GEOACTA, 21, 107-125, 1994

SOLAR VARIABILITY EFFECTS ON CLIMATE

Juan G. Roederer
Geophysical Institute
University of Alaska Fairbanks

ABSTRACT

One of the key tasks in climate research at the present time is to separate anthropogenic effects from natural change. Solar variability is one possible cause of natural change in addition to other external phenomena such as volcanic eruptions and, on a long-term scale, orbital changes and celestial body impacts. The topic of "sun-weather relationships" has followed a long and tortuous history of scientific speculations and controversy since the continuous observation of sunspots began in the 17th century. It was not until recently that more systematic studies with long-term data bases of meteorological, climatological and solar parameters led to an increasing, statistically robust, body of evidence for a causal connection between some manifestations of solar variability and changes in the troposphere and the climate system. While of planetary scale, the strength and sign of pertinent correlations have distinct geographic, seasonal and other temporal characteristics. Most likely, several trigger mechanisms are at work simultaneously, but their relative importance may depend on the time-scale envisaged and on competing processes such as volcanic eruptions. In this paper we will summarize the most frequently formulated criticisms, review the most recent results on relevant solar variability effects at the 11-year, secular, and short-term time scales, and discuss proposed mechanisms such as enhanced Hadley circulation, charged particle ionization effects on clear-air conductivity, and electric field effects on the microphysics of cloud formation.

RESUMEN

En la actualidad, una tarea clave de la investigación climática es la de separar los efectos antropogénicos de los cambios naturales.

La variabilidad del sol es una posible causa del cambio natural junto con otros posibles fenómenos de orden externo como por ejemplo las erupciones volcánicas y en escalas de tiempo mas largas cambios orbitales y el impacto de cuerpos celestes. El tema "relación sol-clima" ha seguido una larga y tortuosa historia de especulación científica y controversias desde que en el siglo 17 se empezaron a observar en forma contínua las manchas solares. Recientemente, sobre la base de estudios más sistemáticos y bases de datos de períodos largos de parámetros meteorológicos, climatológicos y solares se dispone de un cuerpo de evidencias amplio y estadísticamente robusto de la conexión causal entre algunas manifestaciones de la variabilidd solar y cambios en la tropósfera y el clima. En la escala planetaria, la intensidad y signo de las correlaciones poseen distintas características geográficas, estacionales y otras temporales. Muy probablemente, distintos mecanismos de "trigger" operan simultaneamente, pero su importancia relativa puede depender de la escala de tiempo considerada y de procesos competitivos como las erupciones volcánicas. En este trabajo, haremos un resumen de las críticas mas frecuentes, veremos los resultados recientes más relevantes de los efectos de la variabilidad solar de 11 años, secular, y escala de tiempo corta; y discutiremos mecanismos propuestos como la circulación de Hadley ampliada, efectos ionizantes de partículas cargadas sobre la conductividad del aire limpio, y efectos del campo eléctrico en la microfísica formación de nubes.

1. INTRODUCTION

The topic of "sun-weather relationships" has followed a long and tortuous history of scientific speculation and controversy. As soon as sunspots became the subject of systematic observation after Galileo's invention of the telescope, scientists began to wonder about the possible effects of these "blemishes" on terrestrial weather. Yet from the very beginning, the efforts to discover Sun-weather relationships carried the stigma of "bad science". This was in part due to the fact that the only scientific paradigm available was to correlate selected meteorological parameters with time-dependent features on the Sun and hope that the correlation would hold up in the future. And in most cases it didn't!

It was not until the first decades of this century that more systematic studies were begun using long-term data bases of meteorological, climatological and solar parameters. An early example is the study of the global distribution of annual rainfall difference between sunspot cycle maximum and minimum, given in the book by Clayton (1923); later, Roberts (1975) discussed statistics on the occurrence of droughts in Nebraska, which showed a tendency to occur near that sunspot minimum which followed a magnetically negative maximum (of the 22-year Hale cycle). Clayton's study clearly showed that if a physical relationship existed between the 11-year sunspot cycle and tropospheric phenomena, it had to exhibit a regional dependence. Roberts' paper brought in the overall magnetic configuration of the Sun (because of the correlation with the Hale cycle), and it also showed a phase-locking to the quasi-periodic character of the solar cycle (frequency modulation), a statistically significant fact given the low probability that an atmospheric periodicity unrelated to the Sun would track the changes in solar cycle length just by chance.

During the seventies it became apparent that it was necessary to clearly divide the study according to three classes of solar variations: 1. the 11-year sunspot cycle (and, eventually, the related 22-year magnetic Hale cycle); 2. the long-term secular changes of sunspot cycle amplitude (the Gleissberg "cycle"); 3. short-duration, sporadic events, such as solar flares and the Sun-controlled solar wind shocks and reversals of the interplanetary magnetic field. Concerning secular effects, Eddy (1976) published a classical paper revealing the remarkable correlation of winter temperatures in London and Paris with the Gleissberg cycle (traced back in time to the 12th century by using as proxy indicator of solar activity the C-14 concentration in tree rings). This paper received particular attention because it linked the "Little Ice Age" in the 17th century and the cold period in the 15th century with the Maunder and Spörer minima, respectively, in which few sunspots were seen for several decades. Regarding short-term variability, Wilcox et al. (1974) published a study of changes of the northern hemisphere "vorticity area index" ("VAI", a quantitative measure of low-pressure troughs) at the times of magnetic sector boundary passages (interplanetary magnetic field reversals), showing statistically significant decreases of the VAI from 2 days before to one day after the passage during winter months (this correlation, however, did not subsist during the eighties).

These and many other studies (e.g., Hines and Halevy, 1977; Bucha, 1988) could not dispel the general skepticism about the subject per se, particularly on part of the meteorological community. It is only in recent years that the availability of statistically robust results and the formulation of plausible trigger mechanisms to explain the observed correlations have placed the field in a more respectable standing with the general scientific community.

The study of solar variability effects on the immediate human environment is now an important and integral part of the ICSU/SCOSTEP Solar Terrestrial Energy Program 1990-1997 (STEP) (Roederer, 1992). The present paper will discuss the most frequently formulated criticisms of solar activity-climate studies and review some important recent results. For detailed literature references, the reader is referred to the articles cited in this paper.

2. MOST FREQUENT CRITICISMS AND "FREAK FACTS OF NATURE"

The study of solar variability effects on climate has been subjected to severe criticism that can be grouped into several distinct categories. While this criticism was well founded on historical grounds, much of it can now be dispelled on the basis of recent studies.

The power involved in solar variability is too small! Indeed, the total solar electromagnetic power deposited in the Earth's environment is more than 10^{12} MW, whereas the variable components of electromagnetic and particle energy impinging on the Earth system represent only 10^4 MW to 10^6 MW, i.e., only a tiny fraction of the main energy flux. Furthermore, the variable energy components mainly affect the magnetosphere and upper atmosphere which are only very weakly coupled to the troposphere. The energy argument, however, is not valid for highly non-linear, complex systems such as the coupled atmosphere-ocean-biosphere. It is well known that complex systems can behave chaotically, i.e., follow very different paths after the smallest change in initial or boundary conditions, or in response to the smallest perturbation. In a highly non-linear system with large reservoirs of latent energy such as the atmosphere-ocean-biosphere, global redistributions of energy can be triggered by very small energy inputs, a process that depends far more on their spatial and temporal pattern than on their magnitude. Moreover, such a system may exhibit sudden transitions between some "eigenstates" of quasi-equilibrium; recent data from the Greenland ice sheet clearly show this for the Northern Hemisphere atmosphere prior to the Holocene period (Dansgaard et al., 1993).

We cannot think of any mechanisms responsible for solar-climate effects! This is an "unscientific" argument; there are abundant historical examples of processes which were studied, accepted or used, but for which the responsible mechanism was not known for a long time (e.g., dinosaur extinctions, plate tectonics, or atomic spectral line emissions which, although "prohibited" by classical electrodynamics, had been in use in science and technology many decades before their explanation by quantum mechanics). Besides, the argument as a whole does not hold anymore: trigger mechanisms are presently being formulated and studied.

It's all just a coincidence! Historically a quite valid argument, but it does not hold anymore: there are too many statistically robust results the probability of which to occur all by chance is extraordinarily small. This in particular applies to those correlations which stay in phase with the quasi-periodicity of solar activity. Predictions based on recent studies covering a limited period of time were verified when new, later, data came in (e.g., Labittzke and van Loon, 1993) or when more data reaching into the past were included (e.g., Friis-Christensen and Lassen, 1993).

The intervals for which data are available for correlation studies are too short! This is a valid criticism when applied to some of the recent studies of 11-year cycle effects. The band-width of uncertainty in the determination of the period using data from a time-interval of, say, only three solar cycles is ± 2 years; therefore, any intrinsic atmospheric periodicity between 9 and 13 years could be expected to show an acceptable correlation with the solar cycle (e.g., see the discussion in Salby and Shea, 1991). Until now, however, no plausible decadal cycle variation of the atmosphere unrelated to the Sun has been identified or proposed.

The "So what?" question. It has been argued that whereas solar variability-climate effects do exist, they are so small that they are unimportant for forecasting purposes (e.g., Pittock, 1979). Leaving aside the question of whether they are really that small, this is yet another "unscientific" argument: every global and coherent response should be considered important for the physical understanding of the atmospheric system!

In addition to the above criticisms which the researchers in this field may continue to face, there are some "freak facts" of Mother Nature that conspire to make physical interpretations especially difficult.

1. All solar emissions exhibit 11-year periodicity; this makes source identification more difficult. For instance, could an 11-year periodicity in the atmosphere be due to the 11-year variability of solar irradiance, or could it be due to the cumulative effect of short-term events (like solar flares), whose occurrence also exhibits an 11-year periodicity?

2. The amplitudes of the last two solar maxima (both in terms of sunspot number and total irradiance) were nearly identical. This makes it difficult to model solar irradiance (for which absolute satellite measurements are available only after 1978) and its relation to sunspot number, a fact that in turn prevents using the latter as a proxy to estimate total solar irradiance values in the past.

3. The time lag between annual average temperature over land and the mean sea-surface temperature (governed by the ocean response time), and the time lag between the 11-year running means of solar cycle amplitude (the Gleissberg cycle) and solar cycle duration, are both of the order of 10-15 years. This obscures the discriminability among proposed mechanisms (see Section 4.2).

4. The beat period between the 26-28 month quasi-biennial oscillation (QBO) and the biennial period (24 months) is 12-14 years, i.e., of the same order as the solar cycle period. In principle, this can decrease statistical discriminability in QBO-stratified time-series (aliasing effect, Teitelbaum and Bauer, 1990), but it is of little importance due to the strongly quasi-periodic character of the QBO (see section 4.1).

5. GCM calculations show that random stochastic variations of the mean atmospheric temperature can have similar amplitudes and time scales as the observed variations. This makes the distinction between unpredictable intrinsic variations and those driven by external and anthropogenic forcing difficult.

6. The "ultimate freak fact" is that key parameters such as the global temperature, sunspot cycle amplitude, greenhouse gas concentrations and sunspot cycle frequency all show a net increase since the late 1800's. This poses extra difficulties in the determination of the differential climate sensitivity to the various possible forcing factors.

3. THE VARIABLE INPUT CHANNELS

Figure 1 depicts schematically the most important energy input channels from the Sun, their continuous or sporadic variability and the regions in the terrestrial atmosphere that can be affected by this variability. Solar magnetohydrodynamics regulates all forms of variability, including that of the intensity of the cosmic ray flux, the configuration of the high-latitude ionospheric electric field and the precipitation of energetic trapped electrons, all of which are controlled by the interplanetary magnetic field (IMF). The variability of the two photon channels at left is controlled by surface structures, such as faculae and plages (radiation emitters) and sunspots (blockers), whereas the particle channels are mainly controlled by processes in the coronal plasma. Finally, the Earth's internal magnetic field, which is variable on a secular scale (including an approximately 6,000-year oscillation of the main dipole moment), also influences the fluxes of the charged particles penetrating the atmosphere. Not considered at all in the figure (and in this paper) are the variations of solar insolation due to periodic changes in orbital and rotational parameters of the Earth (leading to the Milankovitch cycles). Also not considered are the natural 27-day periodicities related to solar rotation. The atmospheric processes sketched in Fig. 1 will be discussed in Chapter 5.

An important discovery (e.g., Willson and Hudson, 1988) was the small but significant 11-year modulation of the "solar constant" (total irradiance in W/m^2 at 1 a.u.). Figure 2. (from Kyle et al., 1993) shows monthly mean values of the total irradiance as measured with a radiometer on Nimbus 7, together with the Wolf sunspot number. The 11- year modulation comes from two sources of mutually counter-acting effects: enhanced emissions from bright faculae during solar maximum and enhanced blocking by sunspots (obviously the former wins over the latter). In addition, a long-term variation is speculated to exist, caused by the possible growth and decay of the global facular network on the solar surface (White et al., 1992). Fig. 2 shows that the overall relationship between total irradiance and sunspot number during one cycle is clearly non-linear and is different during the descending and ascending phases of solar activity. It is important to note that the UV band of the photon spectrum exhibits a much larger 11-year variability than that shown for the total irradiance (Donnelly,1991).

Since the two recent solar maxima shown in Fig. 2 are of nearly the same amplitude ("freak fact No. 2"), nothing can be concluded from these measurements about the long-term variation of irradiance. Using the satellite measurements and an isolated value of solar irradiance obtained by Kosters and Murcray (1979) in a 1968 balloon flight, Reid (1991) established a linear relationship between total solar irradiance and the envelope of the 11-year sunspot cycle (which he used in model calculations of long-term change of sea surface temperatures; see Section 4.2). This relationship, however, would imply a value of the solar constant 11W/m^2 lower during the cold period of the 1600's, an unrealistically large change. Lean et al. (1992), for instance, estimate that the excess facular radiation from a complex magnetic surface configuration as it exists in the contemporary Sun contributes about 1.5 W/m^2. However, using observations from solar-like stars, these authors conclude that for non-cycling stars (e.g., in a Maunder minimum-like state) the irradiance should be further reduced, below that derived for a total removal of network magnetic flux; their final estimate of the Maunder minimum total irradiance reduction is about 2.7 W/m^2 below the contemporary solar minimum value. These estimations are quite significant from the climatic point of view: GCM calculations show that at least half of the global temperature increase since the Little Ice Age could be explained by solar radiative forcing (Rind and Overpeck, 1993).

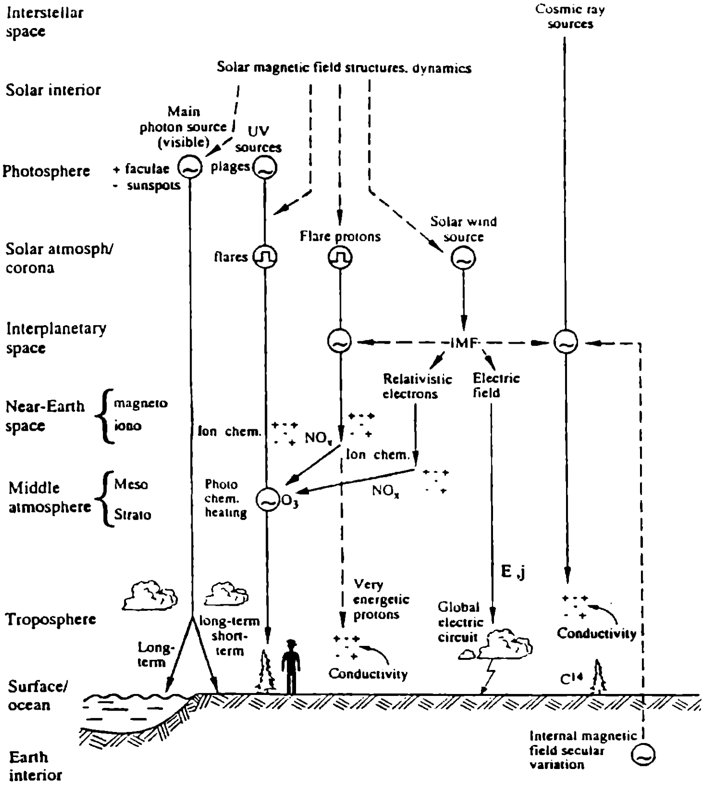

Figure 1. *Sketch of the channels of energy input from the Sun (and beyond), indicating their continuous or sporadic variability (wavy and square-wave signs respectively), and the regions in the terrestrial atmosphere that can be affected by this variability.*

Coronal holes, coronal mass ejections and solar flares are transient manifestations of solar activity at different time-scales; although in themselves aperiodic, they do occur with frequencies tied to the 11-year cycle. They all have important effects on solar wind density, speed and magnetic field configuration (the IMF). Energetic protons from large flares can penetrate the atmosphere down to sea level; shock waves emitted by flares "sweep away" cosmic rays, causing important decreases of their intensity at Earth (the so-called Forbush decreases); sudden changes in the direction of the IMF alter the transfer of solar wind energy into the magnetosphere and can cause magnetic storms and aurorae; changes in the IMF also alter ionosphere, thus affecting the "global electric circuit" in the entire atmospheric system. The higher solar wind speed and the enhanced magnetic irregularities during solar maximum are responsible for a general decrease of the cosmic ray flux, which therefore is anti-correlated to solar activity. During magnetic storms, the magnetosphere "squeezes out" high energy electrons transiently trapped in the Earth's magnetic field (Baker et al., 1987), which subsequently precipitate into the upper atmosphere, preferentially in the region of the so-called South Atlantic Anomaly (where the Van Allen radiation belt comes closest to the atmosphere).

Figure 3 shows the "good old" monthly average sunspot number curve. This is the only indicator of solar activity that can be traced back reliably to the 17th century (Nesmes-Ribes, 1993). It is important to note, however, that it does not represent equally well all aspects of solar variability: the solar activity at different sunspot minima and its direct effects such as geomagnetic disturbance may be rather different, even if it does not manifest itself in the observed sunspot number directly. The sunspot number is not a single-frequency sinusoidal function of time; rather it has a notable amplitude modulation (given by the envelope of the curve in Fig. 3) and a frequency modulation as well (solar cycle length values ranging from 9-12 years). Notice also that the ascending and descending phases have varying slopes. Thus, solar

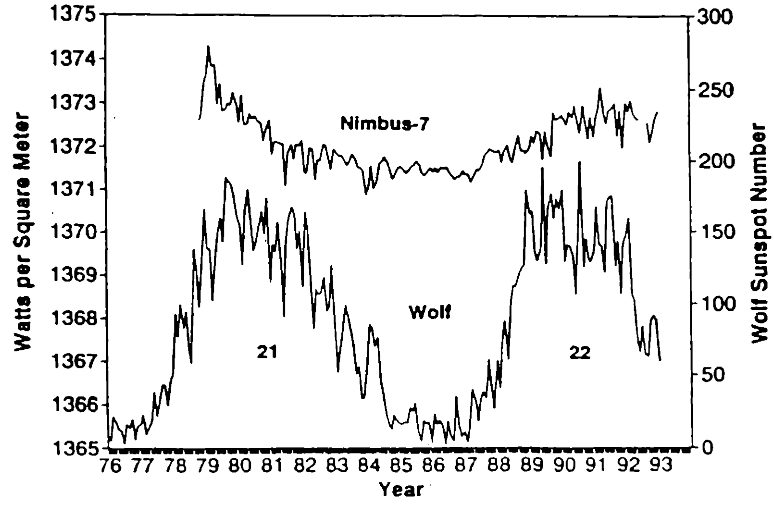

Figure 2. *Monthly mean solar irradiance values and sunspot numbers (from Kyle et al., 1993).*

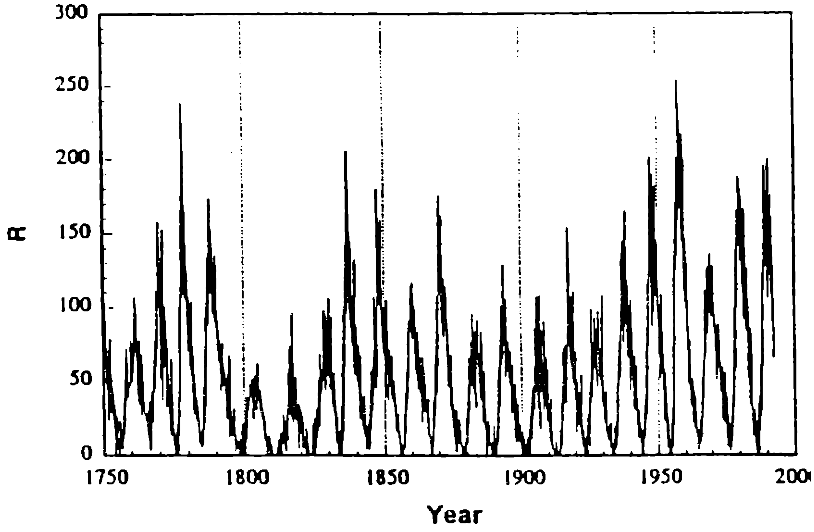

Figure 3. *The monthly mean sunspot number **R** during 2 1/2 centuries.*

cycle amplitude, length and, for instance, maximum time rate of change are important parameters for solar-terrestrial correlations studies. The envelope of the sunspot curve of Fig. 3 follows by 10-20 years a similar-looking curve of (minus) the solar cycle length (shown in Fig. 8); this is approximately the same time-lag that exists between the average global land and sea-surface temperatures ("freak fact No. 3"). Today, the 10.7 cm radiowave flux is often used instead of sunspot number as a measure of solar activity.

4. RECENT RESULTS

4.1 11-Year Cycle

A breakthrough in solar variability effects on climate came in 1987. K. Labitzke of the Free University of Berlin had been engaged in a systematic study of the northern polar stratosphere and its relationship to the quasi-biennial oscillation (QBO) of the equatorial stratosphere (Labitzke, 1982). The QBO refers to the winds at different layers of the tropical stratosphere, which reverse their direction (East to West) with a quasi-periodicity of about 26-28 months (the higher layers leading the lower ones). For the winter hemisphere, this really represents two possible dynamic states of the stratosphere, one in which the equatorial part of a given layer rotates in the same sense as the polar vortex (the west QBO phase) and one in which it contra-rotates; each state presents very different conditions for the propagation of kinetic and thermal energy and momentum toward the higher latitudes. In particular, it was found that the polar vortex was strong and stable, thus colder, during the contra-rotating west phase than during the contra-rotating east phase when major disruptions (major mid-winter warmings) tended to occur in the polar stratosphere. But there were winters in which the reverse situation arose; they happened to correspond to epochs of solar maximum. Indeed, Labitzke (1982) found that all major mid-winter warmings that occurred during the west phase happened only at times of solar maximum.

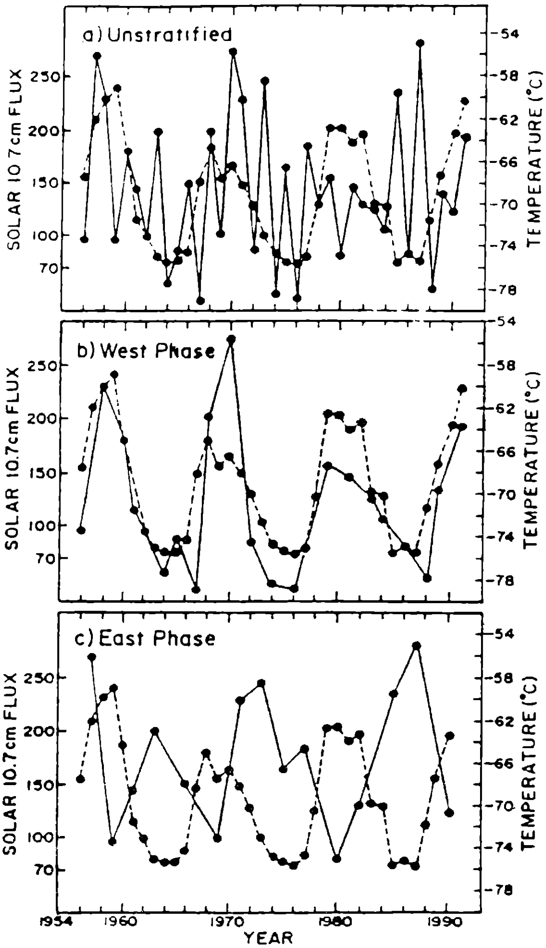

Figure 4. *10.7 cm solar radio wave flux (broken line -in units of 10 W m Hz) and the mean 30 mb temperature at the North Pole during winter (January-February). a): All years; b): Years with west phase of the QBO (westerly winds at the equatorial 50-40 mb level during January-February); c): Years with east phase (easterly winds). Updated from Labitzke (1987).*

Analyzing the temperature of the 30 mb level at the North Pole during January and February for three solar cycles, Labitzke (1987) found an astounding correlation/anti-correlation with solar activity (as expressed by the 10.7 cm flux) when the data were stratified according to the west/east phase of the QBO (defined by the equatorial wind at the 50-40 mb level). Figure 4 is an updated version of these results. It is thus clear that high solar activity introduces a radical perturbation in the process responsible for the global coupling between the tropical and the polar regions of the stratosphere in winter.

Labitzke and H. van Loon from NCAR embarked in a systematic study extending the region under analysis towards lower latitudes, down into tropospheric altitudes, and to the other seasons. Concentrating on the height of the 30 mb level as an important indicator of the integral behavior of the air column below, they identified a clear geographic dependence of the correlation with solar activity; in particular, during mid-winter and east years of the QBO, the basic correlation pattern exhibits a crescent-shape region of positive correlation of the 30 mb level height along 30° N over the Pacific, with a pattern of anti-correlation centered over the Arctic. During the west years, the only important correlation (positive) is found over the Arctic. A consistent average geographic pattern of correlation with solar activity subsists for all data regardless of the QBO phase, as shown in Figure 5 (Labitzke and van Loon, 1993).

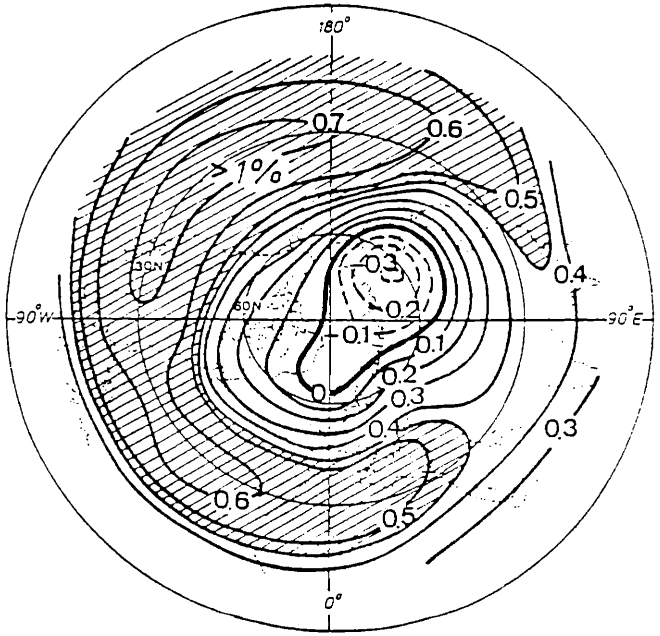

Figure 5. *Contours of equal correlation coefficient between the annual mean height of the 30 mb level (1958-1992) and the 10.7 cm solar flux (hatched area: local statistical significance greater than 1%). From Labitzke and van Loon (1993).*

Figure 6 shows the average change of temperature from solar maximum to solar minimum as a function of geopotential height for Lihue (Hawaii), which is situated under the maximum correlation area of Fig. 5. Different epochs during the year are shown (the reversal at the tropopause at about 100 mb is an inherent property of the atmosphere). The annual mean temperature difference in the troposphere between solar maximum and minimum is substantial, 1.8° C at 300 mb; even at the surface it is 0.9° C in summer. The large positive correlation between 30 mb height and solar flux in the Pacific region is thus mainly due to the difference in tropospheric temperature between the extremes of the 11-year sunspot cycles; this is significant for the identification of trigger mechanisms responsible for solar-climate relationships (see Chapter 5). As to the aliasing effect mentioned in "freak fact No. 4", it does not apply to a strongly period-modulated variation such as the QBO (Tinsley and Heelis, 1993).

An intriguing result is that of Mendoza et al. (1991) concerning statistics of the occurrence of El Niño events from 1700 to 1985. It had been noted earlier (Pérez-Enríquez et al., 1989) that the most intense events have occurred during periods of anomalous solar activity and that in general El Niño/Southern Oscillation (ENSO) events tend to gather around peaks of auroral activity during the descending phase of solar sunspot cycle. Mendoza et al. analyzed the frequency of occurrence of ENSO events according to intervals of sunspot gradient values (time derivative expressed in sunspot number change per year). They conclude that 63% of the ENSO events occur during the descending phase of the solar cycle, and that there are twice as many events occurring one year after the maximum rate of sunspot decline

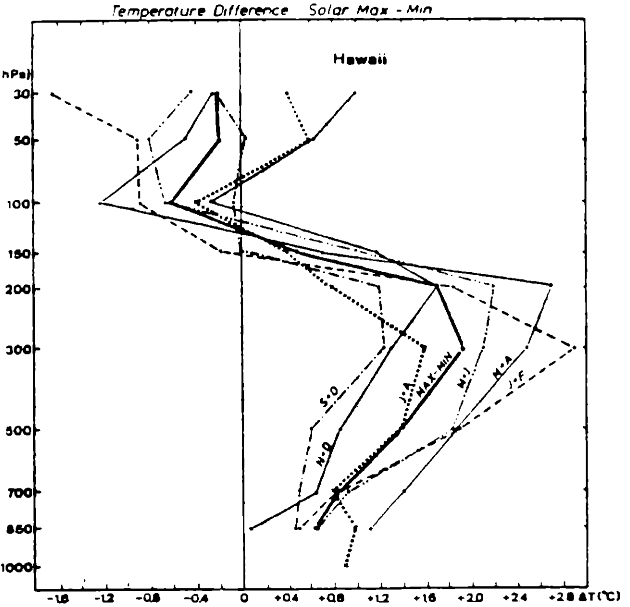

Figure 6. *The average change of temperature from solar minimum to maximum as a function of atmospheric depth, for Lihue (Hawaii), through the year and in the annual mean (heavy line), for a 36-year series. From Labitzke and van Loon (1993).*

(maximum negative gradient) than what would be expected by random occurrence (see Figure 7). While at least part of the first result can be explained by the asymmetry of the sunspot number curve with respect to each maximum (notice in Fig. 3 that the sunspot number rise is usually faster and of shorter duration than the decay), the second result is more significant. Indeed, the authors conclude from a computer experiment that the probability for the peak in Fig. 7 to occur by chance is less than 0.4%.

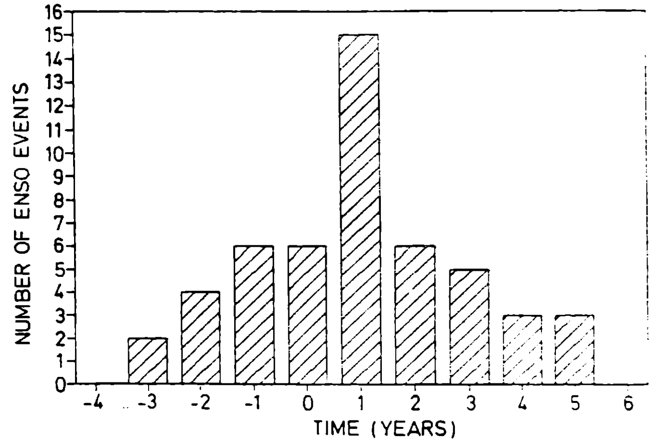

Figure 7. *Histogram of El Niño events around the year of maximum rate of sunspot decrease during the descending phase of the solar cycle, for the period 1700-1985. From Mendoza et al. (1991).*

4.2 Long-Term Variations

The annual average Northern Hemisphere land air temperature curve of Jones et al. (1986) (or Hansen and Lebedeff (1987) for the global temperature), popularly known as the "global warming curve", is usually interpreted by non-scientists (and the media) as being entirely due to an anthropogenically enhanced greenhouse effect. However, it is clear that not all of the global temperature variation in the last 100 years could have been of anthropogenic origin: 1. most of the increase during this century (78%) took place before 1940 when the rate of increase of CO_2 was much lower than at present; 2. there was a steady temperature decrease between 1940 and the early seventies while the CO_2 kept increasing at an accelerated rate; 3. the global temperature had already been rising during the two previous centuries since the Little Ice Age. Still, it came as a surprise when Friis-Christensen and Lassen (1991) published a paper on the correlation between the global temperature and the length of the solar cycle, which at first sight seemed to indicate that the entire temperature behavior during the last 100 years could be due to solar variability (although the authors never stated this - see "freak fact No. 6").

Recently, Friis-Christensen and Lassen (1993) extended this correlation back to 1750 using Northern Hemisphere temperature data by Groveman and Landsberg (1979). They also re-computed the solar cycle length curves using a filtering procedure to take into account earlier criticism of the determination of solar cycle length as a time-dependent parameter. As Figure 8 shows, the remarkable relationship between average annual temperature and solar cycle length persists. In their 1991 paper, the authors noted the similarity of the global temperature anomaly and the 11-year running mean of the sunspot number, but pointed out the fact that the temperature curve was leading the sunspot number curve by up to 20 years, which of course ruled out any causal connection between the two. But as shown in Fig. 8, this time shift disappears when sunspot cycle length is used, implying that it is the cycle length and not the sunspot number that appropriately represents the climate-relevant part of solar variations (for instance, the changes in irradiance). Both the 11-year running mean sunspot number and cycle length do track each other well, but the cycle length leads by 10-15 years (see "freak fact No. 3").

Another study of long-term correlations was conducted by Reid (1991), who used the sea surface temperature record as a climate change indicator. Despite some questions about quality of data before the beginning of the century, global average sea surface temperature variations have certain advantages: because of the thermal inertia of the ocean, they do not exhibit short-term variations as the land temperatures; they do not have to be corrected for effects from recent urban growth; they exhibit greater spatial coherence; and they represent samples from an area covering 70% of the Earth's surface. As mentioned in the previous Chapter, Reid derived an empirical relationship between solar irradiance and sunspot number and, using a one-dimensional columnar model of the coupled atmosphere-ocean system, computed the theoretical average sea surface temperature, shown in Figure 9 together with the experimental data. The calculation consisted of the integration of a heat diffusion equation in the ocean column governed by a set of coefficients and subjected to given coupling and boundary conditions at the top and bottom of the ocean, respectively; this equation was integrated forward in time from the late 1600's, and is shown to fit well the measured average temperature variations (Fig. 9).

As Reid pointed out, the experimental average sea surface temperature follows in general lines the global land temperature, but it is delayed with respect to the latter by 10-15 years. This is three times the value of the time constant for radiative equilibrium of the ocean mixed layer used by Reid in the model calculation. It is reasonable to assume that if one were to use the solar cycle length, which precedes the envelope of the sunspot number by 10-15 years (see "freak fact No. 3"), one would have to use a larger (and perhaps more realistic) value of the ocean response time to achieve a good fit. Another point worth re-examining is the unrealistically large secular variation of solar irradiance used in Reid's calculations (Chapter 3); an appropriate amplification mechanism might have to be introduced on the atmospheric side of the model (see next Chapter).

Finally, concerning very long-term changes, Anderson (1992) reported on a possible connection between surface winds, solar activity and the Earth's internal magnetic field. The author has identified an association between an enhanced 100-200 year solar cycle periodicity (as revealed in the C-14 record in tree rings) and fluctuations in surface wind intensity on a 200 year time scale (as revealed in the thickness of varved sediments in a Minnesota lake). This association existed only during the mid-Holocene 5000-7500 years ago; when the Earth's magnetic dipole moment went through the last minimum, i.e., when the intensity of the solar-activity-modulated cosmic rays entering the atmosphere was at a maximum.

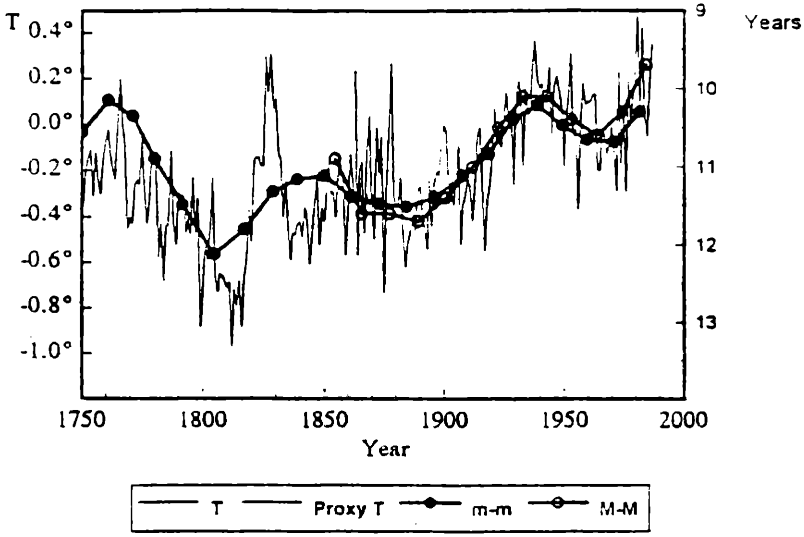

Figure 8. *Annual average values of the Northern Hemisphere temperature (thin curve) and appropriately filtered values of the sunspot cycle length (solid curves, determined independently by means of sunspot minima (m-m) and maxima (M-M); note reverse scale). See Friis-Christensen and Lassen (1993).*

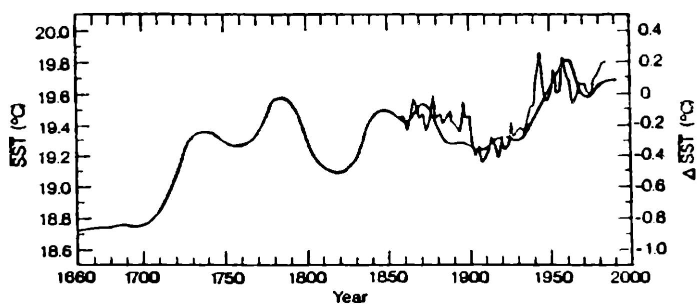

Figure 9. *Global average sea-surface temperature calculated from Reid's model, and the observed time series (Reid, 1991).*

4.3 Short-Term Variations

The studies of possible effects of short-term solar-activity-related variations on the tropospheric system are on much less firm footing than the 11-year cycle and longer-term correlations, despite the fact that, in principle, the former would be easier to identify statistically than the latter because of the unique signatures of each one of the possible solar input on a day-to-day time-scale. These studies are mostly isolated efforts by individual scientists, each one of which has chosen one given pair of solar-atmospheric variables out of the many possible ones; a concerted, internationally coordinated approach is still missing.

One example is Schuurmans' (1991) study of the temperature in the troposphere (500 mb level) and lower stratosphere (200 mb) over De Bilt, The Netherlands, and its behavior during 72 solar proton events that occurred in the interval 1955-1984. The author finds that for the east phase of the QBO there is a clear reduction of the temperature at the 200 mb level of about -2.4° C that persists at least 3 weeks after the proton event . No measurable effect is seen during west QBO. Comparing the behavior of the 200 to 500 mb levels, the author argues that the cooling in the lower stratosphere is not due to a dynamically related warming of the troposphere; rather, he postulates a solar-induced heat sink operating at 200 mb.

Another example is the work by Pudovkin and Babushkina (1992) who studied atmospheric transparency variations during the interval 1961-1984 associated with geomagnetic disturbances, using monthly actinometric data from a network of meteorological observatories in three latitudinal bands of the former Soviet Union. They find a considerable increase in ground-level solar radiation intensity (increase in transparency) in the auroral zone, 1-2 days after the onset of a geomagnetic storm. The authors speculate that this may be due to chemical changes in the stratosphere as the result of storm-associated Forbush decreases of cosmic ray intensity.

5. MECHANISMS

As stated in Chapter 2, because of energy balance considerations, the mechanisms responsible for solar variability effects on climate should be trigger-mechanisms which catalyze or control the release of latent energy, thus leading to a large-scale re-distribution of energy in the atmospheric system. To explain the observed effects, one must look for amplification processes that respond by a factor of at least one million to small variations of energy inputs. There are three basic candidates for mechanisms, sketched in the lower part of Figure 1 and discussed below.

5.1 Total Irradiance Variations

These variations correspond to the main photon channel at left in Fig. 1. Their 11-year variation is shown in Fig. 2; a bigger change can be expected on a long-term, secular scale (Chapter 3). Labitzke and van Loon (1993) proposed the modulation of the Hadley cell circulation as a possible mechanism, based mainly on the fact that the middle and upper troposphere under the crescent-shaped region between 20° N and 45° N in the Pacific-Atlantic area shown in Fig. 5 is consistently warmer during solar maximum (e.g., see Fig. 6). This points to an enhanced Hadley circulation in which the troposphere in the Intertropical Convergence Zone is warmed more by an increased release of latent heat during solar maximum. The reversal of the temperature effect at and above the tropopause in Fig. 6 indicates that the primary enhanced heating process resides within the tropopause (because the vertical temperature difference profiles in Fig. 6 are indeed similar to the profiles of differences between non-solar related strong and weak Hadley circulation periods). The control by the QBO of wave and energy propagation characteristics to higher latitudes in the winter hemisphere (Section 4.1) would be responsible for the 11-year modulation of the polar stratosphere (Fig. 4). In the summer hemisphere, in absence of a well-defined polar vortex, the coupling between the equatorial and polar stratosphere is less complex, and the atmosphere exhibits a behavior parallel to the 11-year cycle regardless of the phase of the QBO (van Loon and Labitzke, 1990). This is yet another argument in favor of the Hadley circulation modulation hypothesis.

In principle, long-term solar variability effects on climate such as shown in Fig. 8 and 9 should be "easier" to explain because the input power variations could be expected to be several times larger than those found for the 11-year cycle (Chapter 3). However, it is not yet clear (although quite suggestive) whether the above Hadley cell modulation mechanism would also be applicable at this time scale. GCM model calculations do show that a total irradiance change of about 2 W/m^2 can have an effect on global temperature of about half that of doubling the CO_2 concentration.

5.2 Solar UV Flux Changes

Given that the UV flux powers the dynamics of the stratosphere via ozone absorption (e.g., Hood et al., 1993), and given the large variability of the solar UV flux, it is reasonable to expect this flux to play a fundamental role in driving stratospheric variability (e.g., Kodera et al., 1991). Such role may be a contributing factor, but it could not explain other important findings by Labitzke and van Loon, such as shown in Figs. 6 and 7. It may, however, be related to the behavior of the polar stratosphere in winter during the west phase of the QBO, when its temperature is positively correlated with solar activity (Section 4.1).

Finally, we may speculate that the long-term variability of solar UV radiation may have an effect on climate via the response to UV of the phytoplankton in the oceans' euphotic zone, which is believed to play a fundamental role in the global control of ocean uptake and release of CO_2.

5.3 Atmospheric Ionization and Global Electric Circuit

The flux of galactic cosmic rays is the dominant source of continuous ionization in the troposphere and the lower stratosphere; because of the modulation of this flux by the interplanetary magnetic field, this ionization is variable. Fig. 1 depicts several possible solar-controlled ways in which the ionization can change. This ionization determines atmospheric conductivity, which in turn regulates the clear-air vertical electric field and electric current between the ionosphere and ground. The latter, at high latitudes, is magnetically connected to the solar wind and its electric potential is controlled by the IMF The electrical connections in the system polar ionosphere/ mid-latitude and equatorial ionosphere/ atmosphere/ ground is called the "global electric circuit". Tinsley has proposed a mechanism (e.g., Tinsley and Heelis, 1993) connecting atmospheric electricity and the rate of contact ice nucleation in clouds, which can operate on a time scale of hours. This theory is based on the fact that the rate at which the charging of water droplets proceeds will depend on the vertical atmospheric current and therefore be responsive to both the cosmic ray flux and the local ionospheric potential. Tinsley's "electrofreezing" mechanism postulates that variations in the amount of such charge affect the rate of initial ice generation at the top of clouds, with ensuing effects on cloud formation.

6. FINAL THOUGHTS

Achieving a scientific understanding of the inner workings of the terrestrial environment is one of the most difficult and ambitious endeavors of humankind, rivaling in complexity the harnessing of nuclear energy, the conquest of space, and the understanding of the human brain. Unfortunately, the more we learn about the environment, climate and anthropogenic effects, the more political problems emerge. Lawyers, judges and politicians are expected to render verdicts and make decisions on the basis of scientific concepts they grasp poorly and without a clear understanding of the scientific method, the inherent experimental uncertainties, and the natural limitations of scientific predictability.

One of the key tasks in climate research at the present time is to learn to separate anthropogenic effects from natural change. Solar variability is one possible cause of natural change in addition to other external phenomena such as volcanic eruptions and, on a long-term scale, orbital changes, celestial body impacts and tectonic plate motion. The unpredictable intrinsic variations of a highly non-linear system such as the atmosphere also must be counted as a "natural" change. Thus, before far-reaching policy decisions are made concerning the impact of future climatic change on human society (and the impact of human activity on climate change), a scientific understanding of possible external influences, however minor at first sight, is of crucial importance. We are still very far from achieving this goal. A truly interdisciplinary, integrated cooperative effort must be developed in STEP in which solar physicists, atmospheric scientists and space physicists work closely together, in coordination with other international programs such as the International Geosphere-Biosphere Program.

Acknowledgments. The author is supported for this work by grant ATM 92-12638 from the Division of Atmospheric Sciences of the National Science Foundation and grant NAGW-1342 from the Space Physics Division of the National Aeronautics and Space Administration. An invitation from the European Science Foundation to present this review at the Workshop on Solar Output and Climate in the Holocene (Bologna, 1-3 April, 1993) was greatly appreciated.

REFERENCES

Anderson, R. Y. (1992): Possible connection between surface winds, solar activity and the Earth's magnetic field. Nature 358, 51-53.

Baker, D. N., Blake, J. B., Gorney, D. J., Higby, P. R., Klebesadel, R. W., and King, J. H. (1987): Highly relativistic electrons: A role in coupling to the middle atmosphere. Geophys. Res. Lett. 14, 1027.

Bucha, V. (1988): Influence of solar activity on atmospheric circulation types. Ann. Geophys. 6, 513-522.

Clayton H. H. (1923): World Weather. Mac Millan New York.

Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., Hvidberg, C. S., Steffensen, J. P., Sveinbjörnsdottir, A. E., Jouzel, J., and Bond, g. (1993): Evidence for general instability of past climate from a 250 kyr ice-core record. Nature 364, 218-220.

Donnelly, R. F. (1991): Solar UV spectrum irradiance variations. J. Geomag. Geoelectr. 43, Suppl., 835-842.

Eddy, J. (1976): The Maunder Minimum. Science 192, No. 4245, 1189-1202.

Friis-Christensen, E., and Lassen, K. (1991): Length of the solar cycle: An indicator of solar activity closely associated with climate. Science 254, 698-700.

Friis-Christensen, E., and Lassen, K. (1993): Two and a half centuries of solar activity variations and a possible association with global temperature. Danish Meteorol. Inst. Scientific Report 93-4.

Groveman, B. S., and Landsberg, H. E. (1979): Simulated Northern Hemisphere temperature departures 1579-1880. Geophys. Res. Lett. 6, 767-769.

Hansen, J., and Lebedeff, S. (1987): Global trends of measured surface air temperature. J. Geophys. Res. 92, 13,345-13,372.

Hines, C. O., and Halevy, I. (1977): On the reality and nature of certain Sun-weather correlation. J. Atmos. Sci. 34, 382-404.

Hood, L. L., Jirikowic, J. L., and McCormack, J. P. (1993): Quasi-decadal variability of the stratosphere: Influence of long-term solar ultraviolet variations. J. Atmos. Sci., in press.

Jones, P. D., Raper, S. C. B., Bradley, R. S., Diaz, H. F., Kelly, P. M., and Wigley, T. M. L. (1986): Northern Hemisphere surface air temperature variations: 1851-1984. J. Clim. Appl. Met. 25, 161-179.

Koster, J. J., and Murcray, D. G. (1979): Change in the solar constant between 1968 and 1978. Geophys. Res. Letts. 6, 382-384.

Kyle, H. L., Hoyt, D. V., Hickey, J. R., Maschoff, R. H., and Vallette, B. J. (1993): Nimbus-7 earth radiation budget calibration history - Part I: The solar channels. NASA Reference Publication 1316.

Labitzke, K. (1982): On the interannual variability of the middle straposphere during northern winter. J. Meteor. Soc. Japan 60, 124-139.

Labitzke, K. (1987): Sunspots, the QBO, and the stratospheric temperature in the North Polar region. Geophys. Res. Lett. 14, 534-537.

Labitzke, K., and van Loon, H. (1993): Some recent studies of probable connections between solar and atmospheric variability. Ann. Geophys., in press.

Lean, J., White, O., Skumanich, A., and Rind, Z. D. (1993): Estimating solar irradiance during the Maunder minimum. Paper delivered at IAU Symposium The Sun as a variable star, Boulder, Colorado, June 1993.

Mendoza, B.. Pérez-Enríquez, R., and Alvarez-Madrigal, M. (1991): Analysis of solar activity conditions during periods of El Niño events. Ann. Geophys. 9, 54-59.

Nesme-Ribes, E. (1993): The Maunder Minimum and the deepest phase of the Little Ice Age: a causal relationship or a coincidence? In: B. Gläser (ed.): Solar output and climate during the Holocene. Publication of the European Science Foundation, in press.

Peréz-Enríquez, R., Mendoza, B., and Alvarez-Madrigal, M. (1989): Solar activity and El Niño: The auroral connection. Il Nuovo Cim. 12C, 223-230.

Pittock, A. B. (1979): Solar cycles and the weather: successful experiments in autosuggestion? In: McCormac, B. M. , and Seliga, T. A. (eds.): Solar-Terrestrial influences on weather and climate, D. Reidel Publ. Co., Dordrecht, Holland, Boston, USA, London, England, 181-191.

Pudovkin, M. I., and Babushkina, S. V. (1992): Atmospheric transparency variations associated with geomagnetic disturbances. J. Atm. and Terr. Electr. 54, 1135-1138.

Reid, G. C. (1991): Solar total irradiance variations and the global sea surface temperature record. J. Geophys. Res. 96, 2835-2844.

Rind, D., and Overpeck, J. (1993): Hypothesized causes of decadal-to-century climate variability: Climate model results. Quat. Sci. Rev., in press.

Roberts, W. O. (1975): Relationship between solar activity and climate change. In: Bandeen, W. R., and Maran, S. P. (eds.): Goddard Space Flight Center Special Report NASA SP-366.

Roederer, J. G. (1992): The Solar-Terrestrial Energy Programme 1990-1997. COSPAR Info. Bull. 123, 47-54.

Salby, M. L., and Shea, D. J. (1991): Correlations between solar activity and the atmosphere: an unphysical explanation. J. Geophys. Res. 96, 22,579-22,595.

Schuurmans, C. J. E. (1991): Changes of the coupled troposphere and lower stratosphere after solar activity events. J. Geomag. Geoelectr. 43, 767-773.

Teitelbaum, H., and Bauer, P. (1990): Stratospheric temperature eleven year variations: solar cycle influence or stratospheric effect? Ann. Geophys. 8, 239.

Tinsley, B. A., and Heelis, R. A. (1993): Correlations of atmospheric dynamics with solar activity: Evidence for a connection via the solar wind, atmospheric electricity, and cloud microphysics. J. Geophys. Res. 98, 10,375-10,381.

White, O. R., Skumanich, A., Lean, J., Livingston, W. C., and Keil, S. L. (1992): The sun in a noncycling state. Publ. of the Astron. Soc. of the Pacific 104, 1139-1143.

Wilcox, J. M., Scherrer, P. H., Svalgaard L., Roberts, W. O., Olson, R. H., and Jenne, R. L. (1974): Influence of solar magnetic sector structure on terrestrial atmospheric vorticity. J. Atmos. Sci. 31, 581-588.

Willson, R. C. and Hudson, H. S. (1988): Solar luminosity variations in solar cycle 21. Nature 332, 810-812.