30 April 1982, Volume 216, Number 4545

SCIENCE

The Sun's Influence on the Earth's Atmosphere and Interplanetary Space

J. V. Evans

The sun's energy comes from thermonuclear reactions, such as the synthesis of helium from hydrogen, which proceed with a loss of mass and liberation of energy. This accounts for the vast release of energy that must have occurred for many millennia and which might be expected to continue well into the future. There is, however, no certainty that the rate of energy release will remain constant; any variations would be of considerable scientific interest and profoundly important for man's well-being.

Summary. The bulk of the sun's radiation is in the visible and infrared. Solar radiation at these wavelengths controls the weather in the lowest levels of the earth's atmosphere. The rate at which this energy is emitted (the so-called solar constant) varies by a few tenths of 1 percent over a time scale of days. Longer period variations may exist, but have yet to be detected. Far more variable are the amounts of energy emitted as ultraviolet, extreme ultraviolet, and x-rays, and in the continuous outflow of ionized solar particles. The latter controls the properties of the space between the earth and the sun as well as those of the earth's magnetosphere. The ultraviolet and particle emissions control the properties of the earth's upper atmosphere, including the global wind circulation and changes therein associated with intense auroral storms. While considerable progress has been made in exploring the solar-terrestrial system since the advent of space research, many problems remain. These include the question of how magnetic energy is converted into ionized particle energy in the sun and in the earth's magnetosphere, the way in which solar and terrestrial magnetic fields join or merge, and how large electric fields are generated and sustained a few thousand kilometers above the earth's poles. Perhaps the most intriguing question concerns the possible relation between solar variability and the earth's weather and climate.

Solar-terrestrial research deals with the response of the earth's upper atmosphere, ionosphere, magnetosphere, and intervening space to the emission of energy from the sun and to any fluctuations therein. Excluded by tradition has been the troposphere, the lowest 10 kilometers of the earth's atmosphere, where the earth's weather occurs (*1*). The earth's ionosphere is formed at altitudes above about 70 km (*2*). Here the gas becomes partially ionized as a result of the absorption of solar ultraviolet and extreme ultraviolet (EUV) radiation, forming a region that affects the propagation of radio waves. The magnetosphere is that region of space surrounding the earth in which the earth's magnetic field dominates solar or other fields. It extends great distances in the antisunward direction but only 10 to 12 earth radii (R_E) on the sunward side.

As early as the 18th century it was noted that changes on the earth (such as magnetic field fluctuations and the appearance of visible auroras) occur that seem linked to changes in the sun (such as the appearance of dark sunspots). Only since the advent of space research has it been possible to establish the coupling mechanism and fully study the earth's upper atmosphere, the magnetosphere, and interplanetary space.

The National Academy of Sciences has recently released a report on the field of solar-terrestrial research and its development during the 1980's (*3*). The report argues (i) that understanding this vast system entails more than the study of its component parts because of the existence of complex linkages and feedback mechanisms and (ii) that there should be appropriate recognition of the unity of the subject by the scientific community, the funding agencies, and teachers at universities. In this article some of these connections are discussed and an attempt is made to sketch what is known about the solar-terrestrial system.

Solar Processes

The energy produced in the sun's core by thermonuclear processes is transported outward by radiation. It undergoes successive transitions to longer wavelengths as high-energy gamma rays are absorbed and reradiated in the form of x-rays and as these in turn are absorbed and reemitted in the form of EUV radiation. At about five-sixths of the distance to the visible surface, or photosphere, the ionization of the gas has decreased to the point where it is convectively unstable. The transport of energy to the surface through much of this outer zone is by means of turbulent convection, not radiation. The convection zone organizes the solar magnetic field that reaches the surface and is a source of mechanical energy; both phenomena have important consequences for the sun's atmosphere (*4*).

In the convection zone the temperature is low enough that some of the free electrons can be recaptured by protons or other nuclei to form atoms. These are

The author is chairman of the Committee on Solar-Terrestrial Research, National Academy of Sciences, Washington, D.C. 20418 and an assistant director of Lincoln Laboratory, Massachusetts Institute of Technology, Lexington 02173.

0036-8075/82/0430-0467$01.00/0

more opaque to radiation than the highly ionized gas below, and this leads to an increased temperature gradient. An upward-moving parcel of gas thus finds itself in a region of decreasing temperature where it is heated by energy released by further electron recombination and its upward motion is accelerated; the reverse holds for a parcel moving downward. The whole process is analogous to the convective processes in thunderstorms. The transport of energy again becomes radiative below the visible surface of the sun, but some of the convective motions penetrate to the photosphere, where they are visible as rising and falling granular structures.

In the chromosphere, a region extending only a few thousand kilometers above the surface, the temperature rises rapidly from about 5,000 to 500,000 K. There is a further increase to 1 million to 2 million K over a distance of about half a solar radius in the corona which constitutes the outermost part of the sun's atmosphere. The rise in temperature in the chromosphere and corona cannot be produced by radiative transfer of energy from the lower cooler layers; the energy is thought to be supplied by mechanical power in the motions of the convection zone. The chromosphere and corona are the source of the ultraviolet and x-ray radiation emitted by the sun. The visible light scattered by the corona can readily be photographed during an eclipse (or from a satellite above the earth's atmosphere). Near sunspot minimum the corona often exhibits considerable gross structure, with irregular equatorial extensions, and may be nearly absent above the poles. Near sunspot maximum the corona appears to be much more structured and jagged, but overall exhibits greater spherical symmetry.

In the convection zone the energy density of the gas motion greatly exceeds that of the solar magnetic field, and thus the motions tend to organize the fields. (Whenever the energy density of the ionized particles in space greatly exceeds that of the magnetic field, the field must follow the particle motions as if frozen into the medium.) The reverse is true in the chromosphere and corona, where the motion of the completely ionized gas is confined and controlled by the magnetic field. Sunspots are thought to be cool regions where the magnetic fields at the photosphere are strong enough to influence the convective motions and upward flow of heat. The 11-year cycle in their number and latitude on the disk may be a result of an observed differential rotation of the convection zone with latitude; the equatorial regions rotate faster than the poles. This differential motion is thought to twist the sun's dipole field until loops or knots of field lines break through the photosphere. Very large loops are sometimes seen on the limb of the sun. These loops represent magnetic "bottles" capable of containing high-density plasma. Solar flares are thought to result when the plasma containment collapses as a result of the merging of magnetic fields in regions where they are strong, complex, and highly stressed. The magnetic energy released is transformed into radiation through heating and acceleration of the trapped ionized gas. In the very largest flares, it is estimated that more than 10^{32} ergs can be released in a period of minutes, but this output remains small compared to the sun's overall luminosity ($\sim 4 \times 10^{33}$ ergs per second).

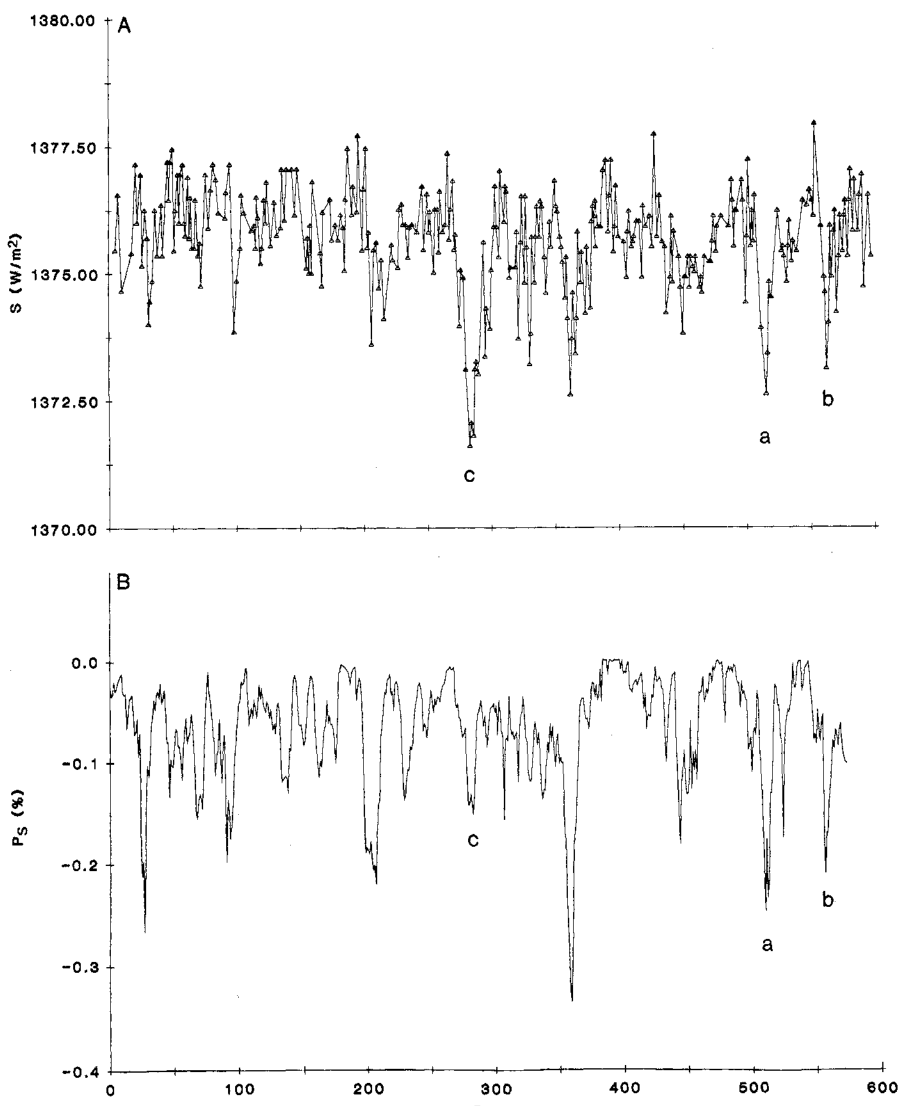


Fig. 1. (A) Nimbus 7 measurements of total solar irradiance from 15 November 1978 through June 1980 (*13*). The dips a and b were detected in observations made with the Solar Maximum Mission satellite (*12*). Point c, the deepest depression seen to date, occurred in August 1979. (B) Predicted variation (P_s) when allowance is made for the fraction of the sun's disk obscured by sunspots. [By permission of P. Foukal (*14*)]

The Solar Constant

The solar constant is the total electromagnetic energy radiated by the sun at all wavelengths per unit time through a given area at the mean distance of the earth; the contemporary value is S = 1.37 ± 0.02 kilowatts per square meter (*5*). The term solar constant is somewhat misleading since the total energy fluctuates by a few tenths of 1 percent. About 99 percent of the sun's energy is emitted at wavelengths between 0.3 and 10 micrometers from the photosphere (which can be characterized as a blackbody radiator with a temperature of roughly 5800 K).

Stellar structure theory suggests that a G-type star such as the sun can be expected to exhibit a change in luminosity of approximately 30 percent during its main-sequence lifetime (~ 10 billion years) (*6*). Evidence that life has existed on the earth without interruption for 2.5 billion years implies that much of the planet has had temperatures between 0° and 100°C throughout this period and places constraints on the rate and extent of secular changes in the solar constant (*7*). There is, however, evidence for climatic changes, such as glaciations, that were striking departures from the norm. Since the cause of these is not well understood, small fluctuations in solar output cannot be ruled out solely on the basis of the climatic record. Computer model studies suggest that fluctuations in the solar constant exceeding a few tenths of 1 percent for a long period would have significant climatic effects (*8*).

Ground-based measurements of the solar constant are limited by the highly varying and uncertain effects of the atmosphere, which influences the transmission of visible light and blocks out most of the energy below 0.3 μm and above 2.5 μm. Despite these difficulties, C. G. Abbot and co-workers at the Smithsonian Astrophysical Observatory made daily measurements of the solar constant from several stations between 1923 and 1952 and obtained a nearly continuous set of data (*9*). Analysis of the data secured at the two best stations rules out any sustained variation of the flux reaching the ground (about 80 percent of the total) exceeding 0.17 percent over 30 years (*10*). However, the removal of slow trends that may be of instrumental origin reveals a significant fluctuation (~ 0.1 percent) with a 28-day period (*11*). The solar constant tends to be low when parts of the disk are covered with dark sunspots and high when the disk is covered by bright features (faculae), which generally are also associated with a locally strong magnetic field.

The latter finding has been confirmed by recent spacecraft observations of the sun. High-precision measurements of total solar irradiance made with the Solar Maximum Mission satellite for 150 days showed two decreases of > 0.1 percent below the mean, each lasting about 1 week (*12*). These same two decreases were seen independently by the Nimbus 7 satellite (*13*) and are shown as points a and b in Fig. 1A. The Nimbus 7 data appear noisier than those reported by Willson *et al.* (*12*) because of limitations in the data sampling system, but extend earlier in time and show a decrease exceeding 0.3 percent (point c in Fig. 1A). Figure 1B shows the variation that would be expected at the earth if the radiation blocked by sunspots is not emitted simultaneously elsewhere from the disk and if allowance is made for the position of the spot relative to the central meridian (*14*). This function is capable of predicting all the major dips, but not their relative amplitudes.

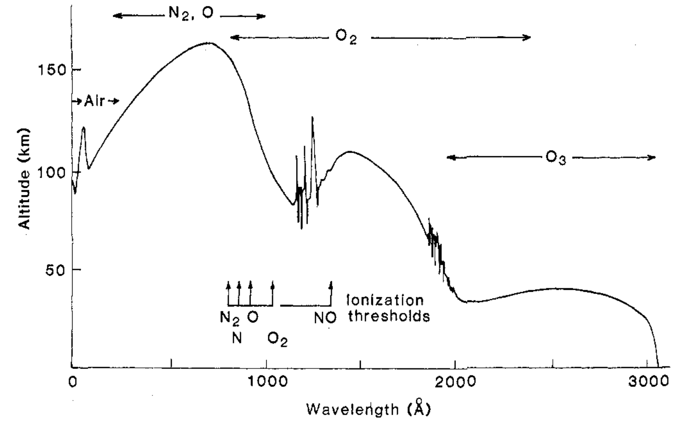


Fig. 2. Altitude at which the intensity of solar radiation drops to 1/*e* of its value outside the earth's atmosphere for vertical incidence. [Adapted from L. Herzberg, in (*53*)]

It seems extremely unlikely that there is any variation in the rate of energy production from nuclear fusion in the sun's core over these time scales. Thus, what is being seen must be a temporary reduction in the rate at which heat is convected through the outer layers of the sun. The "missing" energy presumably is stored as increased thermal and potential energy and released slowly after a spot disappears (*14*).

Ultraviolet and X-ray Emissions

Approximately 1 percent of the total solar irradiance lies in the ultraviolet at wavelengths between 3000 and 1200 angstroms. Over much of this wavelength interval the spectrum appears to be that of a continuum with superimposed absorption lines. This energy does not reach the earth's surface but is absorbed by ozone between 20 and 50 km and by molecular oxygen at still higher altitudes. Ultraviolet at wavelengths shorter than 2424 Å can photodissociate O_2 into its constituent oxygen atoms; these in turn oxidize other O_2 molecules to produce ozone. Thus the energy in this wavelength region is important for controlling the oxygen chemistry in the earth's atmosphere.

Figure 2 shows, as a function of wavelength, the level at which the ultraviolet penetrating the earth's atmosphere from an overhead sun has been attenuated by a factor of 1/*e*. (This is also the level at which the absorption per unit height interval is greatest.) At wavelengths shorter than about 1500 Å, emission lines of gas in the chromosphere and corona dominate the spectrum. The EUV radiation at wavelengths between 300 and 1200 Å is absorbed by O_2, O, N_2, and N in the earth's atmosphere, chiefly above 90 km (Fig. 2). Radiation at wavelengths below 1027, 911, and 796 Å can photoionize O_2, O, and N_2, respectively. Thus, EUV radiation and x-rays govern the production of free electrons in the earth's ionosphere. In all, only about 1 to 3 milliwatts per square meter, or one-millionth of the total solar flux, is absorbed above 120 km, but because the atmospheric density is so low, this energy is capable of raising the temperature of the uppermost levels of the earth's neutral atmosphere (above 300 km) to 700 to 1500 K. The diurnal cycle of heating and cooling creates pressure variations that set up winds at this level with speeds on the order of 100 meters per second.

Recent satellite investigations and ground-based radar studies have led to better understanding of the neutral and ion photochemistry (*15*) and dynamic behavior (*16*) of the region above 90 km, the thermosphere. Less has been learned about the regions below the thermosphere, in part because the photochemistry appears to be more complex (involving many more species, including many trace substances) and in part because of the difficulty in making in situ measurements at these altitudes, which are too high for balloons and too low for satellites.

Despite the improved understanding of the behavior of the thermosphere, our ability to predict the neutral or electron density is hampered by the great variability in the EUV flux. Since this energy emanates from the chromosphere and corona, it is subject to day-to-day variations associated with solar activity, such as the formation of bright plages, and with cyclical variations caused by the sun's rotation (period of 27 days) as active centers grow and decay. The amount of activity is itself subject to the 11-year sunspot cycle, which introduces

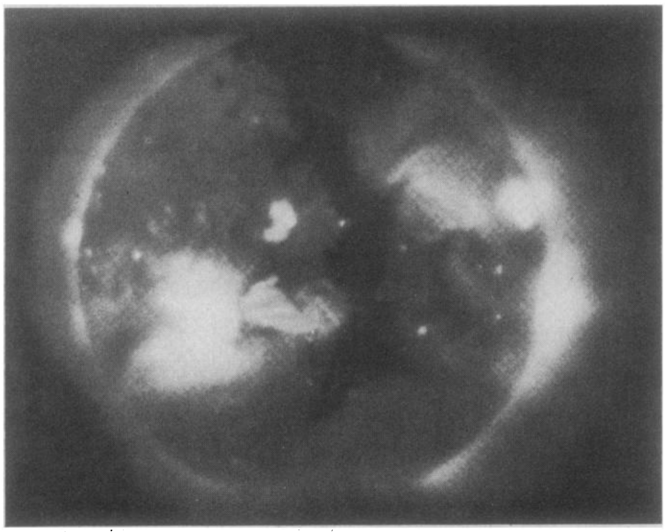


Fig. 3. Image of the sun taken in soft x-rays on 1 June 1973. A large coronal hole is visible as a dark area extending from the north pole down the center of the disk. [Courtesy of American Science and Engineering, Inc.]

about a 2:1 variation in the energy input to the thermosphere over a cycle and a corresponding variation in mean temperature. Direct satellite observations of solar spectral irradiance (*17*) have shown that the variations over the sunspot cycle are greatest at the shorter wavelengths and reach a factor of 10:1 near 300 Å. Even larger variations are seen at x-ray wavelengths; the energy below 10 Å, for example, exhibits a 500:1 variation over the sunspot cycle. The variability over shorter time scales is also largest at the shortest wavelengths; for example, brief increases in x-ray flux of as much as 100-fold have been observed during flares. These short wavelengths penetrate to below 100 km in altitude in the earth's atmosphere and contribute to ionizing these levels. The enhanced ionization increases the absorption of radio waves that otherwise would be reflected by higher levels of the ionosphere, creating a shortwave blackout.

The Solar Wind

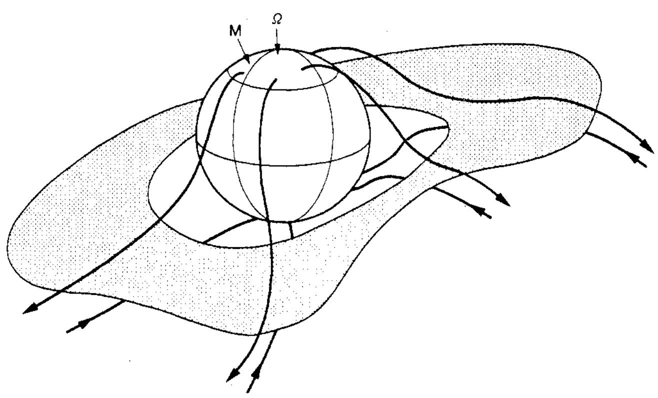
The earth's upper atmosphere derives a further input of energy from the solar wind, a general outflow of ionized particles from the sun (*18*). The first indications of such a particle flux came from studies of comet tail behavior (*19*); later it was shown, on theoretical grounds, that the solar wind is a natural consequence of the conditions in the corona (*20*). The continued increase in coronal temperature out to large distances from the sun causes the gas pressure to fall off less rapidly than the sun's gravitational attraction. This leads to a situation in which there can be no statically stable atmosphere that is gravitationally bound; instead, there must be a continuous outward expansion of the gas. The gas speed increases and becomes supersonic at a distance of a few solar radii. Thereafter the speed and direction remain constant, as there are no forces large enough to change them. Eventually the gas must undergo a shock transition at a distance where the dynamic pressure of the wind drops to the level of the pressure of the gas in interstellar space. This transition is thought to lie beyond the solar system; Pioneer 10, now midway between the orbits of Uranus and Neptune (about 25 times the distance of the earth from the sun), has not encountered any evidence of it.

At the orbit of the earth, the solar wind plasma consists principally of protons and electrons, with a small percentage of helium and other ions, all fully ionized. The number density varies from 2 to 100 ions per cubic centimeter, with a mean of about 10. Likewise, the bulk velocity is found to be variable over the range 200 to 800 km/sec, with a mean of about 450. The plasma carries with it the disordered solar magnetic field in the corona. At the orbit of the earth, the intensity is typically about 5 gammas (1 gamma = 10^{-5} gauss or 10^{-9} tesla) (*21*). This weak, disordered solar magnetic field influences the intensity of galactic cosmic rays reaching the earth. The gas temperature typically is about 10^5 K, but varies from 10^4 K at quiet times when the bulk flow is low to 10^6 K when conditions are disturbed and the bulk speed is high. On average, the total thermal energy is only about 1 percent of the energy of the bulk motion.

Increasingly sophisticated in situ measurements of the solar wind have been made since the early 1960's, when the first deep-space satellites were launched. The early observations were, for the most part, concerned with the microstructure of the local plasma properties that describe the overall flow (*22*). Yet, even before the first satellite observations, the existence of large structures, usually termed streams, had been surmised on the basis of observed variations in the earth's magnetic field (*23*). Recent satellite observations have shown that the solar wind is considerably more complex than might be supposed from these ground-based studies. One does not see discrete streams embedded in steady, low-speed flows. Rather, the latter are few and rarely persist for more than 2 days, while there are several kinds of high-speed flows that might be called streams. Many of these seem to be contiguous or even superimposed on one another (*24*). In the simplest cases, the solar wind velocity increases by several hundred kilometers per second over a period of 1 to 2 days and then decreases monotonically over 2 to 7 days. Compound streams show more complex variations of velocity with time and may be the result of the interaction of two or more simple streams. It is now believed that these recurrent streams originate from a relatively small number of localized regions of the sun known as coronal holes. These regions of relatively low density and temperature are best seen in images of the sun made at x-ray wavelengths (Fig. 3) (*25*). Coronal holes were first observed at a time near the sunspot minimum. The holes over the poles seemed to be semipermanent; those at lower latitudes had lifetimes of several solar rotations (*26*). In recent years, during which sunspots have been approaching their maximum, the polar holes have become less prominent and the mean lifetime of the equatorial holes has decreased to only one or two rotations.

No direct measurements have been made of the magnetic field in coronal holes, but observations of emission features in their vicinity suggest that the field is open and diverging (*27*). Timothy *et al.* (*28*) developed a plausible model for the existence of coronal holes. According to this model, holes form whenever the magnetic fluxes from bipolar magnetic regions interact to produce large regions of locally unbalanced flux, that is, regions that are dominated by a single polarity (this automatically includes the polar regions). In such regions, the field lines tend to extend far into the corona and do not impede the outward flow of plasma, which eventually succeeds in convecting the field away.

Efforts have been made to model the flow through coronal holes (*29*). Thus far it has been necessary to adopt some model for the streamlines a priori. It is generally assumed that the particles and the field lines diverge rapidly above the hole (*30*). Although the solar wind plas-



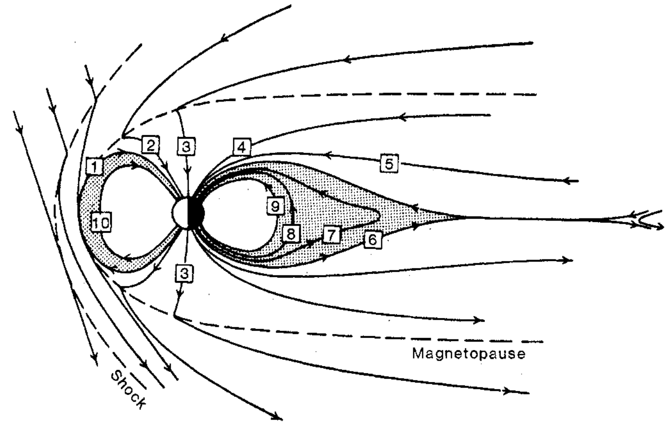


Fig. 4 (left). Sketch of the current sheet responsible for the sector structure. The axis of the current sheet, M, appears to be tilted with respect to the solar rotation axis, Ω (*32*). Fig. 5 (right). Sketch illustrating the distorted configuration of the outer parts of the earth's magnetic field and how this may be maintained by merging of field lines of solar and terrestrial origin. The numbers indicate in sequence the position of the field line after it has merged on the dayside (*35*).

ma moves radially outward from the sun, a stream appears to occupy a spiral path because the plasma in it moves with constant radial velocity while its source rotates with the sun. As may be deduced from Fig. 3 and from photographs taken during eclipses, semipermanent holes are thought to exist at the north and south poles near sunspot minimum when the sun's polar fields are strong. These polar holes are important in controlling the properties of the solar magnetic field at the orbit of the earth.

As observed at the earth, the solar magnetic field carried outward by the solar wind may be directed away from the sun for periods of many days. Comparable intervals then follow, corresponding to many degrees of solar longitude, during which the opposite polarity is observed (*31*). These intervals of similarly directed field have been termed sectors. Except for occasional short intervals in which the polarity is not well-defined, two, four, and occasionally six major sectors are observed per rotation. The transition from one sector to the next, the sector boundary, often precedes the arrival of a fast stream. While satellite measurements have not been made at very high heliocentric latitudes, there are indications that this sector structure disappears some distance above or below the solar equator; the dominant polarity at high latitudes corresponds to the polar magnetic field of the sun in the hemisphere in which the observations are made (*32*). According to this model (Fig. 4), the magnetic fields from the polar coronal holes are kept apart by the outward flow of ionized gas, which forms a thin current sheet. At the earth one dominant polarity or the other is seen depending on whether the observer is above or below the current sheet. It also appears that the current sheet is inclined with respect to the solar equator, so that at the earth two sectors are seen as the sun rotates. Four or six sectors can appear when the current sheet itself develops large-scale warps or undulations (a condition likened to a ballerina's skirt). The warps in the current sheet appear to be caused by the fast streams, which carry out fields from coronal holes located at more central parts of the disk.

Increased disturbance of the earth's magnetic field and increased energy input into the earth's atmosphere are associated with sector boundary crossings. These changes appear to be a result of the fast streams overtaking the slower solar wind ahead and compressing it, creating interaction regions in which the plasma density increases. The magnetic field carried by the plasma is also compressed in the interaction regions and has a tendency to turn southward and then northward. This southward turning of the field is an important parameter governing the interaction of the solar wind with the earth.

Somewhat less is known about flare-produced streams. It is believed that the flare initiates a large, explosive release of coronal material which arrives at the earth after 2 to 3 days as a shock front that has swept up the plasma and magnetic field ahead of it. Behind the shock is a shell of anomalously high He^{++} abundance. The highest flow speed occurs behind the He^{++} shell. It has been suggested that magnetic fields are looped back to the site of the flare, forming an ever-expanding magnetic bottle.

The Magnetosphere

The flow of solar wind around the earth has been a subject of considerable study since the space program began. This research culminated in the International Magnetosphere Study (1976 to 1979), a program involving spacecraft and ground-based observations by many nations (*33*).

The earth and its associated magnetic field represents an obstacle in the path of the solar wind. The result of the interaction of the two magnetized plasmas is to confine the extent of the earth's magnetic field in space. The volume occupied by the earth's field is known as the magnetosphere; its outer boundary is the magnetopause (*34*).

As a result of compression of the earth's magnetic field by the solar wind, the magnetopause is closest to the earth on the sunward side, where it typically extends about 10 R_E (Fig. 5). On the night side the magnetosphere stretches out like a comet's tail for at least 100 R_E, or well past the orbit of the moon.

There has been considerable controversy over the mechanism responsible for the distorted configuration of the earth's magnetic field. One view is that it results from some form of "viscous" drag applied by the solar wind to the magnetopause boundary, presumably through wave-particle interactions. The other view is that the distorted configuration results from the conversion of magnetic energy to particle kinetic energy through merging on the dayside between field lines carried by the solar wind (the interplanetary magnetic field, or IMF) with those of the earth. The merging should be especially effective when the IMF is directed south (Fig. 5) (*35*). In this case, the field lines of the IMF have the correct orientation to sever and connect with the lines of the earth's magnetic field. The effect of the merging is to connect the terrestrial field lines to the solar wind flowing past the earth, which carries them in an antisunward direction across the poles and into the tail, where they become very extend-ed. Eventually, the terrestrial field line reconnects with itself, leaving the IMF field line to be carried off beyond the orbit of the earth. This picture has received considerable support from observations showing that (i) the amount of energy coupled from the solar wind into the earth's atmosphere increases when the IMF turns southward (*36*) and (ii) the pattern of current flow produced in the upper atmosphere at high latitudes depends on whether the IMF is directed toward or away from the sun (*37*). However, the distorted field configuration and weak auroral activity persist even when the IMF is directed northward; therefore, it may be that the viscous mechanism is also operative (*38*).

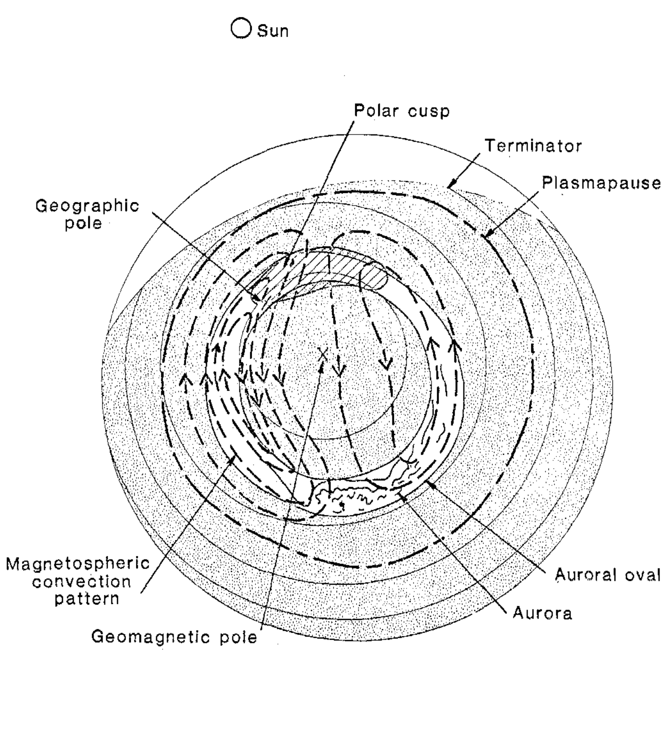
There cannot be a continuous accumulation of magnetic field lines on the nightside; once the field lines have reconnected, they circle back (through either the dawn- or the duskside of the earth) to the dayside, where they may undergo the whole process again. This picture of moving field lines is a convenience in describing the motion of magnetospheric plasma in a process known as convection. As seen from above one of the earth's poles, the feet of the field lines move in two vortices (Fig. 6). At ionospheric heights the velocity of motion of the field lines is typically a few hundred meters per second, but can reach 1 to 2 km/sec during disturbed times (*39*).

At altitudes above about 120 km the ions (and electrons) in the earth's ionosphere are constrained to gyrate around the field lines; thus the motion of the field lines induces a corresponding motion of the ionospheric plasma. Frictional heating of the thermosphere is then produced by the convective motion of the ionized particles through the neutral air whenever the ion speed exceeds a few hundred meters per second. This heating, usually termed joule heating, is greatest along two arcs through dawn and dusk, where the feet of the field lines return to the dayside along a narrow belt of latitudes between 65° and 75° magnetic latitude (Fig. 6) (*40*).

In addition to the heating produced by convection, there is heating associated with the bombardment of the atmosphere by energetic particles that produce visible auroras. These occur along an oval that surrounds the magnetic pole (Fig. 6).

The origin of auroral particles is not clear, since they appear to be considerably more energetic than those in the solar wind. The solar wind particles approaching the earth are slowed as they pass through a shock (the bow shock) about 15 R_E sunward of the earth. Much of the bulk motion of the particles carried through the shock is converted to random, disordered motion. While most of the solar wind flows around the flanks of the magnetosphere, a portion appears to penetrate the magnetopause in a thin region on the dayside (the entry layer) and to flow along the inside of the magnetosphere, forming what is called the plasma mantle (Fig. 7). The plasma in the mantle may also include ions of ionospheric origin released when the terrestrial field lines on which they were convecting became broken and connected to the solar wind. We do not yet have a complete picture of the motion of the plasma in the mantle, but it seems likely that this plasma becomes distributed throughout the magnetosphere. Current thinking is that some of the mantle plasma becomes captured by terrestrial field lines after they reconnect and is then carried toward the equatorial plane, where it forms what is termed the plasma sheet. The subsequent shortening of the terrestrial field lines as they return to a more dipolar configuration serves to energize the plasma sheet particles through the betatron effect. The accelerated ions precipitate into the atmosphere, giving rise to the diffuse or mantle aurora, although some are scattered into trapped orbits on permanently closed field lines; these became Van Allen radiation particles. However, this is not the whole story, since the most intense auroral arcs appear to be formed by the precipitation of electrons that have been accelerated by electric fields (of several kilovolts) oriented parallel to the earth's magnetic field lines and a few thousand kilometers in altitude. These parallel electric fields presumably are set up by the need to close the electric circuit linking the ionosphere and the magnetosphere, and they are the subject of considerable research (*41*).

The magnetospheric convection process is not smooth, but undergoes repeated intensifications (substorms) at intervals of 2 to 3 hours. Following an increase in the velocity of the solar wind or a southward turning of the IMF, the rate of transfer of magnetic flux from the dayside to the nightside increases without a corresponding immediate increase in the return flux. As a result, magnetic flux accumulates on the nightside until some form of instability triggers its release. In the process, some of the mag-



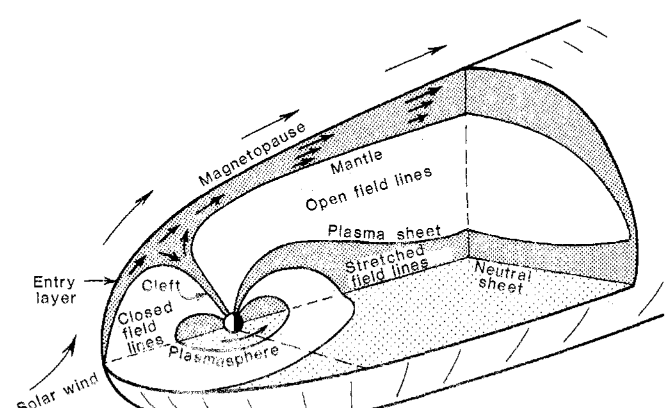


Fig. 6 (left). Path of the feet of the field lines (dashed lines) in the ionosphere above the north pole as a result of the convection process. The actual pattern and speed depend on the orientation of the interplanetary magnetic field. Also indicated is the location of the oval where visible auroras are most often seen. Fig. 7 (right). Topology of the magnetosphere in terms of its particle populations. The plasmasphere is a region of closed field lines populated by H^+ ions. These are produced in the ionosphere and diffuse to fill the flux tubes. The other populations are discussed in text (*33*). [Courtesy of J. Roederer] netic energy is converted to the energy of the plasma sheet particles in what may be the terrestrial analog of a solar flare. There is then a short-lived intensification of the aurora and an increase in the speed with which the field lines return to the dayside.

The Upper Atmosphere and Ionosphere

Although the properties of the upper atmosphere are largely controlled by the sun's ionizing ultraviolet radiation, below about 150 km there are pressure variations and winds caused by propagating atmospheric tides. These tides are the result of the absorption of ultraviolet by ozone and water vapor in the atmosphere below 100 km (*42*). Above about 200 km, the diurnal cycle of heating and cooling gives rise to a temperature variation of several hundred degrees (*43*). The resulting pressure variations establish winds that follow roughly great circle paths from the hottest region on the dayside to the coolest on the nightside. That is, unlike winds at the earth's surface, they blow directly from regions of high to low pressure, rather than circling them. This results from the presence of the ionospheric ions, which are constrained to follow the earth's magnetic field lines. Ion drag thus serves to balance the pressure force and limit the wind speed attained (*44*). The speeds increase on the nightside of the earth, where ion density decays.

Above about 100 km, the constituents of the earth's atmosphere are not mixed and each establishes its own density distribution, with the abundance of lighter constituents decreasing less rapidly with altitude than that of heavier constituents. To satisfy flow continuity and conservation of mass, the horizontal wind velocity increases with height so as to compensate for the altitude variation of the principal species (N_2) in the region where the winds are established. Since lighter constituents (such as O and He) decrease in abundance less rapidly than N_2, there is an overall flux of these lighter constituents both from day to night and (at solstice) from the summer to winter hemisphere. This transport creates diurnal and seasonal variations in the neutral composition at a given location.

The wind pattern established by absorption of ultraviolet radiation is modified by the effects of auroral heating and ion motion. The total heat input into the auroral regions as a result of joule heating and particle precipitation is thought to be only 1 to 10 percent of the global heating caused by ultraviolet and EUV (about 10^{12} W). Its importance lies in the fact that it is concentrated over a narrow interval of latitude and hence can introduce large latitudinal changes in pressure, unlike solar heating. Figure 8 illustrates the change in the mean meridional motion of the upper atmosphere which occurs at equinox as a result of increased auroral heating (*45*). Associated with this change is a redistribution of the light constituents of the upper atmosphere. In this case, high latitudes are depleted of atomic oxygen and equatorial regions are enriched (*16*). There is a corresponding change in the ionosphere, which is formed principally through ionization of this species.

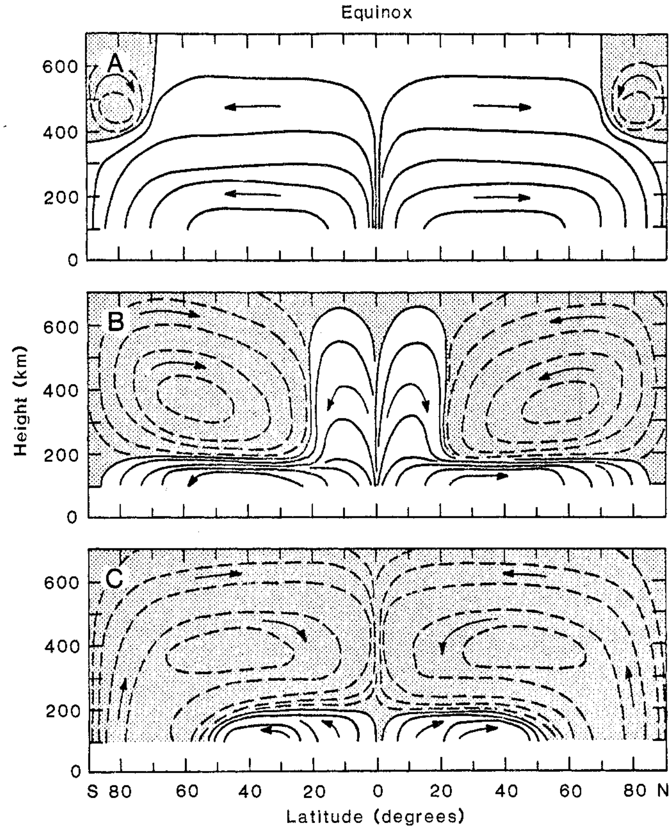


Fig. 8. Mass flow lines indicating the net transport of air (mean meridional flow) at equinox brought about by winds above 100 km during (A) quiet times with no significant auroral heating, (B) moderately disturbed periods, and (C) times of intense magnetic disturbances. [Courtesy of R. G. Roble]

It is believed that the change in mean meridional motion is brought about chiefly by an increase in the speed of the equatorward nighttime winds. Winds at 300 km as large as 500 m/sec have been observed at night blowing southward over North America during magnetically disturbed conditions (*46*). The rapid heating of the air at auroral latitudes during substorms is also observed to launch gravity waves in the atmosphere which propagate to lower latitudes and perturb the ionosphere.

The Lower Atmosphere and Weather

The idea that the earth's weather may in some way depend on solar events goes back for a century, yet there remains no conclusive evidence for this. Correlations that suggest a connection have been noted (*47*, *48*), but until a mechanism can be identified these must be treated with considerable caution.

The shortest period variations in atmospheric properties which appear to be correlated with a solar-dependent quantity are tied to the passage of a sector boundary in the solar wind; both the storminess at the level of the jet stream (defined by a vorticity-area index) and the vertical fair-weather electric field appear to decrease at such times. On a longer time scale, there is evidence for 11- and 22-year variations in rainfall rate that are correlated with the sunspot cycle. There have also been studies suggesting climatic variations. For example, it has been suggested that cold periods in Europe during the Middle Ages (such as the so-called "little ice age") were related to the absence of spots on the sun for several cycles (such as the Maunder minimum).

Most meteorologists are quite skeptical about such connections, since the variable fraction of the sun's energy appears to be quite small. In addition, they have questioned the validity of many of the correlations that have been found (*49*). It appears that if such a coupling does exist, then there must be a trigger mechanism whereby a small amount of energy organizes a much larger amount.

Among the coupling mechanisms that

have been suggested are (i) the manufacture at auroral latitudes of NO molecules through auroral precipitation and their subsequent transport to lower latitudes and altitudes by winds and diffusion (with consequent effects on the ozone abundance); (ii) changes in the electrical conductivity of the atmosphere that modulate the strength of the fair-weather electric fields produced by thunderstorms (this could have effects on cirrus cloud formation and thunderstorm frequency); and (iii) a direct forcing in the troposphere of long-period planetary waves by the 27-day rotational modulation of the solar constant (which might change the heat flow poleward from midlatitudes) (*50*). None of these or other proposed mechanisms has yet been shown to work. Indeed, the only well-established effects of solar variability detected below 100 km are (i) a modulation over an 11-year period of the intensity of cosmic-ray particles reaching the surface of the earth (and possibly affecting cloud formation), apparently caused by changes in the variability and average intensity of the IMF (increasing at sunspot maximum) and (ii) a decrease in ozone abundance at a level of 60 to 70 km, induced by solar flare–produced particles penetrating to this level of the atmosphere (*51*).

In the mistaken belief that the only way in which the sun could cause terrestrial magnetic storms was through the emission of magnetic waves, Lord Kelvin in 1892 concluded that the sun was not the cause of these disturbances (*52*). His mistake lay in assuming that he knew all the facts. Lest we fall into the same trap, we must be wary before reaching conclusions on a sun-weather connection. In the meantime, the topic remains intriguing, tantalizing, and far from settled.

Conclusion

Solar-terrestrial research is moving from the exploratory phase, in which the processes operating in different regions were first examined, to one in which there is an increasing effort to quantify our understanding and develop an overall picture. The accumulation of observational data has spurred attempts to construct models for the various processes involved. This, in turn, has led to the development of numerical computer simulations to test particular aspects of our understanding. These numerical simulations are still rudimentary, dealing only with limited aspects of the models they represent; much more theoretical work is required. This progress has also raised new questions, such as the mechanism by which the chromosphere and corona are heated, how (and where) the solar and terrestrial magnetic fields merge, how magnetic energy is converted to particle energy in the sun and in the earth's magnetosphere, and how intense electric fields are established above the earth's poles parallel to the terrestrial magnetic field. There is a complex and imperfectly understood set of linkages between the source of our energy—the sun—and the earth's atmosphere, and the study of these phenomena remains an important intellectual challenge.

References and Notes

1. The troposphere is beginning to be included in definitions of solar-terrestrial research as evidence is gathered suggesting a connection between sunspots and other forms of solar variability and the weather or climate.
2. Solar-terrestrial researchers distinguish between the middle atmosphere [which includes the stratosphere (10 to 50 km) and the mesosphere (50 to 90 km)] and a region above 90 km, the upper atmosphere [which includes the thermosphere (90 to 400 km) and the exosphere (> 400 km)].
3. Committee on Solar-Terrestrial Research, *Solar-Terrestrial Research for the 1980's* (National Academy of Sciences, Washington, D.C., 1981).
4. E. G. Gibson, Ed., *The Quiet Sun* (NASA SP-303, Government Printing Office, Washington, D.C., 1973).
5. C. Frolich, in *The Solar Output and Its Variation*, O. R. White, Ed. (Colorado Associated Univ. Press, Boulder, 1977), pp. 93–109.
6. G. Newkirk, in *Proceedings of the Conference on the Ancient Sun*, R. O. Pepin and J. A. Eddy, Eds. (Merrill, Boulder, Colo., 1979), p. 293.
7. It seems unlikely that the earth was ever completely ice-covered, since it would not recover from this situation; this probably rules out fluctuations in the sun's energy exceeding about 10 to 15 percent, provided that the earth's cloud cover and atmospheric constituents did not also change. See the recent review by G. R. North, R. F. Cahalan, and J. A. Coakley [*Rev. Geophys. Space Phys.* **19**, 91 (1981)].
8. M. Budyko, *Tellus* **21**, 611 (1969); *J. Appl. Meteorol.* **8**, 392 (1969). Such studies are, however, model-dependent. The sensitivity of the terrestrial atmosphere to changes in the input of solar energy can be demonstrated independent of climate models by the presence of variations in ocean cores that correspond to periods of the orbital variation of the earth [J. D. Hays, J. Imbrie, N. J. Shackleton, *Science* **194**, 1121 (1976)].
9. C. G. Abbot and L. B. Aldrich, *Ann. Smithson. Astrophys. Obs.* **6**, 85 (1942); L. B. Aldrich and W. Hoover, *ibid.* **7**, 26 (1954).
10. T. Sterne and N. Dieter, *Smithson. Contrib. Astrophys.* **3** (No. 3), 9 (1958).
11. P. Foukal, P. Mack, J. Vernazza, *Astrophys. J.* **215**, 952 (1977); P. Foukal and J. Vernazza, *ibid.* **234**, 707 (1979).
12. R. C. Willson, S. Gulkis, M. Janssen, H. S. Hudson, G. A. Chapman, *Science* **211**, 700 (1981).
13. J. R. Hickey, F. Griffin, H. Jacobowitz, L. Stowe, P. Pellegrino, R. M. Haschhoff, *Eos* **61**, 365 (1980).
14. P. Foukal, paper presented at the Workshop on Sunspots, Sacramento Peak Observatory, New Mexico, 14 to 17 July 1981.
15. D. J. Torr and M. A. Torr, *J. Atmos. Terr. Phys.* **41**, 797 (1979).
16. H. G. Mayr and I. Harris, *Rev. Geophys. Space Phys.* **17**, 492 (1979); J. V. Evans, *ibid.* **16**, 195 (1978).
17. H. E. Hinteregger, in *Advances in Space Research* (Pergamon, Oxford, England, 1981), pp. 39–52.
18. C. F. Kennel, L. J. Lanzerotti, E. N. Parker, Eds., *Solar System Plasma Physics* (North-Holland, New York, 1979), vol. 1.
19. L. Biermann, *Z. Astrophys.* **29**, 274 (1951); J. C. Brandt and D. A. Mendis, in *Solar System Plasma Physics*, C. F. Kennel, L. J. Lanzerotti, E. N. Parker, Eds. (North-Holland, New York, 1979), vol. 2, pp. 255–291.
20. E. N. Parker, *Astrophys. J.* **128**, 664 (1958); in *Solar Terrestrial Physics*, J. W. King and W. S. Newman, Eds. (Academic Press, London, 1967), chap. 2.
21. W. C. Feldman, in *The Solar Output and Its Variability*, O. R. White, Ed. (Colorado Associated Univ. Press, Boulder, 1977).
22. H. J. Völk, *Space Sci. Rev.* **17**, 255 (1975).
23. Investigations of the fluctuations of the earth's magnetic field revealed two general types of disturbance. One is a sudden decrease in the horizontal component of the earth's magnetic field which lasts one or more days and tends to follow large flares on the sun. The other disturbance also lasts for one or more days but does not follow a flare, is less intense, and recurs one or more times at 27-day intervals. These events are called sudden commencement storms and recurrent storms, respectively. See S. I. Akasofu and S. Chapman, *Solar-Terrestrial Physics* (Oxford Univ. Press, Oxford, England, 1972).
24. L. F. Burlaga and K. W. Ogilvie, *J. Geophys. Res.* **78**, 2028 (1973); L. F. Burlaga, *Space Sci. Rev.* **17**, 327 (1975).
25. J. W. Harvey and M. R. Sheeley, *Space Sci. Rev.* **23**, 139 (1979).
26. Many of the best observations were made by Skylab astronauts over the period May 1973 to February 1974. See *Coronal Holes and High Speed Wind Streams*, J. B. Zirker, Ed. (Colorado Associated Univ. Press, Boulder, 1977).
27. A. J. Hundhausen, in *ibid.*, p. 226.
28. A. F. Timothy, A. S. Krieger, G. S. Vaiana, *Sol. Phys.* **42**, 135 (1975).
29. S. T. Suess, *Space Sci. Rev.* **23**, 159 (1979).
30. R. H. Levine, M. D. Altschuler, J. W. Harvey, *J. Geophys. Res.* **82**, 1061 (1977).
31. J. M. Wilcox and N. F. Ness, *ibid.* **70**, 5793 (1965).
32. E. J. Smith, B. T. Tsurutani, R. L. Rosenberg, *ibid.* **83**, 717 (1978).
33. J. G. Roederer, *Science* **183**, 37 (1974).
34. ——, in *Solar System Plasma Physics*, C. F. Kennel, L. J. Lanzerotti, E. N. Parker, Eds. (North-Holland, New York, 1979), vol. 2, chap. 1.
35. I. Axford, *Rev. Geophys.* **7**, 421 (1969).
36. S. I. Akasofu, *Planet. Space Sci.* **27**, 425 and 1039 (1979).
37. J. Heppner, *J. Geophys. Res.* **82**, 1115 (1977).
38. P. H. Reiff, R. W. Spiro, T. W. Hill, *ibid.* **86**, 7639 (1981).
39. Movement of the field lines cannot be detected at ground level and occurs fully only above about 120 km. It is the layer of neutral air between the earth and the ionosphere which permits the field lines embedded in these two conducting regions to move with respect to one another.
40. There are several magnetic coordinate systems in use. That employed in Fig. 6 is defined by tracing a field line out to the equator, where its distance L from the earth's center (in earth radii) fixes the latitude Λ (= arc sec $\sqrt{L}$).
41. S. I. Akasofu, *Am. Sci.* **69**, 492 (1981).
42. J. M. Forbes and H. B. Garrett, *Rev. Geophys. Space Phys.* **17**, 1951 (1979).
43. A. E. Hedin *et al.*, *J. Geophys. Res.* **82**, 2193 (1977).
44. In the lower atmosphere the pressure force is balanced by Coriolis force, causing the winds to circle regions of high and low pressure.
45. R. G. Roble, in *The Upper Atmosphere and Magnetosphere* (National Academy of Sciences, Washington, D.C., 1977), chap. 3.
46. D. W. Sipler and M. A. Biondi, *J. Geophys. Res.* **84**, 37 (1979); G. Hernandez and R. G. Roble, *ibid.* **81**, 5173 (1976).
47. J. Eddy, Ed., *Solar Variability, Weather and Climate* (National Academy of Sciences, Washington, D.C., in press).
48. A recent summary of papers published in support of possible connections between solar variability and weather or climate may be found in J. R. Herman and R. A. Goldberg, *Sun, Weather and Climate* (NASA SP-426, Government Printing Office, Washington, D.C., 1978).
49. The climatic connection has been questioned by M. Stuiver, *Nature (London)* **286**, 868 (1980).
50. H. Volland, in *Solar-Terrestrial Influences on Weather and Climate*, B. M. McCormac and T. A. Seliga, Eds. (Reidel, Dordrecht, 1979), p. 263.
51. P. J. Crutzen, I. S. A. Isaksen, G. C. Reid, *Science* **189**, 457 (1975); D. F. Heath, A. J. Krueger, P. J. Crutzen, *ibid.* **197**, 886 (1977).
52. J. K. Hargreaves, *The Upper Atmosphere and Solar-Terrestrial Relations* (Van Nostrand Reinhold, New York, 1979), p. 267.
53. C. O. Hines *et al.*, Eds., *Physics of the Earth's Upper Atmosphere* (Prentice-Hall, Englewood Cliffs, N.J., 1965).

54. I am grateful to R. Carovillano, P. Foukal, A. F. Nagy, G. C. Reid, and P. H. Reiff for helpful comments on the manuscript. This work was supported by grant ATM 7909189 from the National Science Foundation.