LONG-TERM VARIATIONS IN THE MAGNETIC FIELDS OF THE SUN AND POSSIBLE IMPLICATIONS FOR TERRESTRIAL CLIMATE

M. Lockwood [1, 2], and S. Foster [2]
[1] Rutherford Appleton Laboratory, Space Science Department, Chilton, Oxfordshire, England, U.K.
[2] Southampton University, Department of Physics and Astronomy, Southampton, Hampshire, U.K.

ABSTRACT

Recent studies of the variation of geomagnetic activity over the past 140 years have quantified the "coronal source" or "open" magnetic flux F_s that leaves the solar atmosphere and enters the heliosphere and have shown that it has risen, on average, by 34% since 1963 and by 140% since 1900. This variation is reflected in studies of the heliospheric field using isotopes deposited in ice sheets and meteorites by the action of galactic comic rays. The variation has also been reproduced using a model that demonstrates how the open flux accumulates and decays, depending on the rate of flux emergence in active regions and on the length of the solar cycle. The cosmic ray flux at energies > 3 GeV is found to have decayed by about 15% during the 20th century (and by about 4% at >13 GeV). We show that the changes in the open flux do reflect changes in the photospheric and sub-surface field which offers an explanation of why open flux appears to be a good proxy for solar irradiance extrapolation. Correlations between F_s, solar cycle length, L, and 11-year smoothed sunspot number, R_{11}, explain why the various irradiance reconstructions for the last 150 years are similar in form. Possible implications of the inferred changes in cosmic ray flux and irradiance for global temperatures on Earth are discussed.

Key words: Solar variability; climate change; heliospheric field; solar irradiance; cosmic rays

1. INTRODUCTION

1.1. Solar Irradiance Variations

There has been a growing interest in long-term variations of the Sun. What was once termed the "solar constant" has been shown to vary with the sunspot cycle and thus is now generally referred to as the "total solar irradiance", S. The amplitude of this variation is of order 0.1% [*Willson*, 1997; *Fröhlich and Lean*, 1998]. Irradiance variations are mainly caused by the combined effect of darkening by sunspots with the brightening of associated faculae and of the network. Of the two, the brightening is the dominant effect [*Chapman et al.*, 1997], but both are associated with the magnetic field at the solar surface [*Fligge et al.*, 1998]. Variations of the irradiance on this 11-year timescale have not been considered to be greatly significant for Earth's climate because their effects are smoothed by the heat capacity of the oceans [*Wigley and Raper*, 1990]. However, continuous and homogeneous sunspot observations have been made since 1749, and sunspot numbers show longer-term, secular variations (on timescales of order 100 years and greater) in addition to the 11-year solar activity cycle [*Gleissberg*, 1944; *Pulkkinen et al.*, 2000]. Any associated changes in the total solar irradiance on these longer timescales would be significant for global climate change because they would not be smoothed. Thus an understanding of the ratio of the amplitudes of the variations on 11-year and longer timescales is of great importance. Using various proxy data, attempts have been made to extrapolate the recently observed total solar irradiance variation back in time [*Hoyt and Schatten*, 1993; *Lean et al.*, 1995; *Solanki and Fligge*, 1998; 1999; *Lockwood and Stamper*, 1999; *Lean*, 2000] and these reconstructions have been used to evaluate its rôle in the rise of average surface temperatures on *Earth* [*Lean et al.*, 1995; *Lockwood et al.*, 1999b; *Tett et al.*, 1999].

1.2. Cosmic Rays

The Earth is shielded from galactic cosmic rays by the heliosphere. A number of processes are active, although their relative importance is still a matter of debate. However, the heliospheric magnetic field is the key component of this shield [*Moraal et al.*, 1993; *Potgieter,* 1995], such that the cosmic ray fluxes seen at Earth are very highly anticorrelated with the local heliospheric field (the interplanetary magnetic field, IMF) [*Cane et al.*, 1999]. The field is dragged out of the sun by the continuous, but highly variable, solar wind and fluctuates with the level of magnetic activity on the sun. This results in the fluxes of those cosmic rays that do penetrate the shield and reach the Earth, showing a clear solar cycle variation, with a strong anticorrelation with sunspot numbers.

Cosmic rays are a key part of the global electric circuit that is driven primarily by thunderstorms [*Bering et al.*, 1998]. They generate air ions in the subionospheric gap which allows current to flow between the tops of thunderclouds and the ionosphere and also between the ionosphere and the ground in association with the fair-weather electric field. It is not known what sort of modulation to this circuit could be brought about by the changes in cosmic ray fluxes, not what the implications of this might be.

Svensmark and Friis-Christensen [1997] and *Svensmark* [1998] have reported a solar cycle variation in the global fraction of terrestrial cloud cover seen from space. These authors proposed that this is because

Proc. 1st Solar & Space Weather Euroconference, 'The Solar Cycle and Terrestrial Climate', Santa Cruz de Tenerife, Tenerife, Spain, 25-29 September 2000 (ESA SP-463, December 2000)

cloud cover is directly influenced by the galactic cosmic rays incident on Earth. The cloud data used by *Svensmark and Friis-Christensen* [1997] and *Svensmark* [1998] were compiled from a variety of sources. Concerns about this compilation of diverse and uncalibrated cloud data were pointed out by a number of authors (see review by *Soon et al.* [2000]) and the correlation was made weaker when data for after 1992 became available [*Kristjánsson and Kristiansen*, 2000]. There was also no clear dependence on cloud type, as would have been expected. Another criticism has been that the polar regions were excluded from the study because cloud cover there could not always be distinguished from the ice caps. However, subsequent work has shown that there is a strong correlation for one subset of the cloud cover, as seen at infrared wavelengths [N. Marsh and H. Svensmark, private communication, 1999; *Marsh and Svensmark*, 2000]. These observations are possible at all latitudes (for both day and night), and so true full-globe averages could be used. The data are the "D2" set compiled and inter-calibrated by the International Satellite Cloud Climate Project (ISCCP) [*Rossow et al.*, 1996]. The strong correlation with cosmic ray fluxes is for clouds that are inferred to be at low altitudes and so is associated mainly with stratus and stratocumulus cloud types. The correlations are strongest for maritime regions at midlatitudes and are poor in areas dominated by other phenomena such as the El Niño-Southern Oscillation (ENSO). The correlations are compelling, especially with the higher energy of cosmic rays (but not so high in energy that the amplitude of the solar cycle variation becomes negligible). The correlation coefficient for 12-month smoothed data exceeds 0.85.

Several more cycles of data are required before we can be confident that the solar cycle variation of cloud cover is not just a chance occurrence. The mechanism or mechanisms that could result in a correlation between cosmic ray fluxes and the global cloud cover are highly controversial and are certainly not yet understood, although clear possibilities are emerging [*Svensmark*, 2000, this volume]. If present, such a mechanism would mean that long-term changes in the heliospheric field were of great importance.

1.3. Geomagnetic Activity

As well as modulating the solar irradiance and the heliospheric field (discussed above in sections 1.1 and 1.2), solar activity causes geomagnetic activity through energy extracted from the solar wind flow by Earth's magnetosphere. This has been monitored in a continuous and homogeneous manner since 1868 using antipodal observing stations in southern England and Australia. The level of activity was quantified by *Mayaud* [1972], who developed the *aa* index, using the range of the fluctuations of the horizontal component of the field observed in 3-hour intervals at these stations. The *aa* index was designed to be very well correlated with other planetary indices of geomagnetic activity; for example, the variations in annual means of *aa* and the planetary *Ap* index are almost identical since the start of the *Ap* data series in 1932 [*Mayaud,* 1972; *Ahluwalia,* 1997]. This is true for both the solar cycle and the longer-term variations. The long-term changes in *aa* are also highly correlated with related phenomena, such as the occurrence of low-latitude aurorae [*Pulkkinen et al.*, 2000]. The *aa* data sequence shows a systematic long-term rise in geomagnetic activity since 1900, a variation which shows some intriguing similarities to the observed rise in global mean surface temperatures [*Cliver et al.*, 1998b]. *Clilverd et al.* [1998] have analyzed in detail a variety of potential causes for this rise in *aa* and eliminated most of them. For example, the long-term drift of the geomagnetic field has caused a systematic shift in the magnetic latitudes of the stations: The Northern Hemisphere station used to generate *aa* has drifted about 4° equatorward in magnetic coordinates since 1868, whereas that in the Southern Hemisphere has drifted about 2° poleward. However, the data show that the activity at both stations has risen in an almost identical manner and thus the rise is not caused by this effect. *Stamper et al.* [1999] studied interplanetary data from solar cycles 20, 21, and 22 (1963-1996) in a search for the cause of the rise in *aa*. They used the theory of solar wind power extraction by Earth's magnetosphere [*Vasyluinas et al.*, 1982] to show that the largest of several contributions to the long-term change in *aa* was a rise in the magnitude of the interplanetary magnetic field around the Earth. *Lockwood et al.* [1999a] also used this theory to, effectively, separate the effects of nonrecurrent geomagnetic activity (for example, induced by coronal mass ejections) and recurrent geomagnetic activity (due to fast solar wind streams) and thereby developed a procedure to compute the magnitude of the interplanetary field at Earth B_E from the *aa* data.

1.4. Solar Cycle Length

There is another important indicator of long-term solar change derived from a continuous measurement sequence. This is the solar cycle length *L*, derived from the sunspot numbers *R*. *Gleissberg* [1944] noted that there was a long-term (of order 80-100 years) quasi-periodic behavior in *R* and that this was related to *L*. He employed a long-timescale secular filter to smooth the values of *L*, which were determined from the times of adjacent minima and maxima of *R*. In recent years a variety of other methods to determine *L* have been developed. For example: *Hoyt and Schatten* [1993] determined from each annual mean of *R* the fraction of the cycle that had been completed and then measured the delay to the same point in the next cycle; *Fligge et al.* [1999] have employed Morlet wavelets; *Mursula and Ulich* [1998] have used median activity times; *Friis-Christensen and Lassen* [1991] used a method and filter that were similar to Gleissberg's; and *Lockwood* [2000] has used an autocorrelation technique.

These various techniques produce variations in *L* that are similar but with distinct differences. The variation found by *Friis-Christensen and Lassen* [1991] anticorrelates both with global average temperatures on Earth and the *aa* geomagnetic index. This is because it falls more gradually over the last 100 years than, for example, the variations deduced by *Lockwood* [2000] and *Fligge and Solanki* [1998] that show more step-like falls, followed by periods of more constant L.

The irradiance constructions of *Hoyt and Schatten* [1993] and *Solanki and Fligge* [1998; 1999] depend on L. These reconstructions are quite similar to those from other methods, and some possible reasons why L is a valuable proxy are discussed in section 4.

1.5. Coronal Source Flux

The coronal source surface is where the magnetic field becomes approximately radial. It is a roughly spherical surface at a heliocentric distance r of about $2.5R_s$ (where R_s is a solar radius; that is, the photosphere is at $r = 1R_s$). This surface can be considered as the boundary that separates the solar corona from the heliosphere. The magnetic flux threading the coronal source surface is called the open solar flux or the coronal source flux, F_s.

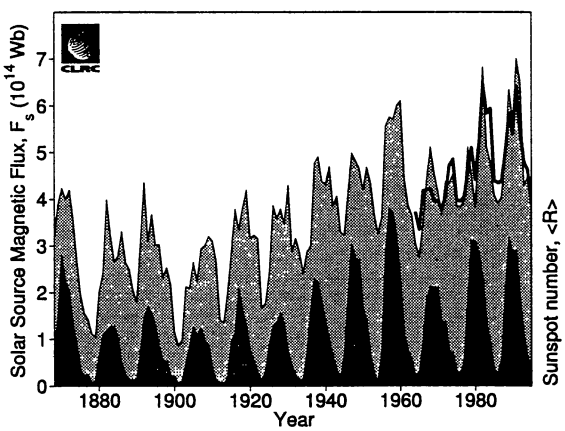

Figure 1: Variation of annual means of the coronal source flux F_s as derived from the aa index by the method of Lockwood et al. [1999a] (thin line bounding grey shaded area) and from interplanetary measurements of the radial component of the interplanetary magnetic field near Earth (thick line). The dark area gives the variation of the smoothed sunspot number.

On annual timescales, the interplanetary magnetic field (IMF, the heliospheric field in the ecliptic plane), obeys the Parker spiral orientation with a garden hose angle γ that is approximately constant at a given r [*Gazis*, 1996; *Stamper et al.*, 1999]. Thus the IMF's radial component near Earth $B_{rE} = B_E \cos(\gamma)$ can also be estimated from the *aa* geomagnetic activity data using the method of *Lockwood et al.* [1999a]. This has significance for the heliosphere away from the ecliptic plane because the Ulysses spacecraft has shown that sheet, rather than volume, currents dominate, to the extent that latitudinal gradients of the average radial heliospheric field are small [*Balogh et al.*, 1995; *Lockwood et al.*, 1999b]. This means that the radial field B_r, at a heliocentric distance r and at any latitude, is approximately equal to $(r/R_1)^2 B_{rE}$, where $R_1 = 1$ AU. Because, on average, the Parker spiral theory of the heliospheric field applies and because this does not predict significant flux crossing the current sheet at $r < R_1$, the total magnetic flux leaving the Sun and entering the heliosphere (the coronal source flux) is $F_s = (1/2)(4\pi R_1^2)<|B_{rE}|>$. The method of *Lockwood et al.* [1999a] for deriving annual means of F_s from the *aa* index assumes that both Parker spiral theory and the uniformity of B_r were valid at all times since 1868. The method was developed using data from solar cycles 21 and 22 and then tested against independent interplanetary measurements from cycle 20 [*Lockwood and Stamper*, 1999]. Their results are shown in figure 1. The lightly shaded area is the coronal source flux estimated from the *aa* index; the thick line is the corresponding value derived from the mean radial field $<|B_{rE}|>$ measured by spacecraft near $r = R_1$. The area shaded black gives the variation of smoothed sunspot number for comparison. Figure 1 shows that the average F_s has risen by 34% since 1964 and by 140% since 1900.

2. LONG TERM CHANGES IN THE PHOTOSPHERE

The changes in the total open flux of the sun relate specifically at the source surface ($r \approx 2.5R_s$) and could result from changes that are restricted to the corona ($1R_s < r \leq 2.5R_s$), with little implication for the total flux emerged through the photosphere (which contains both open and closed solar flux and is of order is of order $10F_s$ [*Wang et al.*, 2000b]), nor possibly even for the distribution of that flux over the solar surface. As discussed above, changes in the total solar irradiance, on timescales of 20 years or less, are well explained by magnetic phenomena in and below the solar photosphere ($r \approx 1R_s$) and so the implications for irradiance of the drift in open flux are not immediately clear. However, *Lockwood and Stamper* [1999] obtained a correlation of the coronal source flux F_s with solar irradiance measurements of 0.84. In fact, they correlated the data from the various monitors separately and used the regression fits to intercalibrate the instruments. If the intercalibration of the instruments by *Fröhlich and Lean* [1998] is adopted, this correlation falls to 0.79. From this correlation, Lockwood and Stamper derived a long-term irradiance variation that is strikingly similar to that of *Lean et al.* [1995]. This agreement is even closer for the revised reconstruction presented recently by *Lean* [2000]. Lean used a method which employs sunspot numbers added to a long-term drift with a waveform given by 11-year running means of sunspot numbers, and with an amplitude based on a comparison of the Maunder minimum and noncyclic Sun-like stars. The extrapolation by Lockwood and Stamper is also quite similar to the reconstruction by *Solanki and Fligge* [1998], who add contributions by the active regions (based on sunspot numbers, R) to a quiet sun (network) variation (based on cycle length, L).

The possibility that F_s could be used as a valid single proxy for irradiance variation reconstruction (as opposed to a composite of added terms) is also suggested by the fact that the changes in F_s, as shown in Figure 1, do have some marked similarities to some changes seen in the solar photosphere; thus they appear to reflect changes in the magnetic field there and in the subsurface layers. Figure 1 shows that the peak sunspot number R at the maximum of the solar cycles has risen in association with F_s. Figures 2a and 2b show that this is also true for have the solar-cycle averages of sunspot number R_{11} [*Lockwood et al.,* 1999b]. Figure 2c shows that there is also some similarity to the variation of solar cycle length, L, as deduced by the autocorrelation technique of *Lockwood* [2000]. (To aid comparison, Figure 2c shows {12.5 - L}, note that 12.5 years is merely a convenient reference value and has no significance). In addition, careful inspection reveals that there has been a small upward drift of the minima in R (figure 1).

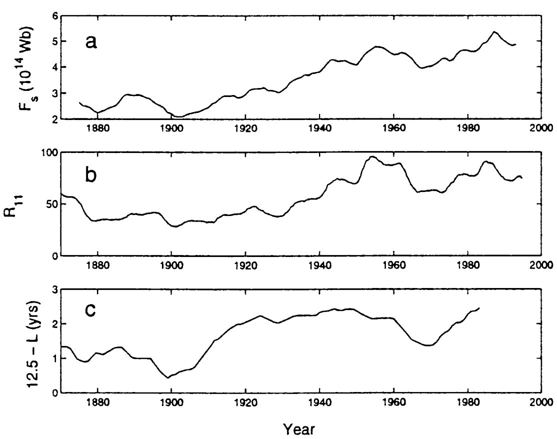

Figure 2: (a) The 11-year running means of the coronal source flux F_s, derived from the aa index of geomagnetic activity (see Figure 1); (b) 11-year running means of the sunspot number, R_{11}; (c) (12.5 – L), where L is the length of the solar cycle in years and is derived from autocorrelation functions of sunspot number R, using the method of Lockwood [2000].

Foster and Lockwood [2000] have shown that there is also a strong correlation between annual means of F_s and the standard deviations of sunspot latitudes. They employed the observations of sunspot groups that were made at Greenwich between 1874 and 1981 and extended this sequence using the observations made in a compatible way at Mt. Wilson after 1967. Figure 3 shows that the spread of sunspot latitudes increases with a very similar waveform to F_s, although the long-term (~100-year) drift is a smaller fraction of the solar cycle variation than it is for F_s. In other words, the photosphere shows more spots, spread over a greater area of its surface, when F_s is higher. Thus the variations in F_s do appear to be a symptom of changes in the surface field. These findings indicate that there have been considerable changes in the processes that generate and distribute photospheric and coronal magnetic fields.

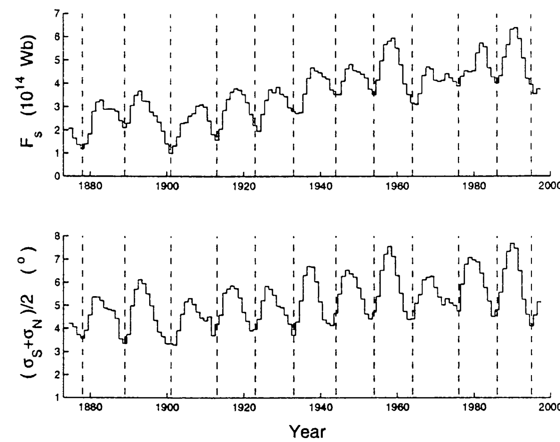

Figure 3: (a) The variation of annual means of the coronal source flux, F_s, as shown in Figure 1, compared with (b) the average of the standard deviations of sunspot group latitudes, $(\sigma_N+\sigma_S)/2$, as deduced by Foster and Lockwood [2000] from Greenwich and Mount Wilson sunspot group observations. σ_N and σ_S are the standard deviations for the Northern and Southern solar hemispheres, respectively, and are effectively the width of the wings in the butterfly diagram.

The magnetic field at the photosphere includes both open flux (that threads the coronal source surface) and the larger closed flux (that does not) and is generated by dynamo processes in the solar interior, in particular at the base of the convection zone (at around $r = 0.7R_s$). This field emerges in active regions (as bipolar magnetic regions, or BMRs) [*Harvey and Zwaan*, 1993] at a total rate which increases with the number of sunspots. Most of the flux associated with BMRs is annihilated by diffusion toward the neutral line between regions of opposite field polarity, but some is added to the fields associated with the supergranulation network. This field is transported over the solar surface by the differential rotation in the outer Sun, by convection associated with supergranules, and by meridional poleward flow in the surface layers. As the centers of the two polarity regions of a BMR separate in the differential rotation, the magnetic field loops rise through the coronal source surface, and the coronal source flux increases. Open flux (that threads the coronal source surface) accumulates near the poles, where it forms the large coronal holes seen at the subsequent sunspot minimum [*Wang et al.*, 2000a].

Recent theoretical modelling by *Solanki et al.* [2000] predicts a rôle for both cycle length and emergence rate in the long-term variation of coronal source flux shown in Figure 1. These authors used the sunspot number to quantify E, the rate of flux emergence through the photosphere in BMRs. The basic equation is the continuity of open network flux:

$$\mathrm{d}F_s/\mathrm{d}t = \gamma E - F_s/\tau_N \quad (1)$$

where E is the rate of flux emergence through the photospheric active regions, τ_N is the time constant for destruction of open network flux and $\gamma = (1 + \tau_t/\tau_a)^{-1}$ and thus depends on the time constants for flux annihilation in active regions τ_a and for transfer of flux from active regions to the network τ_t. *Solanki et al.* [2000] estimate γ to be 0.015 and derive a best fit with a time constant τ_N of 4 years. Because this τ_N is sufficiently large, the cycle length L will influence the amount of residual open network flux because shorter cycles will mean that this decay does not progress as far as during longer cycles. This prediction is confirmed by Figure 4, which contrasts the variation of $\{12.5 - L\}$ (bottom graph) with the rate of increase in the coronal source flux, $\mathrm{d}F_s/\mathrm{d}t$ from the variation of F_s shown in figure 1 (top graph). Because 20-year intervals were used to evaluate

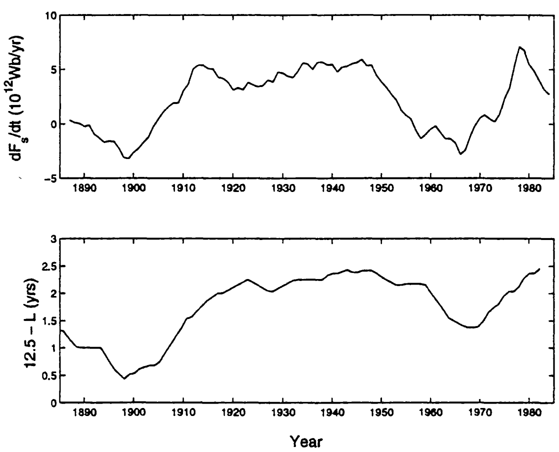

Figure 4: (top) Rate of increase in the coronal source flux (dF_s/dt). Running means over 20-year intervals are presented. (bottom) The variable {12.5 - L}, where L is the length of the solar cycle, in years, derived from the autocorrelation function of sunspot number R using 20-year intervals (as in Figure 2c).

L, running means of $\mathrm{d}F_s/\mathrm{d}t$ over 20-year intervals are presented. A clear similarity is evident. In particular, when L is a maximum (minima in $\{12.5 - L\}$), $\mathrm{d}F_s/\mathrm{d}t$ is negative. Some lag is present in Figure 4, but considering that these are smoothed data sets, this is probably not significant. This relationship between the cycle length L and $\mathrm{d}F_s/\mathrm{d}t$ was inherent in the modelling of the F_s data sequence by *Solanki et al.* [2000]. Thus Figure 4 confirms their conclusion that cycle length L and the rate of flux emergence (related to sunspot number R) in active regions are both key elements in the long-term accumulation of open solar flux (when average sunspot numbers R_{11} are high and/or when cycle lengths L are low) and its decay (when R_{11} is low and/or when cycle lengths L are large).

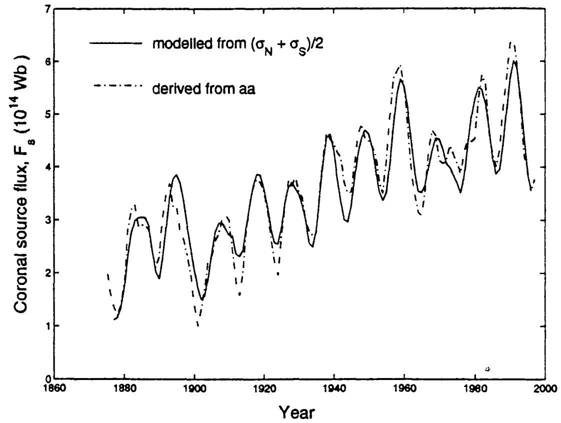

Figure 5: The variations of the coronal source flux F_s, derived from geomagnetic activity by Lockwood et al. [1999a] and as shown in Figure 1 (dot-dash line); and as predicted using the model of Solanki et al. [2000], as adapted by Foster and Lockwood [2000] (solid line).

Figure 5 shows the observed variation of F_s and a best-fit model prediction using the Solanki et al. model and equation (1). In this plot, the emergence rate E has been quantified in a slightly different way, using the spread of sunspot latitudes (figure 3) and the best-fit τ_N is 3.8 years. [*Foster and Lockwood*, 2000], but the result is almost exactly the same as presented by *Solanki et al.* [2000]. It can be seen that the agreement is excellent. Thus we have a new understanding of how variations in the flux emergence rate and solar cycle length have produced the observed long-term drift in the coronal source flux.

3. LONG-TERM VARIATIONS IN COSMIC RAYS

Because of their large gyroradii and complex paths, cosmic rays sample much greater regions of the heliosphere than do near-Earth satellites. Thus we would expect a strong anticorrelation of the coronal source flux F_s with cosmic ray fluxes. The products of galactic cosmic ray bombardment of the atmosphere have been detected using neutron monitors continuously since 1953. Cosmic rays are also shielded by the geomagnetic field, such that the range of energies studied depends on the latitude of the observing station. For example, for the mid-latitude station at Climax, USA, primary cosmic ray particles of energy exceeding 3 GeV can be detected, whereas for the lower-latitude stations at Huancayo, Peru and Haleakala, Hawaii the greater shielding by the geomagnetic field means that only primaries of energy exceeding 13 GeV can be detected. The two stations at Huancayo and Hawaii provide a homogeneous data sequence, the data series being continued at Hawaii, after monitoring ceased at Huancayo in 1993.

Lockwood [2000] has shown that there is a strong anti-correlation between the fluxes of cosmic rays and the coronal source F_s deduced from the *aa* geomagnetic index by *Lockwood et al.* [1999a]. The correlation coefficients for count rates at Climax (C) and Huancayo/Hawaii (H) are $c_C = -0.874$ and $c_H = -0.897$. The best-fit, least squares linear regressions are $C/10^3 = 5.22 - 0.278F_s$ and $H/10^3 = 1.866 - 0.345F_s$. The correlations show that $c_C^2 = 76.4\%$ and $c_H^2 = 77.3\%$ of the variations in the cosmic ray fluxes seen at Climax and Huancayo/Hawaii, are explained by the strength of the heliospheric field alone. Other factors, such as drifts that depend on the polarity of the heliospheric field [*Ahluwalia and Wilson*, 1996; *Usoskin et al.*, 1998] and the solar wind flow [*Sabbah*, 2000], collectively contribute a maximum of $(1 - c_C^2) = 23.6\%$ and $(1 - c_H^2) = 22.7\%$ to the variations.

Using the high correlations and the corresponding least-squares regression fits between C and F_s and between H and F_s, along with the data sequence of F_s shown in Figure 1, the cosmic ray fluxes can be extrapolated back to 1868. The results are shown in Figure 6, the solid line being for the Climax data (>3 GeV), and the dashed line being for the Huancayo/Hawaii data (>13 GeV). In both cases, the count rates have been normalised to the average value seen during solar cycle 21. The plot indicates that the average fluxes of cosmic rays above 3 GeV were approximately 15% higher in 1900 than they are now, whereas the fluxes at energies above 13 GeV were higher by about 4%.

Also shown in Figure 6 is the variation of cosmic ray fluxes since 1937 deduced from a collection of ionisation chambers at high latitudes, and near sea level [*Ahluwalia*, 1997] (solid line joining dots). The sites employed are Cheltenham (1937-1956), Fredericksberg (1956-1972) and Yakutsk (1954-1994). The geomagnetic energy cutoffs at these stations are 2.2 GeV, 2.2 GeV and 1.7 GeV, respectively, but fluxes are limited by the higher atmospheric cutoff of about 4 GeV in each case. The muons detected relate to somewhat higher energy primary cosmic rays than for the neutron monitors discussed above, and the median energy observed by these ionisation chambers is 67 GeV. To provide a single sequence, considerable intercalibration factors have been applied to the data from these ionisation chambers [*Ahluwalia*, 1997], and thus long-term drifts may not be accurately represented. Of particular concern is that so little overlap in data (1954-1956) exists for full annual means from Cheltenham and Yakutsk. Nevertheless, there are strong similarities for the variation for >13 GeV predicted from the F_s values (the correlation coefficient is 0.811). In terms of the percentage change, both the solar cycle variations and the longer-term drifts seen by the ionisation chambers are smaller (indicating that they are responding to higher-energy primary cosmic rays that are less influenced by the heliospheric shield). Analysis of global carbon-14 production [*Struiver and Quay*, 1980] implies that the combined ionisation chamber data may underestimate a slight downward drift in the cosmic ray fluxes in the period 1937-1970 [*O'Brien*, 1979].

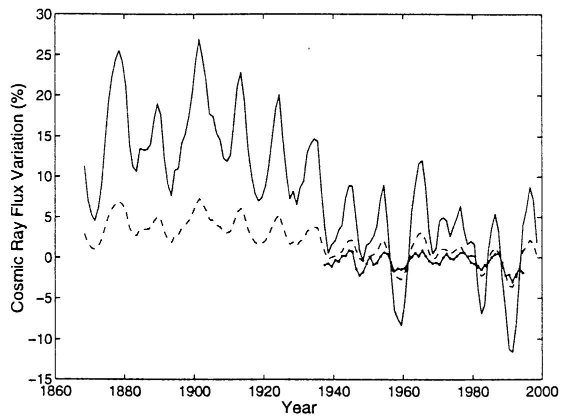

Figure 6: Inferred variation of cosmic ray fluxes since 1868. The solid line is an extrapolation using F_s, based on the correlation with counts C by the cosmic ray neutron detector at Climax (geomagnetic energy cutoff of 3 GeV). The dashed line is based on the correlation with the counts H by the cosmic ray neutron detectors at Huancayo/Hawaii (geomagnetic energy cutoff 13 GeV). The line joining dots is the variation deduced from three ionization chambers by Ahluwalia [1997], which respond to a median energy of 67 GeV. In all cases, the variation is relative to the average value seen during solar cycle 21. After Lockwood [2000].

The high anticorrelations between cosmic ray fluxes and the coronal source flux imply that there should be a strong anti-correlation between F_s and the abundance of the ^{10}Be isotope in ice cores, produced by cosmic ray bombardment. The ^{10}Be isotope is formed as a spallation product in the upper atmosphere when galactic cosmic rays impact oxygen and nitrogen nuclei. This isotope is deposited in ice sheets by precipitation over a subsequent extended period. Thus a lag is expected. *Lockwood* [2000] found a strong anticorrelation of F_s with the ^{10}Be isotope abundance (derived from the Dye-3 Greenland ice core [*Beer et al.,* 1998]) for annual means from 1868-1985. The peak correlation coefficient is at a lag of 1 year and is –0.64. The best-fit, least squares linear regression is $[^{10}\text{Be in atoms g}^{-1}] = 10^4 (1.323 - 0.133F_s)$. Figure 7 compares the extrapolated variation of cosmic ray flux (at >13 GeV) since 1868 (solid line, as shown in Figure 6), with the ^{10}Be isotope record, scaled using the regressions given above. In both cases, the variation is relative to the average value seen during solar cycle 21. The long-term trend is reproduced in both data sets although the solar cycle variations are not identical. This is consistent with the results of *Fligge et al.* [1999], who found that cycle lengths derived from the ^{10}Be isotope data were not fully reliable, presumably because the extended response

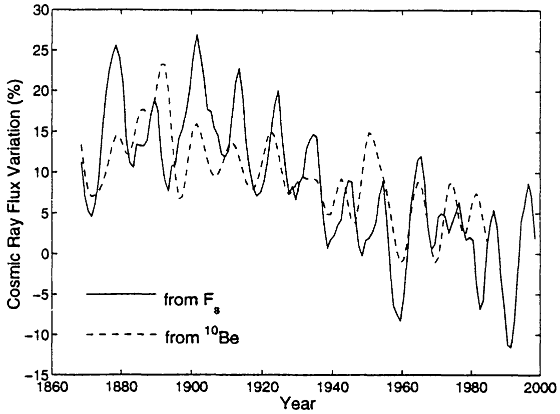

Figure 7: Variation of cosmic ray flux since 1868. The solid line is an extrapolation using F_s, based on the correlation with fluxes at >3 GeV shown in Figure 6. The dashed line is based on the ^{10}Be isotope record, scaled using the correlations with F_s. In both cases, the variation is relative to the average value seen during solar cycle 21.

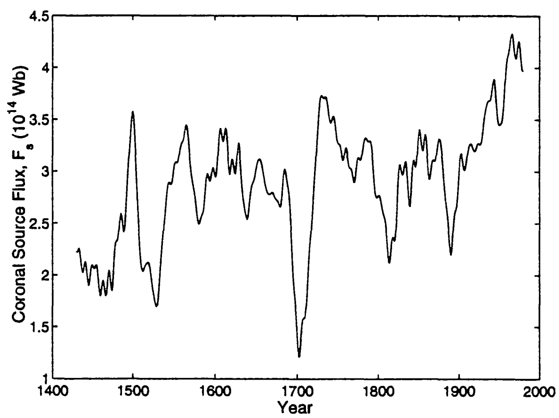

Figure 8. Extrapolated coronal source flux F_s from the ^{10}Be isotope data, using the regression analysis of data since 1868.

function is variable (due to a spread of isotope deposition timescales caused by climate variability) and the difficulties in dating the ice.

We can use the full data sequence of the ^{10}Be isotope data, along with the regression with F_s for data since 1868, to estimate the coronal source flux back to near 1400. This extrapolation is shown in Figure 8: These data are 11-year running means in which the solar cycles have been smoothed out. Comparison with Figure 2 shows that the drift in average F_s since 1868 is very well reproduced. The plot shows that variations in F_s seen since 1900 are similar in amplitude to those seen at prior times. Particularly rapid changes are seen late in the "Maunder minimum" of sunspot activity (roughly 1645-1715) [e.g., *Cliver et al.*, 1998b], with F_s falling to its lowest value of 1.2×10^{14} Wb near 1700, toward the end of this period. This corresponds to a value of about one quarter of present-day values. That F_s should reach a minimum at the end, rather than in the middle, of the Maunder minimum is consistent with the model of *Solanki et al.* [2000]. This is because the F_s would decay throughout the minimum because the flux emergence rate in active regions was exceptionally low.

Other studies similarly confirm the drift of F_s discussed here. For example, *Bonino et al.* [1995] use the Titanium isotope ^{44}Ti, produced in meteorites by galactic cosmic rays and having a half life of about 96 years, to deduce a rise in the heliospheric field throughout the 20th century: the rise found is four times larger than would be deduced from sunspot numbers using the anti-correlation with cosmic ray fluxes seen since 1953.

4. CONCLUSIONS AND IMPLICATIONS

Recent work has revealed that there have been important long-term (here meaning ~100 year) changes in the magnetic field in the solar atmosphere. Some of this field permeates the coronal source surface and enters the heliosphere. Because of the work by *Solanki et al.* [2000], an understanding of these changes is becoming available to us, in terms of the emergence of flux in active regions, its transfer to the network, and the balance between loss of open network flux and its accumulation in coronal holes. In particular, this gives us an understanding of why the length of the solar cycle has important implications. Shorter solar cycles facilitate a rise in the coronal source flux; longer cycles allow it to decay. However, the accumulation of the coronal source flux is also strongly dependent on the rate of flux emergence in active regions. In general, the peak and cycle-averaged sunspot numbers are larger when cycles are shorter. Thus shorter cycles can also be associated with larger flux emergence rates and there is less time available for the open flux to decay. Together, these effects mean that the net effect is that shorter cycles correlate with increasing coronal source flux.

The change in the coronal source flux was revealed by studies of the long-term drift in geomagnetic activity [*Lockwood et al.*, 1999a; b]. Because this work was based on the quantitative theory of solar wind-magnetosphere coupling, the change in the open solar flux has been quantified and found to be considerable (a factor of 2.4 change since 1900). The variation is confirmed by studies of isotopes produced by cosmic ray bombardment in ice cores and meteorites. Cross-correlation allows us to calibrate these changes and so extrapolate the solar changes further back in time. For example, the ^{10}Be isotope in ice cores shows us that by the end of the Maunder minimum the coronal source flux had fallen to about a quarter of present-day values.

In addition to influencing the transfer of energy from the solar wind to the magnetosphere, the rise in F_s will have caused the cosmic ray flux incident on the Earth to have fallen. In this paper, we have estimated that the cosmic ray fluxes above 3 GeV were 15% higher, on average, around 1900 than they are now. The corresponding figure for >13 GeV particles is about 4%. The potential implications of this are not yet understood.

4.1. Potential Effects of Cosmic Ray Variations

Cosmic rays generate air ions in the subionospheric gap which allows current to flow in the global electric circuit. This connects thunderclouds with the ground via lightening, by closing the loop via the ionosphere [*Bering et al.*, 1998]. It is not yet known what sort of modulation to this circuit is caused by the changes in cosmic ray fluxes, nor what influence this might have.

The apparent correlation between cosmic rays and global cloud cover will remain just that until we have sufficient data to confirm or disprove its significance. That having been said, the strong correlation for the global low-altitude cloud cover, [*Marsh and Svensmark*, 2000] is very self-consistent and interesting and would be very important indeed were it to reflect a genuine and active mechanism. Potential mechanisms that might explain how cosmic ray products grow in size to allow water droplets to condense are now beginning to emerge. Given that we now know that cosmic ray fluxes have declined systematically over the past 100 years with the increasing heliospheric field, it is now very important to investigate these mechanisms, both experimentally and theoretically.

4.2. Solar Irradiance Reconstructions

Lockwood and Stamper [1999] found a correlation of the coronal source flux F_s with solar irradiance measurements and used this to generate a reconstruction of past variation of irradiance. Figure 9 contrast this to other reconstructions of the irradiance. A fair degree of agreement can be seen between the various variations inferred, despite the fact that they are based on different assumptions and proxies. Figure 2 offers a simple practical reason as to why these solar irradiance reconstructions are similar. The coronal source flux F_s , which was used as a proxy for total irradiance by *Lockwood and Stamper* [1999], is highly correlated with the 11-year running means of sunspot number, which *Lean et al.* [1995] and *Lean* [2000] use to give the waveform long-term drift. The coronal source flux is also related to the cycle length (Figures 2 and 4), which was used as a proxy for total irradiance by *Hoyt and Schatten* [1993] and to give the quiet sun variation by *Solanki and Fligge* [1998]. Thus these methods are not as independent as they initially seem, in terms of the waveform of the irradiance variation that they predict on 100-year timescales. However, the amplitude of the long-term drifts derived are also similar, and this is not derived the same way for the three cases. By linking emergence rate through the photosphere in active regions (covered by sunspots and faculae), solar cycle length and the coronal source flux, the work of *Solanki et al.* [2000] may offer the origins of a theoretical explanation of the agreement between the reconstructions of Lean et al. and Solanki and Fligge with the extrapolation by Lockwood and Stamper.

4.3. The Effects of inferred Solar Irradiance Variations

The concept that significant climate change is caused by

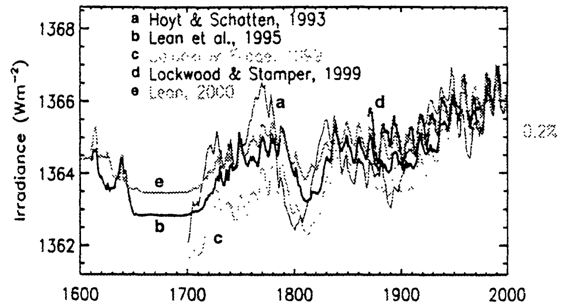

Figure 9: A comparison of the various solar irradiance reconstructions (courtesy Judith Lean).

changes in the solar irradiance is certainly not new [e.g. *Blanford*, 1891]. The observed solar cycle variation in irradiance is small, S varying between 1367.0 Wm^{-2} and 1368.3 Wm^{-2}, a variation of just under 0.1%. The most conservative of the estimates of the drift in the 11-year running mean of the irradiance, S_{11}, are by *Lockwood and Stamper* [1999] and *Lean* [2000] for which ΔS_{11} = 1.65 ± 0.23 Wm^{-2} over the last century, a rise of 0.12 ± 0.02%. Allowing for Earth's albedo, a, and spherical geometry this gives a change in the radiative forcing at the top of the atmosphere of $\Delta Q = \Delta S_{11}(1-a)/4 \approx$ 0.29±0.04 Wm^{-2}, where a is here taken to be 0.3 [see discussion of the concept of Q by *Hansen et al.*, 1997]. The effect of any change in S on global mean surface temperatures will be complex because it will be made up of contributions that are much stronger at some wavelengths than at others and because a variety of other effects (for example, changes in anthropogenic greenhouse gases, tropospheric sulphate aerosols and volcanic dust in the stratosphere, ozone absorption of UV etc.) will also be active and will interact with each other in complex feedback loops [*Rind and Overpeck*, 1993; *Hansen et al.*, 1997]. This estimate of ΔQ due to solar change is similar to the estimate in the last IPCC (Intergovernmental Panel on Climate Change). By way of comparison, the IPCC's ΔQ estimates for the same interval due to CO_2, other greenhouse gasses, and aerosols are roughly 1.5 Wm^{-2}, 1.1 Wm^{-2}, and –1.3 Wm^{-2}, respectively [e.g. *Wigley et al.*, 1997].

In order to make a simple evaluation of the effect of the solar irradiance change, we can adopt an estimate of the "climate sensitivity", dT/dQ, where T is the global mean of the surface temperature. Using an estimate of the climate sensitivity of dT/dQ = 0.85 ± 0.15°C/Wm^{-2} [*Rind and Overpeck*, 1993], we predict a temperature rise of ΔT = -0.16+(0.147±0.026)×([S in Wm^{-2}]-1365). This represents a consensus view of the Earth's climate sensitivity, being the average of values from several large numerical models of the coupled atmosphere-ocean circulation system, with an uncertainty set by the range of the estimates. From this we infer that the sun's brightness change, on its own, could have caused a temperature rise of 0.24±0.04°C since 1900.

Lockwood et al. [1999b] illustrate how the inferred variations from irradiance changes, ΔT, are highly correlated with the observed global average of the surface temperature, ΔT_o [e.g. *Parker et al.*, 1994] the correlation coefficient being 0.93 and at a lag of $\delta t = 2$ years, which is consistent with the heat storage effect of the oceans [*Wigley and Raper*, 1990]. However, great care must be taken not to over-interpret this correlation. It would imply a climate sensitivity of dT/dQ = 2.2°C/Wm^{-2}, in order to explain the observed temperature rise ΔT_o in terms of solar irradiance variations alone. This is a higher value than any of the published estimates from modelling studies (one of the largest is dT/dQ = 1.7°C/Wm^{-2}, by *Nesme-Ribes et al.* [1993]) and roughly twice the consensus value. Multi-variable analysis that accounts for several mechanisms, including anthropogenic effects, natural climate variability and volcanoes does reveal correlations are improved if some solar drift is included [*Wigley et al.*, 1997; *Tett et al.*, 1997].

Lockwood et al. [1999b] and *Lean et al.* [1995] compared the inferred temperature rise ΔT (for the consensus prediction of the climate sensitivity) with that observed ΔT_o we find no significant difference for the period 1870-1910. On the other hand, the change in solar luminosity alone can account for only about 50% of the rise in ΔT_o over the period 1910-1960 but leas than 31% of the rapid rise in ΔT_o over 1970-present. In the same interval, industrially-produced CO_2 in the atmosphere increased from about 280 to 355 ppmv. The implications are that that the onset of an man-made contribution to global warming was disguised by the rise in the solar irradiance and that the anthropogenic effect may have a later, but steeper, onset than previously thought. Such an effect is consistent with the predictions for combined greenhouse and aerosol pollutants [*Wigley et al.*, 1997; *Hansen et al.*, 1997]. Recently, *Tett et al.* [1999] have used a set of simulations made by a coupled atmosphere-ocean global circulation model to deduce a shift from solar forcing to anthropogenic effects as this century has progressed.

Acknowledgments: The author is grateful for discussions with a number of scientists, in particular: S. Solanki, N. Marsh, H. Svensmark, Y.-M. Wang, G. Cini Castognoli and J. Lean. He is also grateful to J. Beer, who supplied the ^{10}Be isotope ice core data, and to the World Data Center system. This work was supported by the U.K. Particle Physics and Astronomy Research Council.

REFERENCES

Ahluwalia, H.S., Galactic cosmic ray intensity variations at a high-latitude sea-level site 1937-1994, *J. Geophys. Res.*, *102*, 24,229-24,236, 1997

Ahluwalia, H.S., and M.D. Wilson, Present status of the recovery phase of cosmic ray 11-year modulation, *J. Geophys. Res.*, *101*, 4879-4883, 1996.

Balogh, A., E.J. Smith, B.T. Tsurutani, D.J. Southwood, R.J. Forsyth, and T.S. Horbury, The heliospheric field over the south polar region of the Sun, *Science, 268*, 1007-1010, 1995.

Beer, J., S. Tobias, and N. Weiss, An active Sun throughout the Maunder minimum, *Sol. Phys., 181*, 237-249, 1998.

Bering, E.A., A.A. Few, and J.R. Benbrook, The global electric circuit, *Phys. Today, 51*, 24-30, 1998.

Blanford, H.F., The paradox of the sunspot cycle in meteorology, *Nature, 43*, 583-587 (1891)

Bonino, G., G. Cini Castognoli, N. Bhabdari and C, Taricco, Behavior of the heliosphere over prolonged solar quiet periods by ^{44}Ti measurements in meteorites, *Science, 270,* 1648-1650.

Cane, H.V., G. Wibberenz, I.G. Richardson, and T.T. von Rosenvinge, Cosmic ray modulation and the solar magnetic field, *Geophys.Res.Lett., 26*, 565-568, 1999.

Chapman, G.A., A.M. Cookson, and J.J. Dobias, Solar variability and the relation of facular to sunspot areas during cycle 22, *Astrophys. J.*, 842, 541-545, 1997.

Clilverd, M.A., T.D.G. Clark, E. Clarke, and H. Rishbeth, Increased magnetic storm activity from 1868 to 1995, *J. Atmos. Sol.-Terr. Phys., 60*, 1047-1056, 1998.

Cliver, E.W., V. Boriakoff, and K. Bounar, Geomagnetic activity and the solar wind during the Maunder minimum, *Geophys. Res. Lett., 25,* 897-900, 1998a.

Cliver, E.W., V. Boriakoff, and J. Feynman, Solar variability and climate change: Geomagnetic activity and global surface temperature, *Geophys. Res. Lett., 25*, 1035-1038, 1998b.

Fligge, M., S.K. Solanki, Y.C. Unruh, C. Fröhlich and C. Wehrli, A model of solar total and spectral irradiance variations, *Astron. Astrophys., 335*, 709-718, 1998.

Fligge, M., S.K. Solanki, and J. Beer, Determination of solar cycle length using the continuous wavelet transform, *Astron. Astrophys, 346*, 313, 1999.

Foster, S. and M. Lockwood, Long-Term Changes in the Solar Photosphere Associated with Changes in the Coronal Source Flux, *Geophys. Res. Lett.,* submitted, 2000

Friis-Christensen, E., and K. Lassen, The length of the solar cycle, an indicator of solar activity closely associated with climate, *Science, 192*, 1189-1202, 1991.

Fröhlich, C., and J. Lean, The Sun's total irradiance: Cycles, trends and related climate change uncertainties since 1976, *Geophys. Res. Lett., 25*, 4377-4380, 1998.

Gazis, P.R., Solar cycle variation of the heliosphere, *Rev. Geophys., 34,* 379-402, 1996.

Gleissberg, W., A table of secular variations of the solar cycle, *J. Geophys. Res., 49,* 243-244, 1944.

Hansen, J., M. Sato, and R. Ruedy, Radiative forcing and climate response, *J. Geophys. Res., 102*, 6831-6864, 1997.

Harvey, K.L., and C. Zwaan, Properties and emergence of bipolar active regions, *Sol.* Phys.,148, 85-118, 1993.

Hoyt, D., and K. Schatten, A discussion of plausible solar irradiance variations 1700-1992, 18895-18906, 1993.

Kristjánsson, J.E., and J. Kristiansen, Is there a cosmic ray signal in recent variations in global cloudiness and cloud radiative forcing?, *J. Geophys. Res., 105*, 11,851-11,863, 2000.

Lean, J., Evolution of the Sun's spectral irradiance since the Maunder minimum, *Geophys. Res. Lett., in press*, 2000.

Lean, J., J. Beer, and R. Bradley, Reconstruction of solar irradiance since 1610: Implications for climate change, *Geophys. Res. Lett., 22*, 3195-3198, 1995.

Lockwood, M., Long-Term Variations in the Magnetic Fields of the Sun and the Heliosphere: their origin, effects and implications, *J. Geophys. Res.*, in press, 2000

Lockwood, M., and R. Stamper, Long-term drift of the coronal source magnetic flux and the total solar irradiance, *Geophys Res. Lett., 26*, 2461-2464, 1999.

Lockwood, M., R. Stamper, and M.N. Wild, A doubling of the sun's coronal magnetic field during the last 100 years, *Nature*, 399, 437-439, 1999a.

Lockwood, M., R. Stamper, M.N. Wild, A. Balogh, and G. Jones, Our changing Sun, *Astron. Geophys, 40*, 4.10-4.16, 1999b.

Marsh, N., and H. Svensmark, Cosmic rays, clouds and climate, *Space Sci. Rev.*, in press, 2000.

Mayaud, P.N., The *aa* indices: A 100-year series characterising the magnetic activity, *J. Geophys. Res., 77*, 6870-6874, 1972.

Moraal, H., Cosmic ray modulation studies in the outer heliosphere, *Nucl. Phys. B, 33A,B,, proc. suppl.*, 161-178, 1993.

Mursula, K., and T. Ulich, A new method to determine the solar cycle length, *Geophys. Res. Lett., 25*, 1837-1840, 1998.

Nesme-Ribes, E., E.N. Ferrierer, R. Sadournay, H. LeTreut, and Z.X. Li, Solar dynamics and its impact on solar irradiance and the terrestrial climate, J. Geophys. Res., 98, 18923-18935, 1993.

O'Brien, K., Secular variations in the production of cosmogenic isotopes in the Earth's atmosphere, J. Geophys. Res., 84, 423-431, 1979.

Parker, D.E., P.D. Jones, A. Bevan, and C.K Folland, Interdecadal changes of surface temperatures since the late 19th century, *J. Geophys. Res., 99*, 14,373-14,399, 1994.

Potgieter, M.S., The long-term modulation of galactic cosmic raysnin the heliosphere, *Adv. in Space Res, 16(9),* 191-203, 1995

Pulkkinen, T.I., H. Nevanlinna, P.J. Pulkkinen, and M. Lockwood, The Earth-Sun connection in time scales from years to decades to centuries, *Space Sci. Rev.*, in press, 2000.

Rind, D. and J. Overpeck, Hypothesized causes of decade-to-decade climate variability: climate model results, *Quaternary Sci. Rev., 12*, 357-374 (1993)

Rossow, W.B., A.W. Walker, D.E. Beuschel, and M.D. Roiter, International Satellite Cloud Climatology Project (ISCCP): Documentation of new datasets, WMO/TD 737, World Met. Org., Geneva, 1996.

Russell, C.T., and R.L. McPherron, The magnetotail and substorms, *Space Sci. Rev., 15,* 205-225, 1973.

Sabbah, I., The role of interplanetary magnetic field and solar wind in modulating both galactic cosmic rays and geomagnetic activity, *Geophys. Res. Lett., 27*, 1823-1826, 2000.

Solanki , S.K. and M. Fligge, Solar irradiance since 1874 revisited, *Geophys. Res. Lett., 25*, 341-344, 1998.

Solanki, S.K. and M. Fligge, A reconstruction of total solar irradiance since 1700, *Geophys. Res. Lett., 26*, 2465-2468, 1998.

Solanki , S.K., M. Schüssler, and M. Fligge, Secular evolution of the Sun's magnetic field since the Maunder minimum, *Nature*, in press, 2000.

Soon, W., S. Baluinas, E.S.Posmentier, and P. Okcke, Variations of solar coronal hole qarea and terrestrial lower tropospheric air temperature from 1979 to mid-1998: Astronomical forcings of changes in Earth's climate?, New Astron., *4*, 563-579, 2000.

Stamper, R., M. Lockwood, M.N. Wild, and T.D.G. Clark, Solar causes of the long term increase in geomagnetic activity, *J. Geophys Res., 104,* 28,325-28,342, 1999.

Stuiver, M., and P.D. Quay, Changes in atmospheric carbon-14 attributed to a variable Sun, *Science, 207*, 11, 1980.

Svensmark, H., Influence of cosmic rays on Earth's climate, *Phys. Rev. Lett., 81*, 5027-5030, 1998.

Svensmark, H., and E. Friis-Christensen, Variation of cosmic ray flux and global cloud coverage: A missing link in solar-climate relationships, *J. Atmos. Sol. Terr. Phys., 59*, 1225-1232, 1997.

Tett, S., P.A. Stott, M.R. Allen, W.J. Ingram and J.F.B. Mitchell, Causes of twentieth century temperature change, *Nature,388*, 569-572, 1999.

Usoskin, I.G., H. Kananen, K. Mursula, P. Tanskanen, and G.A. Kovaltsov, Correlative study of solar activity and cosmic ray intensity, *J. Geophys.Res., 103*, 9567-9574, 1998.

Vasyliunas, V.M., J.R. Kan, G.L. Siscoe, and S.-I. Akasofu, Scaling relations governing magnetospheric energy transfer, *Planet Space Sci., 30*, 359-365, 1982.

Wang, Y.-M., N.R. Sheeley Jr., and J. Lean, Understanding the evolution of Sun's magnetic flux, *Geophys. Res. Lett., 27*, 621-624, 2000a.

Wang , Y.-M., J. Lean, and N.R. Sheeley Jr., The long-term evolution of the Sun's open magnetic flux, *Geophys. Res. Lett., 27,* 505-508, 2000b.

Wigley, T.M.L., and S.C.B. Raper, Climatic change due to solar irradiance changes, *Geophys. Res. Lett., 17*, 2169-2172, 1990.

Wigley, T.M.L., P.D. Jones, and S.C.B. Raper, The observed global warming record: what does it tell us? *Proc. Natl. Acad. Sci., USA, 94*, 8314-8320, 1997

Willson, R.C., Total solar irradiance trend during cycles 21 and 22, *Science, 277,* 1963-1965, 1997.

M. Lockwood, World Data Centre C-1 for Solar-Terrestrial Physics, Space Science Department, Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, England, U.K. (m.lockwood@rl.ac.uk)