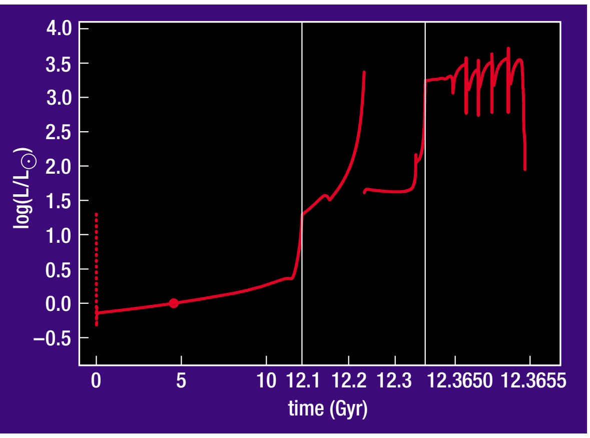
Solar variability and climate change: is there a link?

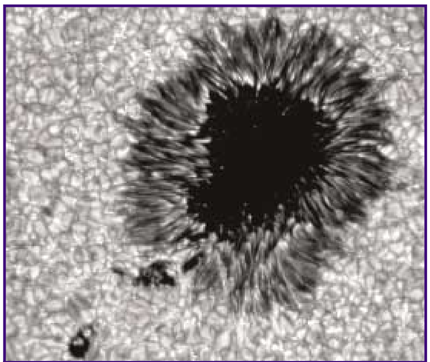
Sami K Solanki presents the Harold Jeffreys Lecture on the links between our climate and the behaviour of the Sun, from the perspective of a solar physicist.

Abstract

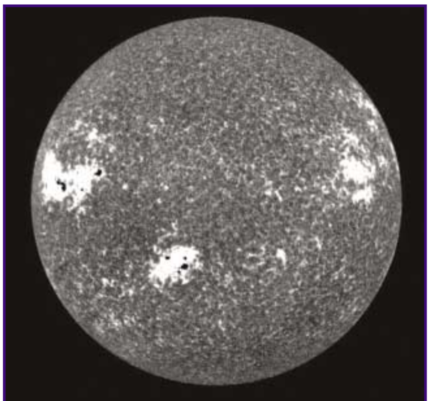
Radiation from the Sun makes Earth a habitable planet. Fluctuations in the solar output are therefore likely to affect the climate on Earth, but establishing both how the output of the Sun varies and how such variations influence Earth's climate have proved tricky. But increased amounts of data from the Sun and about the climate on Earth over recent years mean that rapid progress is being made. In this paper, I review the current debate on the influence of the Sun and summarize the state of play in this area of solar physics.



1: Computed luminosity of the Sun normalized to L$_\odot$, the luminosity at present, plotted as a function of time, starting from the Sun's contraction on to the main sequence (dotted line) and ending with the final contraction towards a white dwarf. Note the three different time scales, as the evolution speeds up. (Adapted from figure 4 of Sackmann *et al.* 1993, by permission.)



2: Image of a sunspot and the surrounding photosphere. The inner, darker part is called the umbra, the outer striated part is the penumbra. The sunspot is surrounded by granules, convection cells with a bright core harbouring hot upflows and dark boundaries (so-called intergranular lanes) composed of cool downflowing gas. (Figure kindly provided by T Berger, Lockheed-Martin Solar and Astrophys. Lab.)



3: An image of the full solar disc taken in the light of singly ionized calcium. The dark features are sunspots, the bigger bright patches are faculae and the small bright features present all over the solar disc are network elements.

The Earth is a warm and cosy cradle dangling in cold and largely empty space made even less hospitable by harmful high-energy particles and short-wavelength radiation. Two of the factors critical for the survival of life are the protection provided by the Earth's atmosphere and the warmth bestowed by the Sun's radiation. Since the Sun is by far the largest supplier of energy to the Earth's surface, any change in the radiative output of the Sun also affects the energy balance of the Earth's surface and atmosphere, so that at some level it influences our climate. Changes in the solar spectrum, in particular in the UV, could enhance (or dampen) this influence, by affecting stratospheric chemistry: most importantly the balance between ozone production and destruction (each driven by radiation at different wavelengths). Finally, the Sun may influence the Earth's climate also in other, more convoluted ways, e.g. by modulating the flux of cosmic rays, which have been proposed to increase the coverage by low-lying clouds.

But how strongly does the Sun vary and to what extent does it influence the Earth's climate? It is important to find a firm answer to these questions in order to put the solar contribution into proper perspective within the global warming debate and in particular to determine its weight relative to that of the man-made greenhouse gases. At present, the question of the Sun's contribution to global warming can only be partly answered, but progress has been rapid in the last years and shows no sign of slowing down. In the following I'll attempt to give a brief tour of the current status from a solar physicist's point of view. Let me begin by describing the relevant features of the Sun.

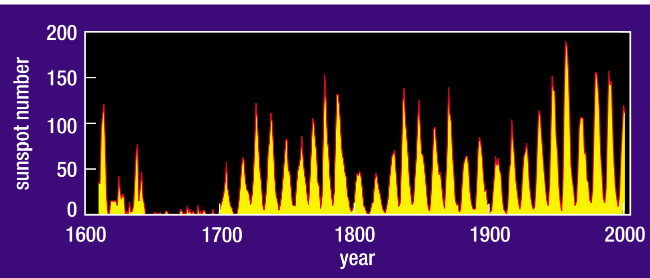
Solar variability

There are two major causes of solar variability: one is solar evolution, driven by conditions in the Sun's core; the other is the magnetic field of the Sun, or rather the field located in the solar convection zone (i.e. in the outer part of the solar interior) and in the Sun's atmosphere. The Sun is typical among stars of similar mass, composition and age. It is currently near the middle of its approximately 10-billion-year tenure on the main sequence. During this time the Sun is expected to roughly double in brightness and also to increase substantially in radius. The evolution of the Sun's luminosity is plotted in figure 1. Upon leaving the main sequence the evolution of the Sun speeds up, with large and rapid excursions both in brightness and radius. Obviously the Sun will eventually play havoc with the Earth's capability of harbouring life as we know it. This is far in the future, however, and need not concern us here. But what about the past? The luminosity of the Sun at the beginning of its tenure on the main sequence was roughly 30% lower than its current value. In the absence of other changes this would imply that the Sun provided insufficient energy to keep the Earth's surface from becoming largely covered with ice. Since ice reflects almost all the incoming radiation, this enhanced albedo would make it impossible for even the bright present-day Sun to melt this ice cover. The fact that the Earth is not wrapped in ice suggests that something else did change in the course of the Earth's history. Carl Sagan was among the first to propose that the presumable evolution of the composition of the Earth's atmosphere nicely balanced the evolution of the Sun's luminosity. The concentration of greenhouse gases was much larger in the atmosphere of the young Earth, so that over billions of years solar variations and changes in greenhouse gas concentrations have equally shaped the Earth's climate.

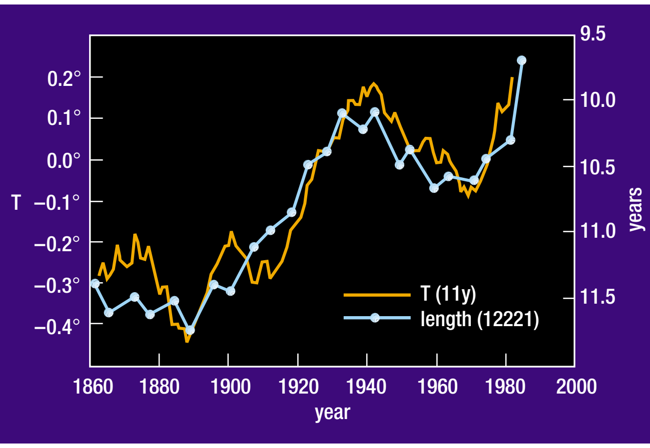
The other major force leading to a variation in the solar output is the Sun's magnetic field. It is generated by a dynamo located at the bottom of the convection zone, which lies roughly 30% of the distance from the solar surface to the solar core. The major ingredients determining the strength and structure of the resulting magnetic field are the (differential) rotation of the Sun and the (turbulent) convection at and below the solar surface. The differential rotation produces a mainly toroidal field near the base of the convection zone. With time the strength of this field increases. Above a certain critical strength the field becomes unstable and individual loops start to rise towards the solar surface, which they finally reach and pass through. At this point the magnetic field becomes accessible to observation. The interaction with the convection leads to the concentration of the field in filaments or bundles of field lines called flux tubes. The largest of these, the sunspots, have diameters similar to that of the Earth and are visible as dark features on the solar surface. A relatively symmetric sunspot is shown in figure 2. The more common smaller flux tubes appear as bright points having diameters below 300 km and are called magnetic elements. Concentrations of large numbers of these are visible as bright faculae in the active regions (where also the sunspots are found) and as a network distributed over the whole solar surface (figure 3).

Whereas most of the magnetic field lines piercing the solar surface form loops and head back to the Sun within a few solar radii (they form what is called the closed magnetic flux), a small fraction is carried out by the solar wind into interplanetary and finally interstellar space (usually called the open magnetic flux, more from a parochial solar physicist's point of view than with any aim of undermining Maxwell's laws of electrodynamics). In particular the closed magnetic field gives rise to a large number of phenomena, such as sunspots, chromospheric plages, hot coronal loops, filaments and prominences, flares and the associated high-energy

4: The yearly sunspot-number record since the beginning of telescopic observations. The period in the second half of the 17th century that was practically free of sunspots is called the Maunder minimum.



5: Eleven-year running mean of the annual average northern hemisphere land-air temperature relative to the average temperature 1951–1980 and the filtered length of the sunspot cycle. (From Friis-Christensen and Lassen 1994, by permission.)



radiation, and coronal mass ejections, to name but a few. These phenomena are collectively described under the heading of solar activity.

The solar magnetic field, and hence also the associated activity, is strongly time-dependent. The most prominent feature of this time dependence is the solar activity cycle with a period of roughly 11 years. This cycle is beautifully illustrated by the number of sunspots visible on the solar disc at any given time, which increases tenfold or more between activity minimum and maximum (see figure 4). At the same time sunspots provide the longest direct record of solar activity.

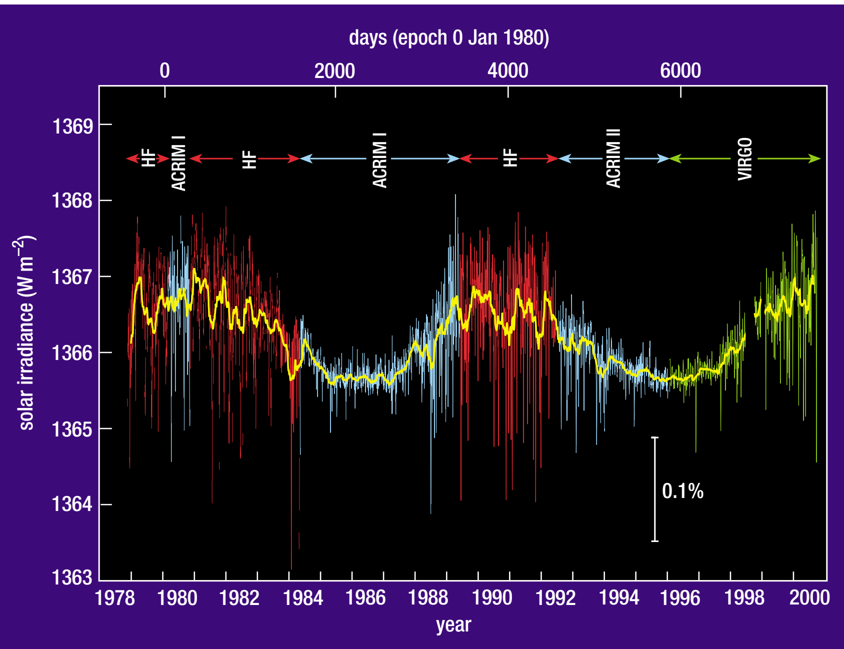
It is obvious from figure 4 that some cycles are stronger than others. In addition, some cycles are also longer than others, although these variations are not sufficiently large to be clearly visible from the figure. In extreme cases cycles become so weak that they cease to be recognizable as such. This was the case during the Maunder minimum in the latter half of the 17th century, when practically no sunspots were present on the solar surface. Incidentally, this coincided with the so-called Little Ice Age in Europe, a time of severe cold and great hardship, when the Thames froze regularly and alpine glaciers grew deep into the valleys.

The sunspot record and other, often more indirect, proxies of solar magnetism have invited comparison with records of climate. For several such records a significant correlation is indeed present. One of the most striking correlations was found between the air temperature above land masses in the northern hemisphere and the length of the sunspot cycle by Friis-Christensen and Lassen (1991), shown in fig. 5.

If the correlation is not a product of chance, which cannot be ruled out completely for time series of this length, then there must be a physical link with an unknown mechanism. One of the relevant forms of solar variability must therefore be closely connected with the length of the solar cycle, a parameter previously not known to play a decisive role in solar variability. The forms of solar activity that may be relevant for climate are total and spectral irradiance (i.e. the flux spectrum seen from the Earth) and the strength of the Sun's open magnetic field, which modulates the cosmic-ray flux reaching the Earth.

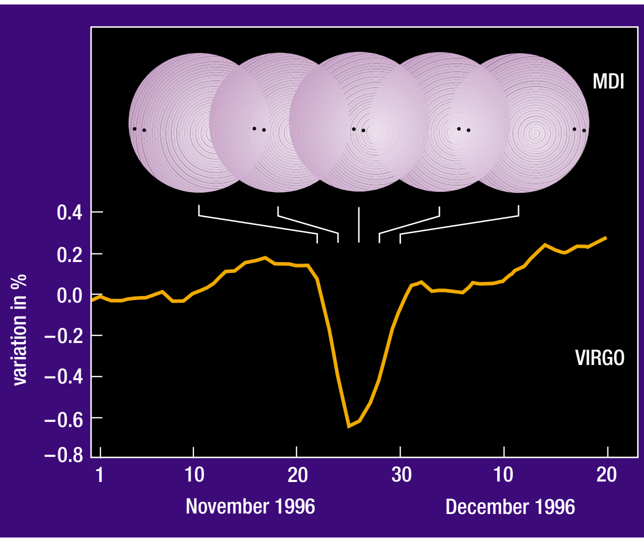
Solar irradiance

Let us now consider one of these forms of solar variability in greater detail, namely fluctuations in the brightness of the Sun seen as a star, as measured from above the Earth's atmosphere, the so-called solar irradiance. The total solar irradiance (i.e. irradiance integrated over all wavelengths) has been measured since 1978 (using radiometers onboard spacecraft) with sufficient accuracy to detect intrinsic solar changes and is displayed in figure 6. The plotted record is pieced together from different data sets because no single instrument survived the whole period. In addition to the presence of the 11-year cycle with an amplitude of roughly 0.1% in total irradiance, significant shorter fluctuations are present which give the diagram a noisy appearance. These fluctuations are mainly solar; the amplitude of this "noise" changes in phase with the cycle and the major excursions all point downward, implying a short-term darkening. One such excursion is plotted on an expanded



6: Composite of total solar irradiance covering more than two solar cycles, as measured by radiometers flying on spacecraft. Data from four instruments have been used to create this composite. Details of the procedure employed to merge the data sets are provided by Fröhlich and Lean (1998).

7: Dip in total solar irradiance measured by the VIRGO instrument on the SOHO spacecraft (orange curve). The face of the Sun in white light on five selected days during the dip is also shown. Note the sunspots on the solar disc.



horizontal scale in figure 7, together with continuum images of the solar disc on five days during the temporary darkening. These reveal a pair of small sunspots crossing the solar disc, which are the cause of the darkening. The other darkenings in figure 6 are also all associated with sunspots or sunspot groups. Sunspots are dark because their strong magnetic field suppresses convection, which is the dominant form of energy transport just below the solar surface. The vertical transport by radiation is too inefficient to compensate for the loss of convective transport and sunspots are too large for a horizontal inflow of radiation to have any significant effect on their brightness.

Figures 6 and 7 indicate that the energy blocked by a sunspot is stored within the Sun to be released at another time. This temporary storage of heat is only possible because the convection zone has a very high thermal conductivity and heat capacity, as pointed out by Spruit (1982). This combination implies that heat blocked by a sunspot from reaching the surface will quickly diffuse throughout the convection zone (due to the high thermal conductivity) whose temperature will be raised only imperceptibly (due to the large heat capacity), so that all in all the solar brightness is decreased. The stored heat is eventually radiated away, but only very gradually over a period of 10^5 years (corresponding to the thermal relaxation time of the convection zone).

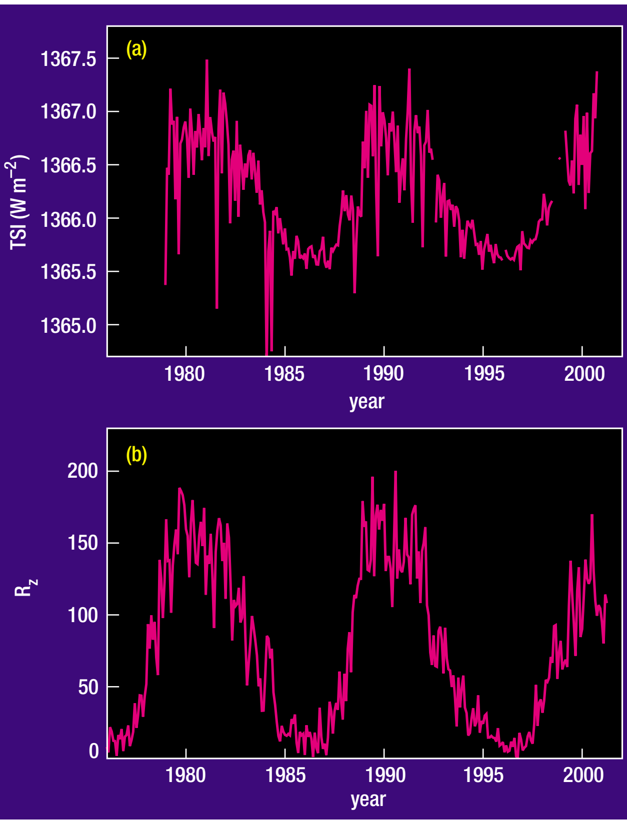
This important insight leads to the next question: when the Sun darkens in the presence of a sunspot, why is the Sun on average brighter during solar activity maximum, i.e. when more sunspots are present on the solar disc, as can be seen from figure 8? One school of thought maintains that the magnetic field influences convection and/or other processes in the solar interior sufficiently to enhance the energy flux at times of large solar activity. Recent work has, however, given more credence to the other point of view that it is the magnetic field at the solar surface that produces this brightening. The basis for this hypothesis lies in the fact that sunspots are only the largest and by far the rarest of the broad range of magnetic features (magnetic flux tubes) in the solar photosphere. The smaller flux tubes forming the faculae and the network (see figure 3) are bright. Like sunspots they also inhibit convection in their interiors, but they are sufficiently narrow that radiation flowing in from the sides into the highly evacuated flux tubes more than compensates for this loss. Like the sunspots, the number of these small magnetic elements also increases from activity minimum to maximum. Furthermore, the area covered by these elements increases by a far larger amount than that covered by the sunspots. Consequently, the brightening due to the faculae outweighs the darkening due to the sunspots in the long run, i.e. when averaged over multiple solar rotations.

Support for the important role of the magnetic field at the solar surface is provided by the fact that the irradiance variability can be reproduced quantitatively by a simple three-component model, with the individual components representing the quiet Sun, faculae and sunspots. Figure 9 shows a comparison between measured and reconstructed solar irradiance. Note the excellent agreement between model and data, both on short time scales on which the Sun rotates and active regions evolve (upper two frames) and on the longer time scale from solar activity minimum to maximum (lower frame). This suggests that the basic premise underlying such a model is correct and that it is indeed the manifestations of the magnetic field at the solar surface (i.e. sunspots and faculae) that are responsible for the irradiance variations.

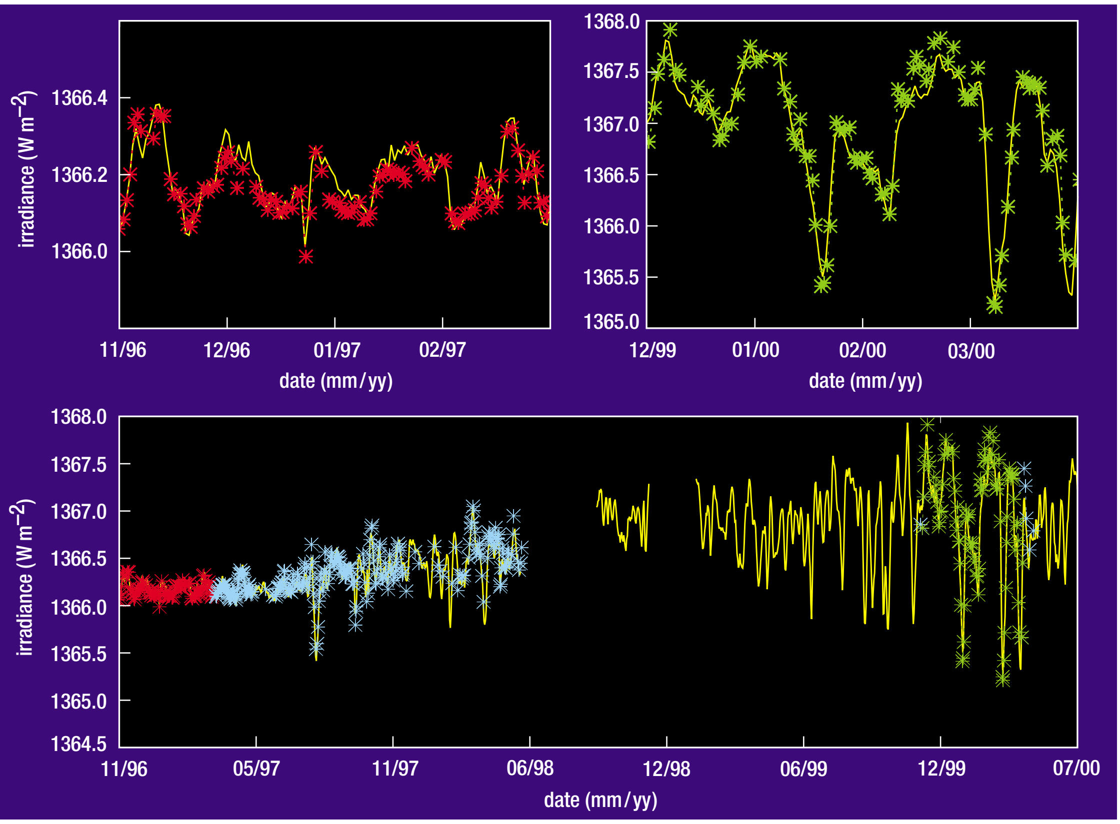
Obtaining an understanding of the processes leading to the variability on a solar-cycle time scale is not sufficient, however. Firstly, a purely periodic variation of solar brightness does not produce a long-term change in solar brightness and hence also no contribution to global warming. Now, the solar cycle is not quite periodic and its amplitude has varied considerably in the course of the last centuries, but that alone produces an increase of less than 0.05% in solar irradiance over the last 150 years. This is unlikely to be sufficient to influence climate in a significant manner. Is there any reason to believe that the Sun has exhibited larger changes than these over the last few centuries? Different pieces of evidence and lines of thought suggest that this is the case. For example, a comparison between the Sun and other field stars indicates that even at activity minimum the Sun exhibits much stronger signs of magnetic activity than many other stars. Quantitatively, during the last few cycles the Sun's Ca II H and K brightness, which is a measure of a star's magnetic activity, always remained in the top third of the range spanned by field stars. Interestingly, stars with the lowest Ca II brightness also do not exhibit any cyclic variability, which has been interpreted to mean that they have been observed in a Maunder minimum-like state. Assuming that the relationship between solar activity and irradiance found over the solar cycle also acts over longer times, it is then possible to work out that the Sun was between 2 and 4 $W m^{-2}$ less bright during the Maunder minimum than today.

Support for this conclusion is provided by the fact that the magnetic field of the Sun does not disappear at activity minimum when there are no sunspots on the solar surface, but rather that a background of magnetic flux, partly concentrated into the magnetic network, continues to be present. There is evidence from multiple sources that this magnetic background changes with time. The strongest such evidence comes from the increase in the geomagnetic AA-index over the last 150 years, which has been used by Lockwood *et al.* (1999) to reconstruct the interplanetary field, closely related to the Sun's open magnetic flux. Their reconstruction reveals that the Sun's open flux varies secularly and has doubled in the course of the last century.

Somewhat surprisingly, a relatively simple model accurately reproduces the reconstruction of Lockwood *et al.* (1999), while at the same time also returning the evolution of the Sun's total magnetic flux (Solanki *et al.* 2000, 2002). The critical insight entering the model is that although solar cycles delineated by prime indicators of solar activity, such as sunspots, do not overlap, this is not the case for the magnetic field. Indeed, considerable magnetic flux from the old cycle is still present on the solar surface when flux due to the new cycle starts to erupt. One can distinguish between flux belonging to the two cycles from its location on the solar surface and the relative orientation of the positive and negative polarities of the field joined together by a magnetic loop. Consequently, the magnetic flux on the Sun is never allowed to decay completely, unless there is a long gap similar to the Maunder minimum between two cycles. In figure 10 the predictions of the model of Solanki *et al.* (2000) for the open magnetic flux are compared with the reconstructions of

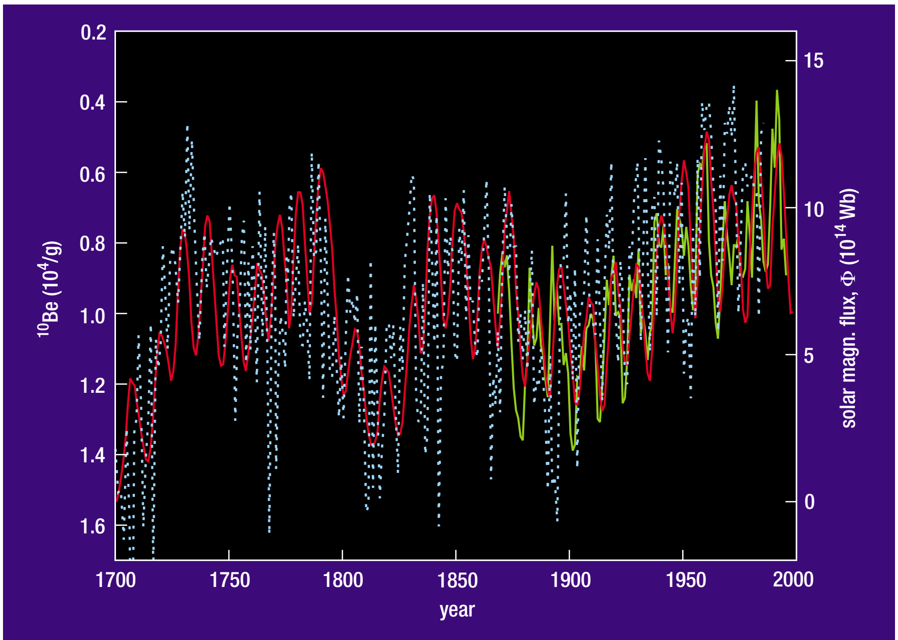


8: Upper frame: the same total solar irradiance record as in figure 6, but now after the application of a one-month running mean to the data. Lower frame: a similarly smoothed record of sunspot number.



9: Total solar irradiance measured by VIRGO (solid curve) and as reconstructed using a model assuming that the magnetic field at the solar surface is responsible for irradiance variations (stars). The period from the last activity minimum to near the current activity maximum is shown in the lower frame. Blow ups of the two shorter periods indicated by the red and green stars are presented in the upper frames.

Lockwood *et al.* (1999) and with the concentration of ^{10}Be. This radioactive isotope is produced by the interaction of cosmic rays with air molecules in the upper atmosphere. Its production rate in historical times may be deduced from ice cores taken from large glaciers. Since the cosmic-ray intensity is modulated by processes that scale with the strength of the Sun's open field, the ^{10}Be production rate is expected to roughly follow the Sun's open flux. Note the good agreement between the three curves (in the case of ^{10}Be the agreement is better on longer time scales due to the high-frequency noise present in the data). One interesting aspect is that according to this model the secular evolution of both the Sun's open and total magnetic flux follows a combination of solar-cycle amplitude and length. This is the first indication on a physical basis that cycle length, which has been found to correlate very well with climate indicators (see figure 5), indeed influences solar irradiance, a parameter that may be of direct relevance to the climate.



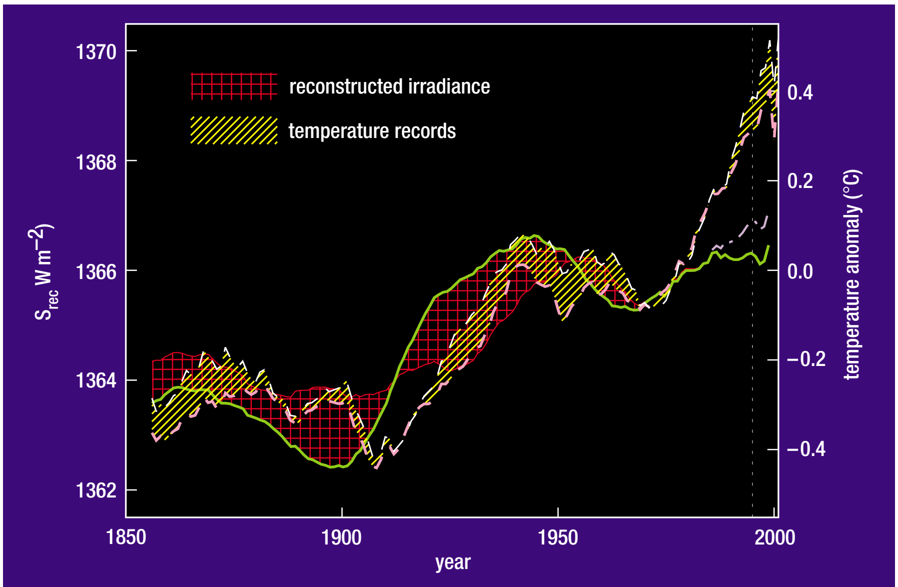
10: Evolution of the open magnetic flux at the solar surface since the end of the Maunder minimum in 1700. Model predictions by Solanki *et al.* (2000) are represented by the red curve, reconstructions by Lockwood *et al.* (1999) based on geomagnetic indices by the green curve and the ^{10}Be concentration in ice cores (corresponding to the inverted scale on the left *y*-axis, Beer *et al.* 1990) by the dotted curve.

Solar spectrum changes

With the help of the Sun's total and open magnetic flux and of historical records of sunspot numbers and other proxies of solar activity, it is possible to reconstruct not just the total irradiance of the Sun and its variation over the last couple of centuries, but also changes in the solar spectrum (in particular the irradiance in the UV) and in the cosmic-ray flux during this period. All these quantities, which represent different paths by which the Sun could affect climate, are found to evolve in a very similar manner. To what extent they actually do affect climate is still a matter of debate, with considerable work being required to pin down the contributions of the various possible mechanisms acting in the Earth's atmosphere. As an illustration it is quite instructive simply to compare solar and climate time series. If they are similar, then it does not prove that the Sun causes climate change, but if they are sufficiently different, this allows limits to be set on the solar influence.

In this spirit, the total irradiance since 1850 is compared with climate records in figure 11. Figures showing the UV irradiance and cosmic-ray flux are similar and lead to the same conclusions as can be drawn from the plotted curves. The part of the figure shaded red represents very roughly the uncertainty in the irradiance reconstruction. It lies between two reconstructions based on somewhat different assumptions. The part shaded yellow highlights the difference between global and northern hemisphere temperatures. All curves have been smoothed by 11-year running means. Obviously, prior to roughly 1980 the solar irradiance on the whole ran parallel to and even slightly ahead of the Earth's temperature.

This is consistent with a causal relationship between the two and supports, but by no means proves, the view that the Sun has had an important, possibly even dominant influence on our climate in the past. Other contributors to climate variability are volcanic activity, the internal variability of the Earth's atmosphere and man-made greenhouse gases. After 1980, however, the Earth's temperature exhibits a remarkably steep rise, while the Sun's irradiance displays at the most a weak secular trend. Hence the Sun cannot be the dominant source of this latest temperature increase, with man-made greenhouse gases being the likely dominant alternative. ●



11: Two reconstructions of total solar irradiance combined with measurements, where available (enclosing the red shading) and two climate records (enclosing the yellow shading) spanning roughly 150 years.

Sami K Solanki, Max-Planck-Institut für Aeronomie, Max-Planck-Str. 2, 37191 Katlenburg-Lindau, Germany.

References

Beer J *et al.* 1990 *Nature* **347** 164.
Friis-Christensen E and Lassen K 1991 *Science* **254** 698.
Friis-Christensen E and Lassen K 1994 *The Sun as a Variable Star* (eds) J Pap *et al.* Cambridge University Press p339.
Fröhlich C and Lean J 1998 *GRL* **25** 4377.
Lockwood M *et al.* 1999 *Nature* **399** 437.
Sackmann I-J *et al.* 1993 *ApJ* **418** 457.
Solanki S K *et al.* 2000 *Nature* **408** 445.
Solanki S K *et al.* 2002 *A&A* **383** 706.
Spruit H C 1982 *A&A* **108** 356.