IMPACTS OF SOLAR VARIABILITY

Solar Causes, Terrestrial Impacts, and Societal Effects

Activity-related events on the Sun lead to societal consequences on the Earth through a causal chain of ***solar events***, their physical ***impacts*** on the Earth, and the resultant ***effects*** of these impacts on human life and endeavor.

The initiating ***causes*** can be dynamic events at the Sun such as the expulsion of a CME or the eruption of a flare, or through more gradual forcing factors such as the Sun's 27-day rotation and the year-by-year variation in the extent of magnetic activity on the surface of the star. All of them impose variations—which can be sudden and large—in the amount of energy the Sun releases into space: in the intensity of electromagnetic radiation, the number and energies of solar atomic particles, or most often, in both of these.

Among the solar-induced ***impacts*** on the Earth—or in the now heavily-trafficked zone of space around it—are heightened levels of hazard in the near-Earth and space environment; disruptions and storms in the Earth's magnetosphere; the generation of strong electric currents that thread the atmosphere from top to bottom; changes in the ionosphere's reflective properties and in the chemistry of the stratosphere; and increased heating of the lower atmosphere, the surface of the Earth and the oceans.

The ***effects*** we feel on Earth or in near-Earth space cover a broad scope of national, commercial and personal activities and concerns. These range from satellite telephone calls and the nightly news on television to spaceflight, the transfer of electric power, national security, the transfer of financial and commercial information, weather and climate, and human health and well-being.

Some of the ***impacts*** of solar variability on different regions of the Earth and near-Earth space are summarized, with the solar ***cause*** of each of them, in the accompanying table. They are then explained, individually, in the text that follows it. A subsequent chart at the end of this section identifies some of the ***human activities*** affected by each of these ***impacts***. The specific societal ***effects*** in each of these areas of human activity are covered, one by one, in the next section.

Back to Table of Contents

IMPACTS OF SOLAR VARIABILITY ON THE EARTH AND NEAR-EARTH SPACE AND THEIR CAUSES ON THE SUN

LOCATION	IMPACT
Near-Earth Space	**Greater flux of high-energy atomic particles** *CAUSE: eruptions of solar flares and CME-driven shock waves and the rise in their occurrence as solar activity increases* **Increase in highly-energetic EUV and x-ray radiation** *CAUSE: eruptions of solar flares and an increase in the level of solar activity* **An 11-year cycle in the Earth's receipt of galactic cosmic rays** *CAUSE: effect of the 11-year variation in solar activity and its impact on GCRs*
Magnetosphere	**Initiation of geomagnetic storms and sub-storms** *CAUSE: transfer of energy from CME-driven shock waves; high-speed streams of solar wind plasma and impulsive events on the Sun*
Upper Atmosphere	**Displays of the aurora** *CAUSE: entry of high-energy particles from the Sun or from the magnetosphere into the upper atmosphere, where they strike atoms of oxygen and nitrogen, resulting in the emission of visible radiation in discrete colors* **Changes in the temperature and dynamics of the thermosphere** *CAUSE: day-to-day and impulsive variations in the amount of solar x-ray and ultraviolet radiation that is absorbed at this level by neutral atoms of air* **Variability in the electrical properties of the atmosphere** *CAUSE: the direct transfer of energy from the solar wind into the magnetosphere, combined with impacts of high energy particles from solar flares and CMEs. These generate and sustain electric currents that couple the magnetosphere, thermosphere, mesosphere and stratosphere together, with impacts on the surface electric and magnetic field* **Fluctuations in the concentration of free electrons in the ionosphere** *CAUSE: changes in the Sun's x-ray and extreme ultraviolet radiation that is absorbed by neutral atoms of air in the thermosphere and ionosphere to produce free electrons and charged ions in the ionosphere*
Lower Atmosphere, Oceans, and Land Surface	**Changes in the abundance of ozone and other traces gases in the stratosphere** *CAUSE: fluctuations in solar ultraviolet radiation and high energy atomic particles from solar flares* **Changes in the temperature of the stratosphere** *CAUSE: solar-driven fluctuations in ozone that affect the amount of solar near-ultraviolet radiation that is absorbed at this level* **Changes in the temperature and other meteorological conditions in the troposphere** *CAUSE: persistent changes in the total and spectral radiation received from the Sun, including the near-ultraviolet, which by altering the amount of ozone in the stratosphere can affect the transfer of radiation throughout the lower atmosphere* **Changes in the surface and sub-surface temperature of the oceans** *CAUSE: variability in total and spectral solar radiation, which is absorbed within the topmost (photic) zone of the ocean, to a depth of about 650 feet in the clearest water* **Changes in the surface temperature of the solid Earth** *CAUSE: variability in the total and spectral radiation received from the Sun*

Impacts on Near-Earth Space

The inner heliosphere in which our shielded planet is immersed is an open range of hostile fire: shot through by both harmful short-wave radiation and potentially lethal atomic particles from the Sun and the cosmos.

The dosages of solar electromagnetic radiation and of energetic particles tend to track each other and are highly variable: both from year to year in the course of the 11-year cycle of solar activity, and from day-to-day and even minute-to-minute in response to solar rotation and impulsive events on the Sun and disturbances in the solar wind.

The amount of change in the intensity of emitted radiation or energetic particles is determined primarily by magnetic conditions on the surface and in the solar atmosphere. Since the temperature of the solar atmosphere increases with height, so does the energy and variability of particles and radiation that come from each level. For electromagnetic radiation this means that as we move upward from the photosphere through the chromosphere and into the corona the nature of radiation the Sun emits systematically shifts toward shorter, more energetic and more potentially damaging wavelengths.

The least variable parts of the solar spectrum—which change by no more than a few tenths of a percent from day-to-day or in the course of the 11-year solar cycle—are the visible and infrared portions: the light and heat that we receive from the Sun, which come from the relatively-cool and stable 10,000° F photosphere.

In contrast, the near-ultraviolet radiation emitted from the higher and more dynamic 20,000° chromosphere changes by up to several percent in the course of the solar cycle: a hundred times more variable than radiation in the visible and infrared. Extreme ultraviolet (EUV) and still-shorter wavelength x-ray radiation come from even higher, hotter, and more variable levels: from the transition zone and corona, where temperatures range from about 200,000 to about 2,000,000° F, and from magnetically-heated material in solar flares.

During the rapid three-to-four year rise in solar activity that follows each minimum in the 11-year solar cycle the radiation from the Sun in the highly-energetic EUV increases by 10 to 100 percent or more. In the even more energetic and potentially lethal x-ray region, the Sun's output can increase by factors of more than 1000.

Thus the most variable and impulsive regions in the spectrum of solar radiation are also the most energetic and potentially damaging. And although few particles and no short-wave radiation reach the surface of the Earth—not even at the summit of Mt. Everest—they are present in full strength in near-Earth space, just above the top of our thin blue atmosphere, where spacecraft fly.

The most energetic solar particles—those with energies per particle of a few hundred thousand to a billion electron volts or more—also come from explosive and eruptive events on the Sun. The flux of these, too, though highly variable in both energy and number, follows the ups and downs of the solar cycle with many more of the most energetic when the level of solar activity is high.

The flux of cosmic rays—with energies per particle of 10^8 to 10^9 or more electron volts—is also modulated by solar activity, though in the opposite way. When solar activity is high and conditions in the solar wind more turbulent, fewer cosmic rays reach the Earth. When solar activity is reduced, more cosmic rays arrive. Those with energies in the awesome range of 10^9 to 10^{20} electron volts, which are not uncommon, pass right through the heliosphere and our own upper atmosphere with the greatest of ease, seemingly oblivious to the presence of the Sun or how active it might be.

SOLAR ENERGY RECEIVED AT THE EARTH IN THE FORM OF ATOMIC PARTICLES

SOURCE	FRACTION OF ALL ENERGY RECEIVED FROM THE SUN	ENERGY INCIDENT PER ACRE AT THE MAGNETOPAUSE: WHICH TOTALS 5.5 MILLION WATTS	RANGE OF DAY-TO-DAY AND 11-YEAR CYCLIC VARIABILITY	COMPONENT THAT REACHES THE EARTH'S SURFACE
Coronal Mass Ejections	0.0018 % over periods of tens of minutes	Up to 100 watts	Occur sporadically: energy deposited varies by order of magnitude	None
High Energy Protons From Solar Flares	0.00015 % over periods of minutes to hours	Up to 8 watts	Occur sporadically: energy deposited varies by order of magnitude	Secondary particles of lower energy only
Solar Wind	.00002 %	1.2 watts	An order of magnitude	None
Galactic Cosmic Rays	.0000006 %	0.03 watts	20-25% 11-yr variation	None

This 11-year heliospheric modulation of cosmic rays, combined with the occurrence of explosive solar flares and CMEs that march to the same drummer, continually alter the character and level of risk in near-Earth space.

Magnetic Storms

A major impact of the never-ending flow of solar wind plasma is a reshaping and extension of the lines of force of the Earth's extended magnetic field, which, were the wind turned off, would revert to a simpler and more symmetric form.

The solar wind shapes the magnetosphere on its Sun-facing side by flattening the magnetopause and pushing it much closer to the planet it protects. On the nightside, the transfer of momentum from the solar wind extends the magnetosphere into a long tail that extends more than 20 times its undisturbed length.

These changes are more than superficial, for they can weaken the magnetosphere's defenses by making it easier for solar particles to intrude.

The principal *driver* of major disturbances in the magnetic field of the Earth—called magnetic storms—is the extended magnetic field of the Sun, also called the interplanetary magnetic field or IMF. Magnetic reconnection is also involved, which directly links the considerable kinetic energy of a fast or CME-accelerated solar wind stream to the magnetosphere and ionosphere. The overpowering pull of the solar wind also sets up patterns of large-scale circulation in the magnetospheric and ionospheric plasma which in turn provide the energy for geomagnetic storms.

The principal *impacts* on the Earth of magnetic storms are induced electric currents that run all the way from the magnetosphere to the Earth's surface, and even beneath the ground and on the ocean floor.

* * *

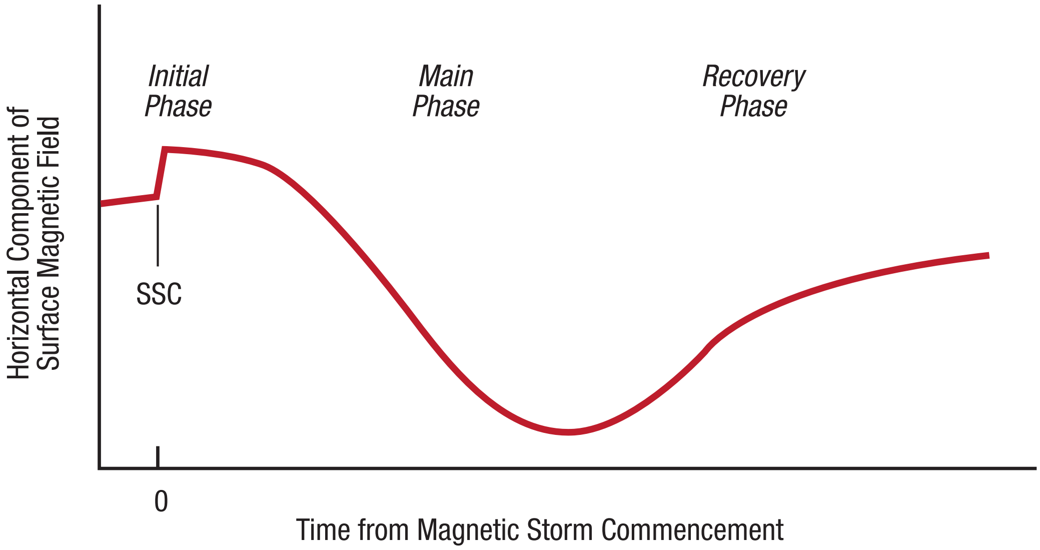
Magnetic storms are often initiated by the arrival of a plasma cloud that has been accelerated by a fast-moving CME, or by high-speed streams in the solar wind. To initiate a storm either of these must bring with it solar magnetic fields with a significant south-pointing component. Both are often preceded by a developing shock wave that sets the storm in motion.

The effect at ground level of these sudden disturbances is an abrupt and sustained jump—sensed around the world—in the strength of the horizontal component of the Earth's magnetic field. These appear very clearly in continuously running records at magnetic observatories, or in erratic behavior of magnetic compass needles.

Back to Table of Contents

These tell-tale warnings—which have been recognized as such since the mid-1800s—signal the onset, far above, of a magnetic storm that will trigger a chain of disruptive physical, chemical and electromagnetic processes in the upper atmosphere, as well as major impacts on a continually increasing number of human activities. They can also flip the switch that turns on the northern and southern lights, in colorful and moving displays of the aurora borealis and australis.

⊙ ⊙ ⊙



Sequence of transient changes in the relative strength of the horizontal component of the Earth's surface magnetic field that together define a typical magnetic storm. In the initial phase, beginning with a sudden storm commencement (SSC), the field abruptly increases in strength by a distinct rise of a percent or so. This is followed by a slow and continuous weakening (main phase), lasting from a few to several hours that depresses the strength of the field by many times this amount, and a subsequent protracted rise (recovery phase) to its previous value, lasting hours to tens of hours.

Geomagnetic storms unfold in a well-established sequence of three steps, the first of which, described above, is the warning phase, which is also known as a sudden storm commencement or *initial phase*. At this time the lines of force of the Earth's field on the Sun-facing side are squeezed inward toward the solid surface of the planet, producing the observed sudden jump in the strength of the field. For several hours following this abrupt initiation, the Earth's magnetic field remains stronger than normal.

⊙ ⊙ ⊙

During the *second* and so-called *main phase* of the storm, lasting hours to a day, the southward magnetic field in the cloud of solar particles makes contact with the Earth's oppositely-directed northward-pointing magnetic field. The two fields join in a process known as magnetic merging or reconnection and for a time the Earth becomes magnetically connected to an active region on the Sun, 93 million miles away.

The process creates open magnetic field lines rooted at either of the two magnetic poles of the Earth, some of which are then swept along the boundary of the magnetotail. With increasing distance from the Earth the magnetotail narrows, squeezing field lines of one magnetic polarity, rooted in the north magnetic pole, closer to those of the opposite polarity, attached to the south magnetic pole. At several fixed places, known as neutral lines, they are squeezed close enough to come in contact and join together, creating a long, closed magnetic loop with its two feet rooted back at the Earth in the two magnetic poles.

Streaming solar wind plasma that comes upon a neutral line in the magnetotail can pass through the magnetopause and be captured and accelerated in the magnetosphere, following one or other leg of the newly-formed closed magnetic loop back toward the Earth into a plasma reservoir, called the plasma sheet, in the magnetotail. There it mixes in with ions and electrons that have been carried up from the ionosphere.

⊙ ⊙ ⊙

Magnetic storms distend and disrupt the magnetotail, strengthening electric fields in the magnetosphere and ionosphere and accelerating particles in the magnetotail. Electric currents can at these times be shunted down into the northern and southern polar regions, where current-carrying particles collide with atmospheric atoms and molecules to produce displays of the aurora.

The systematic motions of positively-charged ions also intensifies an internal current in the magnetosphere known as the ring current that encircles the planet in a closed loop, confined to the plane of the magnetic equator, about 26,000 miles above the surface of the Earth.

At this altitude the ring current runs its circular course a few thousand miles beyond the farthest edge of the outer radiation belt, and about 8000 miles beyond the geosynchronous orbits where many modern spacecraft fly. One of its impacts on the magnetosphere is to suppress the strength of the Earth's magnetic field which, as we have noted, drops during the main phase of the storm to levels far below preceding conditions. The formation of a strong ring current is another defining feature of magnetic storms.

The powerful electric currents carrying more than a million amperes (!) in the high-latitude auroral region and the equatorial magnetosphere persist from

Back to Table of Contents

a few hours to as much as a day, affecting not only the magnetosphere but the thermosphere and ionosphere as well. In the magnetosphere, the radiation environment for spacecraft is made more hazardous. In terms of societal effects, the changes invoked by magnetic storms perturb broad segments of modern activities. Included are civil and defense communications; commerce and industry; the operation and control of manned and unmanned spacecraft; the utility of geographic positioning systems and other aids to navigation, electric power systems, undersea cables and the operation of satellite telephones.

* * *

In the final, *recovery phase* of the storm, which lasts several days, the strength of the Earth's field—and with it, affected parts of technologically-aided life on the planet—gradually return to normal.

THE THREE STAGES OF A GEOMAGNETIC STORM

STORM PHASE	DURATION	CHARACTERISTICS
Initial phase (sudden commencement)	Several hours	Compression of the magnetosphere; abrupt onset and sustained high field strength
Main phase	Hours to a day	Gradual drop in magnetic field strength at low latitudes; displays of aurorae at high latitudes; systematic motion of charged particles within the magnetosphere; creation of an electrical ring current around the Earth; ionospheric currents at mid and high latitudes; effects on modern technology and society
Recovery phase	Several days	Slow return to normal field strength

The Aurora

The aurora borealis, and its counterpart in the Southern Hemisphere, the aurora australis, are celestial displays of glowing light that illuminate a considerable portion of the night sky: sometimes as rays but as often in the form of dancing ribbons that look like swaying curtains of glowing color. They typically persist for several hours if not all night and can appear in pure shades of green, red, blue, or yellow, or when mixed together, white. Though wispy and ethereal, aurorae can be easily seen with the naked eye with some displays a thousand times more luminous than the Milky Way, and as bright as the full Moon.

* * *

It can be no surprise that auroral displays were so often cited in legend, lore and myth. Nothing else that happens in the dark night sky is as colorful and spectacular, as large and so much in motion, as potentially mysterious or quite as awe inspiring. In northern parts of Iceland, Norway, Sweden, Denmark and Finland—where they appear out of nowhere almost every night—aurorae were seen, for millennia, as an undeniable demonstration of the close presence and power of supernatural forces. Moreover, their mystical origin seemed substantiated by the perception of accompanying sound: a soft and eerie crackling or rustling noise that many today still claim to hear when aurorae flutter like wind-blown curtains across the darkened sky.

In parts of Europe, Eurasia and the New World where they were more rarely seen, the appearance of an aurora was long perceived—as were comets and eclipses of the Sun or Moon—as a portent of awful things to come: or depending on the inclination of the observer, as a divine blessing of a just-completed happening, such as a birth or coronation or a battle fought and won.

A "coronal aurora," in which the locations of the light-emitting atoms of air define the general direction in space from which the incoming particles responsible for the induced emission appear to have come, making it look as though the aurora itself streams downward toward the Earth from a point-source in the sky.

As told a few years ago by the beloved historian Shelby Foote, such was the perception of many Confederate soldiers when on the cold dark night of December 13, 1862 they looked up to see what seemed to be a heavenly endorsement—proclaimed in shimmering banners of multi-colored light—of their bloody, hard-fought victory that day in the Battle of Fredericksburg, Virginia. For many of them, born and raised in the deep South, this was very likely the first aurora they had ever seen, and for some the last, as well.

Back to Table of Contents

⊙ ⊙ ⊙

The number of aurorae you might expect to see in a year, or a lifetime, depends first of all on where you are. Those who live within or close to the Arctic Circle—in northern Alaska, Canada's Northwest Territories, Greenland, Iceland, or in the northern parts of Norway, Sweden, Finland and Russia can watch ever-changing displays of the northern lights on almost every clear night. They occur less and less often the farther south one goes from there, such that within much of the continental United States the appearance of a bright aurora is a relatively rare phenomenon. And particularly so in the electrically-lit skies of the modern world.

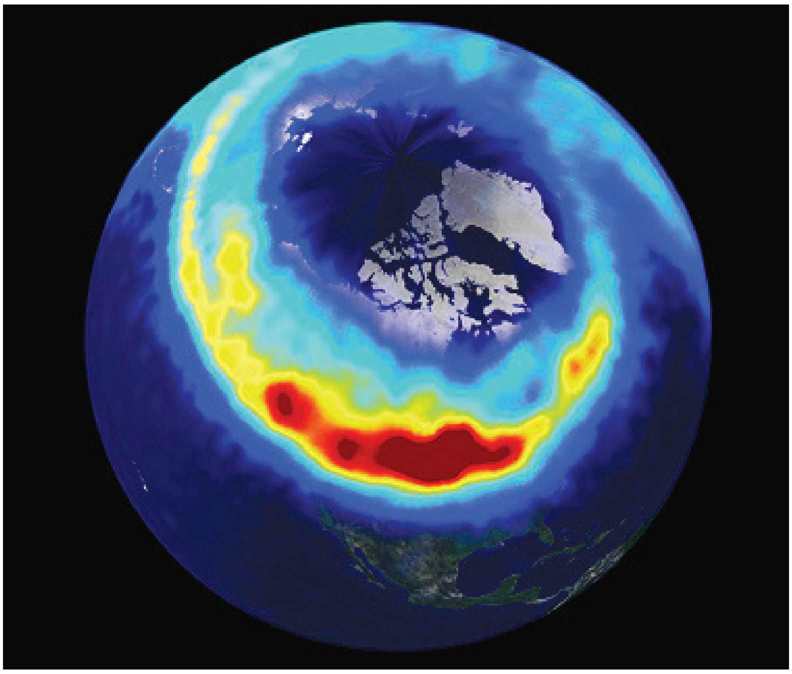
In non-polar regions of the globe the number of aurorae that are there to be seen depends as well on the phase of the 11-year solar cycle: for it is CMEs and high-speed solar wind streams that initiate the chain of events, including magnetic storms, that turn most bright aurorae on and off. Nights on which bright aurorae were seen have been recorded in dynastic histories in China and Korea and in European diaries, personal and professional, since long before the advent of the telescope in 1609 and the discovery of the sunspot cycle in 1843. In succeeding years, reports of bright aurorae have tracked the 11-year solar cycle so well that the more ancient auroral accounts are now used to help extend what is known of the behavior of the Sun through the last 2000 years.

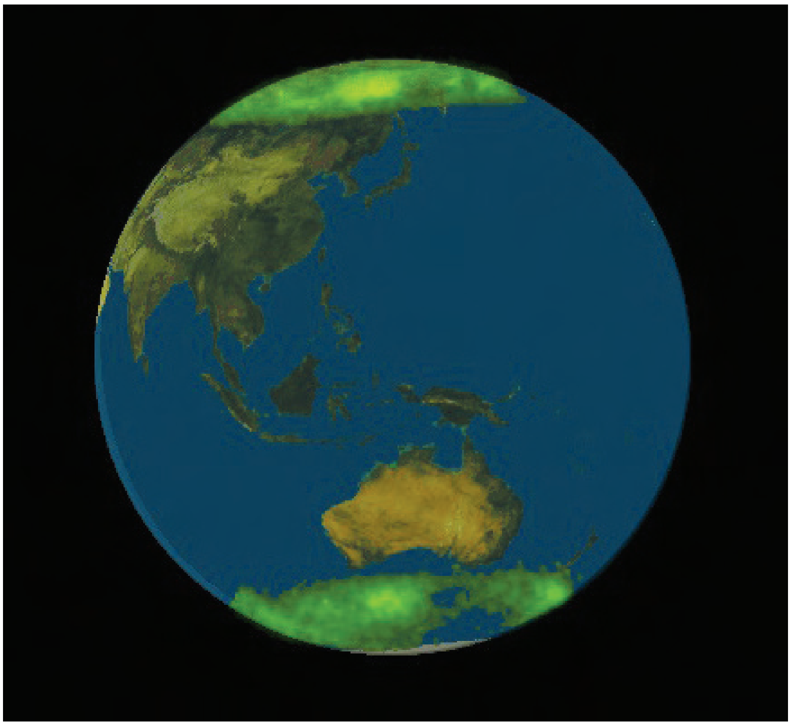
When averaged over the sunspot cycle, and in the absence of clouds, bright outdoor lights or a full Moon, about 100 auroral displays can be seen during an average year in Fairbanks, Alaska; 10 in Minneapolis; 5 in Chicago; 1 in Atlanta; and 1 per decade in Miami.

As the magnetic poles wander, as they always have, these zones of auroral occurrence will follow along, mile for mile. At the present time, the north magnetic pole lies some 800 miles south of the Earth's North rotational pole. Were it to continue its southward drift, the number of auroral displays seen in the sampled cities cited above would of course increase. Were it to reverse direction and drift north as far as the rotational pole, Chicago would see, on average, but one auroral display per year, Atlanta one per decade, and Miami perhaps one in a lifetime.

⊙ ⊙ ⊙

Auroral displays, which are also seen over the poles of Jupiter and certain other planets, are triggered by the arrival of streams of high-speed plasma from the Sun carrying southward magnetic fields. Auroral substorms, which are linked to the brightest and most dynamic aurorae, can occur during magnetic storms or quite on their own.





Images of the aurora borealis and australis, made from the POLAR spacecraft in a far-Earth orbit that crosses the poles of the Earth, offering distant views of all aspects of the entire spherical Earth.

Upper: the Northern Lights in full display, showing the complete auroral oval, here maximizing in brightness over northern North America, with portions of Greenland and northernmost Canada in sight within a more northern interior polar region where far fewer aurora occur.

Lower: the simultaneous display of the northern and southern lights, each centered on one of the two magnetic poles. Representations of the continents in these views of the night-side Earth are added artifacts.

Back to Table of Contents

Auroral substorms require an unstable stretched magnetotail with strong cross-tail currents. Magnetic storms provide these conditions in spades and drive a complicated array of auroral activity. But much weaker and shorter duration southward fields from the Sun can trigger auroral substorms without the need for magnetic storm conditions to develop. In either case, a portion of the current across the magnetotail is suddenly shunted through the auroral ionosphere.

In the course of these sudden events some of the stored electrons are released from the magnetosphere's mighty grasp and propelled downward at high velocities into the thermosphere, ionosphere, and upper mesosphere. There, in the thin air at heights of 50 to 200 miles above the surface of they Earth they collide with neutral atoms and molecules of nitrogen and oxygen.

These collisions release photons of light, in the pure colors that are emitted from individual atoms that have been energized in the encounter. The brightest and most common colors in auroral displays are green and red, both of which come from oxygen atoms. Blue, the next brightest, is produced by excited molecules of nitrogen.

* * *

And the mysterious whispering sounds, that much like UFOs, have long eluded proof of presence? The great distance (in most cases more than 100 miles) which separates the listener from the source, and the long delay (of up to ten minutes) that separates the initial arrival of light from the later arrival of sound makes any physical explanation highly improbable. The most likely answer, offered by the American physicist Elias Loomis in 1866, is that the perceived auroral sounds originate not up in the sky but here on the ground, between our own two ears.

We often hear what we expect to hear, and past experience—while witnessing fireworks displays or explosions, great or small—conditions our mind to *expect* that sound will accompany bright flashes of light: much as people nearly 2000 years ago, according to the Roman historian Tacitus, expected to hear a hissing sound while watching the setting Sun sink slowly into the ocean, just over the horizon, ten or twenty miles away; and often claimed they did.

* * *

Regular observations of aurorae from the vantage point of space allow a global view, seen from afar, of where they occur, confirming and extending what had been learned from the ground regarding the extent and characteristic patterns and differences of auroral displays.

Most aurorae in the Northern Hemisphere occur within an almost circular belt of latitude, 200 to 800 miles in width, known as the auroral oval, that is approximately centered on the north magnetic pole. A similar auroral oval, centered on the south magnetic pole defines the zone in which most aurorae occur in the Southern Hemisphere. The size of the auroral ovals is not fixed but elastic, their dimensions varying in response to the level of magnetospheric disturbances on any given day or year.

The north magnetic pole, as noted earlier, is now almost 800 miles south of the geographic pole and about 300 miles west of Greenland. At times of reduced solar activity, as at minima of the 11-year solar cycle, the northern auroral oval—which is centered roughly on that magnetic pole—can shrink in size until it is less than 1800 miles across: about the distance between Chicago and San Francisco. At these quiet times all of Alaska, most of Canada's Northwest Territories and all but the northern tip of Hudson Bay find themselves, like Denver or Havana, outside the shrunken auroral oval.

When magnetospheric activity increases, the northern auroral oval expands to a much larger size, often extending as far south as Seattle, Milwaukee, and Montreal, and—as was the case in the winter of 1862 (at the peak of the solar cycle that began at the minimum of 1856 and lasted through 1867)—on down to Fredericksburg, Virginia, where a war was being fought. During unusually energetic solar eruptions that trigger major magnetic storms the northern oval can stretch as far south as Mexico City and Panama.

⊙ ⊙ ⊙

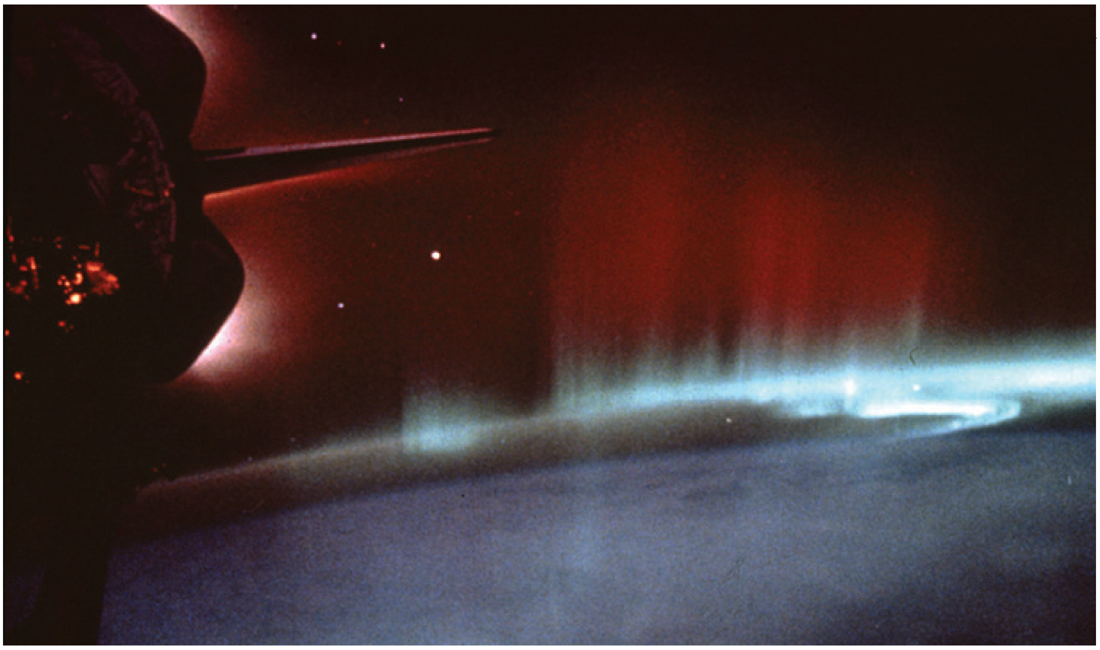
Another class of auroral activity occurs in association with the high-speed solar wind that blows outward from coronal holes. The wind from coronal holes is characterized by rapid variations in the imbedded magnetic field direction, with fields turning rapidly southward then northward then southward again over and over throughout the ten days it takes for a typical high-speed stream to reach the Earth. These short intervals of relatively weak southward magnetic field are sufficient to trigger substorm after substorm, lighting up the auroral oval repeatedly for a week or more. Strong high-speed streams reach a maximum in the declining phase of the solar activity cycle. Because of this, more energy pours into the auroral oval during years near solar minimum than near solar maximum when CMEs more seriously—but much less frequently—disturb the auroral oval.

The bright aurorae produced when the most energetic solar disturbances strike the magnetosphere occur almost entirely within the narrow band of the auroral oval. Seen from afar—as from the vantage point of space—the auroral oval

Back to Table of Contents

appears as an expandable belt of brightness that slides up and down in latitude from day to day and year to year as magnetospheric activity wanes and waxes. Nor is the band equally wide all around the globe, for it is three or four times wider on the *night* side of the Earth than on the sunlit, *day* side where unseen aurorae also continually occur.

These weaker, daytime aurorae—most of which are caused by the precipitation of trapped particles—cannot be seen from the ground, against the glare of the daytime sky. But they are always there, and always have been: like the many flowers, in the words of Thomas Gray, that are born to blush unseen, and waste their sweetness on the desert air.



A nighttime view from the Space Shuttle *Discovery* while flying over the southern hemisphere. Visible, at the right, a reddish haze above white spikes of light at the curved southern horizon is a display of the aurora Australis, or southern lights: the counterpart of the northern lights seen near the opposite pole of the Earth. Both are features of the low thermosphere at heights of about 50 to 80 miles above the surface of the Earth, well below the Space Shuttle's altitude of about 200 miles. More brightly glowing at the left are ionized exhaust gases from the Shuttle's engines, and their light reflected off the spacecraft's vertical stabilizer.

Displays of another weaker class of aurorae—produced when high-speed solar wind electrons work their way into the magnetosphere through openings in the magnetotail—are an almost nightly phenomenon in polar and sub-polar regions. These polar cap aurorae occur poleward of the bounded region of the auroral oval, in the space that separates it from the magnetic pole, and are seen throughout the solar cycle.

TYPES OF AURORAL DISPLAYS

TYPE	INTENSITY	WHERE SEEN
Active aurorae associated with geomagnetic substorms	Brightest and most dynamic	Within the auroral oval
Daytime aurorae	Weak	At lowest latitudes in the auroral oval
Polar cap aurorae	Weak	Polar cap regions

Impacts on the Upper Atmosphere

Atoms and molecules of oxygen and nitrogen in the upper atmosphere—at altitudes from about fifty to several hundred miles above the surface—absorb almost all of the short wavelength, extreme ultraviolet and x-ray radiation that the Sun delivers to the Earth. In this transfer of energy, the Sun heats the air in the high atmosphere by a thousand degrees or more to create and sustain the thermosphere, 200 to 500 miles above the surface, while stripping bound electrons from neutral atoms and molecules to create the ionized layers within it which are called the ionosphere.

Since they are created and sustained by short-wave solar radiation, both the thermosphere and ionosphere are directly and instantly affected by any variation in the radiative energy that the Sun emits in this highly-variable region of the solar spectrum. Solar radiation in the extreme-ultraviolet varies in the course of the solar cycle or in response to solar eruptions by 10 to 20 percent, and in the x-ray region by orders of magnitude.

As a result, the thermosphere and ionosphere take the hardest hammering from changing solar radiation of any region in the atmosphere. They are also far and away the most disrupted and variable, including during the daylit hours when they must respond to the Sun's daily movement across the sky.

Changes in the amount of solar short-wave radiation provoke rapid responses in the temperature of the thermosphere. One of the largest perturbations is due to the daily rising and setting of the Sun, which in a matter of minutes drive the thermospheric temperature up or down through a range of from several hundred to several thousand degrees Fahrenheit. The pace of this diurnal switching—as different parts of the thermosphere are carried by the Earth's rotation into and then out of the sunlight—is not to be compared to the gradual changes in temperature that we sense at dawn or dusk on the surface of the Earth. In the thin air of the thermosphere the swings in temperature are far wider and are accomplished much faster.

Back to Table of Contents

TYPICAL DAY AND NIGHTTIME SUMMER TEMPERATURES IN SAN ANTONIO AND IN THE UPPER THERMOSPHERE

CIRCUMSTANCE	IN SAN ANTONIO, TEXAS	IN THE UPPER THERMOSPHERE
Nighttime	75° F	600° F
Daytime, quiet Sun	95	900
Daytime, active Sun	95	2100
Daytime, during a solar flare	95	3100

The day-lit half of the thermosphere must also respond, equally fast and vehemently, to any significant change in the energy that the Sun delivers in the EUV. The occurrence of a solar flare can raise the temperature of the thermosphere by another 1000° F or more; as can the evolving background change in solar activity, from years of minima to times of maxima in the cycle.

Adding extra heat to the already hot thermosphere causes it to swell and expand upward. For example, the 1500° difference between nighttime and daytime temperature due to the Earth's rotation introduces a hemispheric bulge on the Sun-facing half of the thermosphere: a congenital lopsidedness which has been a feature of our planet from the time it had an atmosphere.

In addition to this daily oscillation, the vertical *extent* of the thermosphere has for billions of years been rising and falling on other scales of time: breathing out and breathing in—as solar activity waxed and waned—in response to day-to-day and year-to-year changes on a star 93 million miles away.

For the most part, these heat-driven waves at the top of the atmosphere were until the middle of the last century largely unnoticed or ignored—like the heaving of the middle ocean before the days of sail. But when man-made satellites began to circle the Earth, these variations came to be important.

When the thermosphere expands upward, it intrudes into the more tenuous regions through which many Earth-orbiting spacecraft fly. For the *Hubble Space Telescope*, the *International Space Station* or any of the other spacecraft orbiting the Earth in the 150 to 400 mile range, any increase in the density of the medium through which it flies will slow it down. This in turn reduces the diameter of its orbit and the height at which it flies, and hastens the day when it will plunge into the deeper atmosphere and return, in pieces, to the surface of the Earth.

Perturbing the Earth's Electric Field

Solar-driven changes in the temperature, density and dynamics of the thermosphere can also perturb the Earth's electric field: a fundamental property of the planet which we witness from time to time in the course of thunder and lightning storms.

Flashes of lightning mark times and places where the negatively-charged surface of the Earth is momentarily connected, through an electrical discharge, to the storm-generated, positively-charged upper layers of clouds that float above it. These spectacular happenings are but the most visible element of a global electric circuit that couples all of the atmosphere, from top to bottom, to the Earth beneath it: one of the few *direct* connections that link the top of the atmosphere to the biosphere, hundreds of miles below.

Thunderstorms are one of the three mechanisms that generate the electrical power that establishes and maintains a voltage of about 300,000 volts between the negatively-charged ground and the positively-charged ionosphere. One of the others is the dynamo that generates electric currents when the solar wind interacts with the Earth's magnetic field.

Winds in the thermosphere power a third electric current generator. These are driven, like the winds we feel at the surface, by differences from place to place in the local temperature and density, which in this case are induced by variations in the intensity of solar EUV radiation. At these lofty heights the thermospheric material that is carried in the winds is weakly-ionized plasma: ions and free electrons that act as an electrical conductor and current generator when they move through the lines of force of the Earth's magnetic field.

The currents added to the global circuit from any of these sources affect the difference in voltage between the ionosphere and the Earth's surface. Charged atomic particles, whether from the Sun or from more distant cosmic sources, can bring about a similar effect by reducing the electrical *resistance*—normally almost infinite—in the layers of air that separate the ionosphere from the solid or liquid surface of the planet. Alone, or acting together, these two impacts—one from the Sun's EUV radiation, the other from incoming charged particles—directly affect the global electric circuit, and perhaps the number of thunderstorms.

Back to Table of Contents

Restructuring the Ionosphere

The ionosphere, or ionized atmosphere, is the name given to the concentrated layer of charged particles—electrons and ionized atoms and molecules—that extends from the top of the mesosphere through the lower part of the thermosphere: from 40 to about 400 miles above the surface of the Earth, with a maximum charge density at a height of about 200 miles.

The electrons and ions are created and sustained there through the action of solar short-wave radiation, solar particles and cosmic rays on neutral atoms and molecules of nitrogen and oxygen, the most abundant species in the atmosphere. In the process of these encounters the neutral atoms and molecules are deprived of some of their bound electrons and become electrified, or ionized.

The outstanding societal effects of the Sun's creation and control of the ionosphere are its impacts on electronic communications of all kinds and at almost all frequencies, including radio, television, satellite telephone, navigation, data transfer systems, spacecraft control and countless military and other national security systems.

Disruptions in these and other telecommunications systems are caused by solar-driven changes in the concentrations of charged particles at different heights in the ionosphere, which alter how electromagnetic waves are reflected, absorbed or allowed to pass through it.

⊙ ⊙ ⊙

The number of electrons and ions in the ionosphere is determined by an ongoing give and take between their rate of production—governed by the varying intensity of solar short-wave radiation and incident particles—and the rate at which the newly-freed electrons and ions once again *recombine* to form reconstituted particles of neutral charge.

Thus, as a portion of the ionosphere is carried by the Earth's rotation into the darkened half of the planet, the rate at which electrons and new ions are produced falls rapidly toward zero: as does the number of charged particles within it, such that by the end of the night the ionosphere's structure and composition have changed dramatically. So much so that most of the lower ionosphere—beneath a height of about 70 miles—will have vanished, leaving only an upper part that extends high into the thermosphere, called the F-region.

Soon after sunrise, incoming sunlight restores the lost lower layers while increasing the thickness and density of charged particles in the much higher all-night region by about a factor of ten: in all, a dawn and dusk ritual, high in the sky, which has been acted out, unseen, since long before dinosaurs first appeared on the planet.

The reality of day-to-night changes in the structure of the ionosphere is quite apparent to anyone who when tuning in after dark to AM radio hears unexpected broadcasts from far away stations with strange sounding names: words, often in foreign tongues, that have been bounced and rerouted part-way around the Earth by the high nighttime ionosphere.

Other more disruptive and less predictable changes in the density of the ionosphere are the direct impacts of solar variability: specifically, changes in the amount of energy that the Sun releases in the form of EUV and x-ray radiation and the impacts of charged atomic particles. During a solar flare, for example, the ultraviolet and x-ray radiation from the Sun increases from 10 to more than 100%, as it does in the course of the 11-year solar activity cycle, and the energies of incoming solar particles by as much or more than that.

⊙ ⊙ ⊙

Since the late 1920s, the structure of the ionosphere has been described in terms of three distinct horizontal *layers*, distinguished by composition and vertical extent and labeled, from the lowest upward, the D, E, and F regions.

All three are made up of ions and free electrons, but the expected lifetimes of any of these short-lived atomic particles vary from layer to layer, due principally to the ever-decreasing density with height in the neutral atmosphere. In the F-region, which extends from 100 to 400 miles above our heads, the air is so thin that ions and free electrons less frequently come in contract with other particles, and as a result can exist an hour or so before they recombine.

The greater concentration of particles in the E-region (60 to 80 miles high) shortens the expected lifetimes of ionized particles to a few minutes; and in the most dense and crowded D-region (40 to 60 miles high) charged atomic particles remain that way for but a few seconds. It is because of these short lifetimes that densities of particles in the D-region and a part of the E soon fall to zero at nightfall.

Two factors work to keep the F-region of the ionosphere from disappearing altogether in the dark of the night. The first, noted above, is the fact that ions, once created, survive a while longer in less dense regions of the atmosphere

Back to Table of Contents

before recombining: an extended lease on life that lengthens with increasing height in the ionosphere.

The second is the replenishment of a fraction of the lost ions by the impact of nighttime cosmic rays on neutral atoms and molecules, supplemented by enfeebled solar ultraviolet radiation that has been scattered and redirected at high levels into the darkened half of the ionosphere from the adjoining daylit half. This weak leakage of light from day to night is not unlike the diffuse nighttime glow we commonly see, just over the horizon, from the lights of a neighboring city.

Some ions and free electrons are found in the upper atmosphere both above and below the portion that we call "ionosphere." With increasing altitude more and more of the neutral atoms and molecules become ionized, and the atmosphere is transformed from wholly neutral to almost completely charged.

The gradual changeover from neutral to ionized begins at an altitude of about 35 miles—not far above the stratopause—and continues all the way to the top: where the "atmosphere" ends and the exosphere and magnetosphere begin. At this ill-defined boundary electrons and ions flow both up and down: many of those carried up into the magnetosphere by the daytime thermal expansion of the thermosphere are drawn back down again at night as the thermosphere cools and shrinks.

What distinguishes the ionosphere from other regions of the atmosphere that also contain electrons and ions is the *number* of charged particles which are found there. For although the *ratio* of ionized to neutral particles is less in the ionosphere than in the high thermosphere and exosphere, there are many *more* of them per cubic inch or cubic mile. It is also in this unique band of altitudes—between about 40 and 400 miles above the surface—that the optimum is reached between the *intensity* of solar short-wave radiation (greater the higher you go) and the *number* of neutral particles to be ionized, which goes the other way.

⊙ ⊙ ⊙

But why D, E, and F? These familiar identifiers of the three layers in the ionosphere were assigned in the late 1920s, in the course of the first remote soundings of the region. In the heady days of initial discovery, it was thought prudent to hold the *first* three letters of the alphabet in reserve, in case additional, lower layers might be found. But as it turned out, neither these nor any of the letters that follow F were ever called into service.

THE LAYERED IONOSPHERE

REGION	HEIGHT IN MILES	CONTENT	TYPICAL LIFETIMES OF PARTICLES	PRINCIPAL IONS	MAIN SOURCE OF IONIZATION
F	100-400	electrons, ions	1 hour	O^+	solar EUV radiation
E	60-100	electrons, ions	1 minute	NO^+, O_2^+	solar EUV and x-ray radiation
D	40-60	ions and in the daytime, electrons	1 second	NO^+, O_2^+	solar EUV and x-ray radiation; solar particles and cosmic rays

Disturbing the Biosphere: The Lower Atmosphere, Oceans, and Land Surface

The most energetic and highly variable radiation from the Sun—in the short-wave EUV, x-ray and gamma–ray spectrum—is wholly absorbed in the upper reaches of the atmosphere and never makes it down into the biosphere, where all of life is found. Nor do any of the highest energy atomic particles that rain down on the Earth from the Sun and other cosmic sources.

What does reach the lower atmosphere, oceans, and solid surface of the planet are less energetic particles, light (visible radiation), heat (the infrared), and a small portion of the invisible near-ultraviolet: in all, about 60% of the energy the Sun delivered at the top of the atmosphere.

Some of the 60% is reflected back into space by clouds and land and sea and ice, but by far the largest part of this—about half of what the Sun delivers—is utilized in the biosphere. Among these tasks are heating the oceans and the solid surface, and through these, the lower atmosphere; creating winds and clouds and providing rain; regulating CO_2 and other greenhouse gases through the Earth's solar-driven carbon cycle that returns some of these pollutants to the oceans and ocean sediments, where through plate tectonics and volcanism they are eventually released; driving the hydrologic cycle that makes the rivers flow; fueling photosynthesis; establishing and sustaining the ozone layer in the low stratosphere; coloring the sky; and illuminating much of what we do.

Because the Sun's outputs of visible, infrared and near-ultraviolet radiation all vary to some degree in response to solar activity, all of these terrestrial services are potentially affected by changes on the Sun. In most of them, including all that involve visible or infrared radiation from the Sun, the immediate impacts of solar variability are far smaller than the changes induced by internal forcing in the system. Among these non-solar sources of variability are the day-to-night changes that result from the Earth's rotation; the perennial march of the seasons; recurrent ocean-atmosphere interactions such as El Niño/La Niña; the varying absorption of clouds, atmospheric aerosols and pollutants; and in the longer term, the introduction of greenhouse gases of human origin.

ORIGINS OF SOLAR RADIATION AND AFFECTED REGIONS OF THE EARTH'S ATMOSPHERE

SPECTRAL RANGE	**SOLAR ORIGIN**	**AFFECTED PART OF THE EARTH'S ATMOSPHERE**	**WAVELENGTH RANGE (IN NANOMETERS OR MICRONS μ)**
Visible and Infrared	Photosphere	Troposphere, stratosphere	400 nm to 10 μ
UV	Upper photosphere and chromosphere	Stratosphere, mesosphere, lower thermosphere	120 to 400 nm
EUV	Chromosphere, transition zone	Thermosphere; E and F layers of the ionosphere	10 to 120 nm
X-ray	Corona	D and E layers of the ionosphere	0 to 10 nm

Since solar radiation in the visible and infrared portions of the spectrum varies so little—less than half a percent in the course of a day, a year or the 11-year cycle—any of these competing weather and climate drivers can easily overwhelm and hide more subtle changes of solar origin, and particularly on shorter scales of time: minutes or months and even years.

As a result, it is often only in the longer term—through slower and more persistent forcing—that the deeper sounds of the Sun emerge through the noisy background chatter of shorter and more ephemeral perturbations.

More immediate and readily recognized are the impacts on the lower atmosphere of the near-ultraviolet solar radiation that reaches the lower stratosphere and upper troposphere. In the region between about ten and thirty miles above the surface, the absorption of solar ultraviolet radiation by molecular oxygen creates a small amount of ozone (about 1 part ozone to 10 million parts of air) which, though variable from place to place and time to time, is sufficient to shield life on the surface of the Earth from more lethal doses of solar ultraviolet rays.

Atmospheric ozone is also the creator of the sinuous vertical temperature profile of the Earth's atmosphere. This it does by absorbing solar short-wave radiation, slowing the steady drop in temperature with height in the troposphere, and at the tropopause, turning it completely around so that the temperature of the air *increases* with height through the next 20 miles of altitude. The result is the Earth's warm stratosphere: a feature of our planet that is unique in the solar system.

Both the thickness of the ozone layer and the temperature structure of the stratosphere are directly affected by changes in solar activity since they are each driven by solar radiation in the near-ultraviolet. But what the Sun giveth it also taketh away, for solar radiation in an adjacent region of the near-ultraviolet breaks ozone molecules apart.

Also at work is the destruction of ozone by chlorine, almost all of which finds its way into the stratosphere in the form of industrially-produced chlorofluorocarbons (CFCs), and by the catalytic influence of nitrogen oxides—much of which is produced by agriculture and the burning of fossil fuels.

The well-established solar control of stratospheric ozone offers as well a potential mechanism that might link solar variability to weather and climate, through connections that tie meteorological conditions in the stratosphere to those in the troposphere.

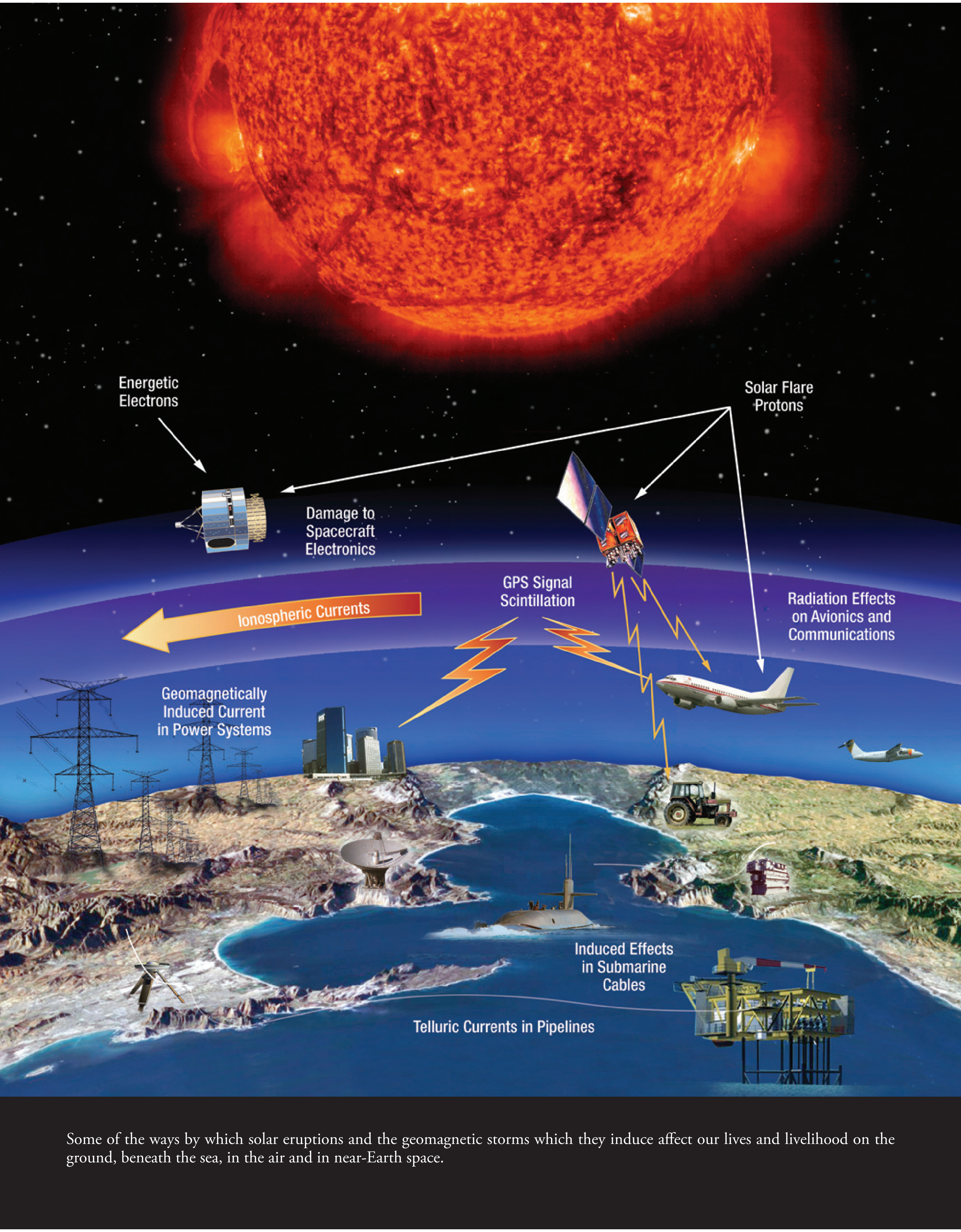
What is undeniable is the role of solar infrared radiation in establishing the temperature and circulation of the lower atmosphere, the temperature of the surface and subsurface ocean, and the temperature of the Earth's surface. Since the Sun's infrared radiation varies systematically in step with the solar cycle—rising slightly with increased activity and falling as it wanes—its influence should and does show up in records of all of these weather parameters.

Back to Table of Contents

IMPACTS OF SOLAR VARIABILITY ON THE EARTH AND NEAR-EARTH SPACE AND THE HUMAN ACTIVITIES THAT THEY ADVERSELY AFFECT

HUMAN ACTIVITIES SOLAR IMPACTS	AIR TRAVEL	HUMAN SPACE FLIGHT	OPERATION OF SPACECRAFT ANDSPACE EQUIPMENT	OBSERVATIONS OF THE EARTH FROM SPACE	TELECOMMUNICATIONS AND NATIONAL SECURITY	GEOGRAPHIC POSITION FINDING AND NAVIGATION	ELECTRIC POWER TRANSMISSION	OPERATION OF OIL AND GAS PIPELINES	GEOLOGIC SURVEYS AND EXPLORATION	CLIMATE
INCREASED LEVEL OF UV AND X-RAY RADIATION	■	■	■	■	■	■				■
INCREASED FLUX OF ATOMIC PARTICLES	■	■	■	■	■	■				
MAGNETIC STORMS	■	■	■	■	■	■	■	■	■	
AURORAL DISPLAYS										
THERMOSPHERIC TEMPERATURE		■	■							
INDUCED ELECTRIC CURRENTS			■	■	■	■	■	■	■	
IONOSPHERIC CHANGES	■	■	■	■	■	■				
STRATOSPHERIC OZONE										■
• AIR TEMPERATURE										■
• OCEAN TEMPERATURE										■
• LAND TEMPERATURE										■

Back to Table of Contents



Some of the ways by which solar eruptions and the geomagnetic storms which they induce affect our lives and livelihood on the ground, beneath the sea, in the air and in near-Earth space.