fields poke through the solar surface appear as dark patches on the photosphere.

Sunspots are not the only magnetic feature seen at the surface of the Sun. The large-scale weak dipolar solar field is dominant at the poles. There is also a network of small and relatively weak magnetic dipoles dotted all over the Sun – the magnetic carpet (Title and Schrijver, 1998). Long, sinuous magnetic neutral lines snake their way across the photosphere, supporting huge relatively cool structures high in the extremely hot ambient solar atmosphere. These structures are known as filaments; they are also called prominences when their height structure is visible above the solar limb.

Strong magnetic fields extend high above the photosphere into the tenuous outer layer of the solar atmosphere, viz., the corona (Fig. 7). The hot coronal plasma is highly ionized and so is trapped by the magnetic fields, thus outlining the fields in X-ray and EUV light so that their evolution and interactions can be traced in detail. This reveals a complicated structure of intertwined and dynamic magnetic fields that change on every observed temporal and spatial scale in response to the dynamic motions in the photosphere, the emergence and dissipation of magnetic structures, and drift caused by
differential rotation. This dauntingly complex magnetic picture of the Sun nevertheless shows a kind of order as the strong magnetic fields change continuously in a quasi-regular way, yielding the solar activity cycle.

**The Solar Cycle**

While there are many ways of measuring the level of solar activity, the most common method is sunspot number (SSN), an approach first used by Rudolf Wolf over 150 years ago (Izenman, 1985). SSN combines the number of sunspot regions with the total number of individual spots in a weighted average that loosely describes the long-term variation of solar activity but does not correlate at all well in the short term with other measures of solar activity such as X-ray flux, magnetic field, or flare rates.

When SSN is plotted as a function of time, a clear cyclic behavior becomes apparent (Fig. 8) as Samuel Schwabe discovered in 1845. The sunspot cycle takes about 11 years, although there have been cycles as short as 8 years and as long as 14 years. The peak amplitudes of the cycle are also highly variable. The reasons for such changes in the cycle length and magnitude are unknown, although there are proximate “symptoms,” such as observed changes in the flow speeds of material in the meridional circulation, from the equator to the poles at the surface, with a presumed return flow in the interior (see, e.g., Zhao and Kosovichev, 2004). It is expected that ongoing helioseismology studies, using higher-resolution data from the Helioseismic and Magnetic Imager (HMI) (Schou et al., 2011) on the Solar Dynamics Observatory (SDO) (Pesnell et al., 2011), will offer new insights for producing more physics-based modeling of the solar cycle.

At solar minimum there are often no sunspots, but only a weak, large-scale dipolar field that stretches from pole to pole and a network of random small magnetic
bipoles. A few old-cycle regions sporadically emerge near the solar equator. They are generally small and short lived. Such conditions can persist for several years, as they did recently in the surprisingly long transition from solar cycle 23 to cycle 24 (Cranmer et al., 2010, and references therein).

The first harbingers of the new cycle appear as small, weak sunspot groups at middle latitudes (30 to 40 degrees) during solar minimum. As the new cycle asserts itself, the old cycle decays until, at the onset of the new cycle, there is a sudden global outbreak of new-cycle activity with larger, stronger magnetic regions appearing at a range of longitudes (Strong and Saba, 2009). Over the next few years, the activity builds to a maximum, often remaining at an elevated level for several years, and then declines to solar minimum conditions once again over the next 4 to 5 years. As the activity level declines, the latitude of sunspots slowly sinks toward the equator in both hemispheres (Fig. 9). Thus the overall toroidal field is slowly destroyed by magnetic reconnection, leaving a new large-scale poloidal field in place but with its polarity the reverse of the previous one.

The full solar magnetic cycle is actually about 22 years because of the reversed magnetic polarities in two successive activity cycles. In a given activity cycle, all the magnetic regions have the same polarity in one hemisphere and the opposite polarity in the other. However, the polarities of the magnetic regions that emerge from one cycle to the next are reversed in each hemisphere. The polarity of the overall dipole field also reverses, around the time of solar maximum.

Throughout the cycle, in comparing one cycle to the next, the only reliable way to determine the level of solar activity is to monitor the magnetic field; however, we have
only just over 30 years’ worth of such data available, and broad conclusions are hard to defend when based on such a short time line. Further, the three recent cycles have been very similar in size and duration. Cycle 24 promises to be different from its three predecessors, perhaps being more like earlier, less active and longer cycles, and so it may offer hope of extending the range of magnetic observations into a new domain as it progresses.

There were periods during the 17th and early 19th centuries when solar magnetic activity declined to exceptionally low levels – the so-called Maunder and Dalton minima (Eddy, 1977; Usokin, 2008). Such extended minima have been regularly observed on other Sun-like stars (Baliunis et al., 1995). However, it is not yet understood how the Sun gets into or out of a minimum-activity state, given our current understanding of how the solar dynamo works.

**The Solar Energy Output**

The important question about solar activity as far as space weather and terrestrial climate are concerned is whether it causes significant changes in the amount or the nature of the radiation and particles striking the Earth. To answer the first part of that question, we must look at the total solar energy output.

The Sun’s output is not constant in time, but the total output (the total solar irradiance – TSI) was once thought to be invariant. With the clarity of hindsight, this is somewhat surprising, as sunspots come and go cyclically and their darkness corresponds to a decrease in the solar energy output. Hence, one might have expected the solar output to decrease at solar maximum when many sunspots are present. However, sunspot groups
are often surrounded by bright patches called faculae, which tend to compensate for the loss of emissions from the dark spots. Which factor dominates, or are they in balance?

In the late 1970s and early 1980s, instruments on two NASA missions (Nimbus 7 and Solar Maximum Mission) began to monitor the solar output with vastly improved precision; their measurements led to the discovery that TSI was positively correlated with the sunspot cycle. Thus, the faculae predominate in determining the variation of TSI throughout the cycle. The long-term variation of the TSI is punctuated by short-lived deep dips (~0.3%) when particularly large sunspot groups transit the solar disk. The problem with looking for longer trends in TSI is that the observations have been taken by several different spacecraft and the task of making a composite to establish a longer baseline is difficult; there have been a number of attempts to do so, but there are wide discrepancies between them, and thus we have no consensus (Frohlich and Lean, 2004; Wilson and Mordinov, 2003).

We now have just over 30 years’ worth of TSI observations, but this represents only three activity cycles, and it is dangerous to put too much weight on any broad conclusions drawn from them. There are a variety of longer-term proxy data for solar activity, but many of the proxy estimates are based largely on unreliable SSN measurements and so may be adding as much noise as signal to our understanding of solar activity patterns.

The change in TSI over a solar cycle is <0.1%. However, the emission in various wavebands can vary much more substantially. For example, UV irradiance can vary by 10-40% over a cycle. Background X-ray emission changes by factors of 100 or more from solar minimum to solar maximum. These short-wavelength emissions are important;
for example they affect the scale height of our atmosphere which changes the atmospheric drag on low Earth orbiting (LEO) satellites.

**Summary**

The Sun is a complex system of systems and until recently, less than half of its surface was observable at any given time and then only from afar. New observational techniques and modeling capabilities are giving us a fresh perspective of the solar interior and how our Sun works as a variable star. This revolution in solar observations and modeling provides us with the exciting prospect of being able to use a vastly increased stream of solar data taken simultaneously from several different vantage points to produce more reliable and prompt space weather forecasts.

Solar variations that cause identifiable space weather effects do not happen only on solar-cycle timescales from decades to centuries; there are also many shorter-term events that have their own unique space weather effects and a different set of challenges to understand and predict, such as flares, coronal mass ejections, and solar wind variations [see in this series Paper II (Strong, Schmelz, and Saba, 2012); and Paper III (Strong, Saba, and Viall, 2012)].

**Acknowledgments**

We have benefited from constructive discussions with Jack Harvey (NSO), Joan Schmelz (University of Memphis), and Karen Fox (ADNET Systems Inc., at NASA/GSFC). This work was made possible by support from Code 670 at NASA Goddard Space Flight Center via SP Systems Inc. and University of Maryland, College Park.
References


Lindsey, C., and Braun, D.C., 1999; “Seismic images of the far side of the Sun,” *Science*, **287**(5459), pp.1799-1801


Shou, J., et al., 2011; “Design and ground calibration of the Helioseismic and Magnetic Imager (HMI) instrument on the Solar Dynamics Observatory (SDO)”, *Solar Physics* (Submitted)


Strong, K.T., Schmelz, J. T., and Saba, J.L.R., 2012; “Understanding Space Weather II: The Violent Sun”, *BAMS*, submitted


**Figure Captions**

Fig. 1. The solar atmosphere viewed in EUV light shows a wide range of magnetic structures. This composite false-color image, combining data captured by the Solar Dynamics Observatory in different filters of the Atmospheric Imaging Assembly, shows active regions (white), coronal holes (blue), and a million-kilometer-long cold filament suspended in 1MK plasma (red).

Fig. 2. Anatomy of the Sun: The Core is dominated by high temperatures and densities that enable the nuclear fusion processes that power the Sun. The location here the temperature and density fall below the critical values needed to sustain nuclear fusion (dashed line) defines the lower boundary of the zone where energy is transported mainly by radiation. The radiative zone rotates as a solid body. At about 0.7 RS the opacity of the plasma increases, causing it to warm and become more buoyant and creating a zone where convection is the predominant transport mechanism. Between the Convective and Radiative Zones lies a relatively thin region called the Tachocline where huge shear forces create a magnetic dynamo. The top of the convective zone is defined where the opacity is low enough for photons to escape into space – the Photosphere. Just above the photosphere, the solar atmosphere reaches a temperature minimum (4700 K) in a region called the Chromosphere. The density drops off rapidly above the chromosphere, and the temperature rises due to non-thermal heating processes, creating the Transition Region. Eventually temperatures approach 1 MK, creating the exquisitely beautiful and dynamic Corona.
Fig. 3. A network of convection cells (granules) in the solar photosphere as viewed with the Swedish Solar Telescope on La Palma in the Canary Islands. These granules are about 1 arcsec in diameter and last for up to 20 minutes.

Fig. 4. The rotation rates derived from observations show the differential rotation at the photosphere; they also show that the Sun rotates internally at varying speeds as a function of radius throughout the convection zone but as a solid body in the radiation zone. We know little about the core dynamics.

Fig. 5. North-south magnetic field lines deep inside the Sun get stretched and twisted as differential rotation carries them around the Sun over a period of years. Eventually they get so stressed that they become unstable and kink, breaking through the solar surface to appear as sunspots.

Fig. 6. Sunspots are the clearest manifestation of solar magnetic activity, heralding the changes of the Sun’s magnetic cycle. However there are also more ubiquitous magnetic variations in the quiet Sun, and short-term changes such as flares and coronal mass ejections. (Image from Big Bear Solar Observatory)

Fig. 7. Coronal magnetic loops observed in EUV by the TRACE satellite. Note the variety of loop sizes and shapes in the arcade overlying the dark (cool) filament channel that meanders across the solar surface. The larger loops imaged here are over 100,000 km high.
Fig. 8. The daily sunspot number (SSN) for the last five solar cycles averaged over a synodic rotation period (27.27 days). SSN clearly shows the approximately 11-year solar cycle and its variable amplitude. On the far right, the onset of cycle 24 is just visible, starting in December 2009. (Sunspot data from NOAA National Geophysical Data Center.)

Fig. 9. The magnetic history of the last three complete solar activity cycles (cycles 21 to 23). Note how the polarity in one hemisphere is reversed compared to the other for a given cycle, and how the polarities reverse from cycle to cycle. The drift of surface magnetic fields to lower latitudes can be seen as the cycle progresses. (Figure courtesy of D. Hathaway, NASA/MSFC.)
Tables:

None.
Figures

Figure 1:
Figure 2:
Figure 3:
Figure 5:
Figure 7:
Figure 8:
Figure 9: