PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Solar Luminosity Variations over Timescales of Days to the past Few Solar Cycles

P. Foukal

Phil. Trans. R. Soc. Lond. A 1990 **330**, 591-599
doi: 10.1098/rsta.1990.0040

Phil. Trans. R. Soc. Lond. A **330**, 591–599 (1990)
Printed in Great Britain

Solar luminosity variations over timescales of days to the past few solar cycles

By P. Foukal

Cambridge Research and Instrumentation Inc., 21 *Erie Street, Cambridge, Massachusetts* 02139, *U.S.A.*

[Plate 1]

Dips in the total irradiance of up to 0.2% and lasting typically 10 days are now well known to be caused by the transit of dark sunspots across the photospheric disc. The large bright magnetic faculae usually associated with spots cause irradiance increases of comparable magnitude although the form of their signal is more subtle. Radiometry from five satellites beginning in late 1978 indicates a minimum in irradiance at the epoch of lowest magnetic activity between solar cycles 21 and 22. Analysis of these radiometric measurements indicates that this irradiance decline between about 1981 and 1986 was caused mainly by decay in the excess radiation of bright faculae in the magnetic network outside of active regions. Empirical models of irradiance modulation extending back to 1874 indicate that the Sun is typically about 0.05% brighter at activity maximum than at minimum.

1. Introduction

Scarcely 10 years ago, it was still unclear whether the total irradiance, S (the 'solar constant'), varied over timescales accessible to direct measurement. Since then, radiometers flown on the *Nimbus*-7 and *Solar Maximum Mission* (*SMM*) satellites have provided us with essentially continuous daily coverage beginning in late 1978. The similarity of the main features in these two concurrent databases gives us confidence that we now have access to the behaviour of the total solar irradiance over all timescales from a few minutes to essentially one complete solar activity cycle.

The results are of interest to solar and stellar astronomers, as well as to climate physicists. For instance, the variations on timescales of a few minutes contain a contribution from modes of approximately 5 min period associated with acoustic oscillations of the solar convection zone and interior. Analysis of these data has proven to be of interest in helioseismological studies of the rotation of the solar interior, of the damping of the acoustic modes, and of possible structure variations in the solar interior over the 11-year cycle that is discussed by D. O. Gough (this Symposium).

In this paper, I concentrate on the longer-term irradiance variations that seem to be of most direct interest in studies of climate. I first discuss the direct irradiance measurements made from satellites over the past decade, and what they seem to tell us about the main sources of solar luminosity variation. The physical interpretation of these variations has been reviewed recently (Foukal 1987, 1988; Spruit 1988) and I only mention the most recent developments here. I then present a simple empirical model of the irradiance variations developed together with J. Lean (Foukal & Lean 1986; Lean & Foukal 1988). This model is based on the understanding gained from cycle 21, and I apply it to reconstructing the variations expected to have been caused by solar activity during cycles 19 and 20. An adaptation of this model is also described that enables us to estimate what the solar cycle variation might have been back to 1874.

Photometric studies of younger Sun-like stars provide us with the most direct information on the behaviour of the solar luminosity and its variability over longer palaeoclimatological timescales. An excellent review of this topic has been given recently by Baliunas (1988).

2. Radiometry of the total solar irradiance 1978–88

The best existing data on solar irradiance variation are shown in figure 1. The longest time series is from the Earth Radiation Budget (ERB) radiometer on the *Nimbus*-7 satellite. These daily data begin in November 1978 and continue to the present. The experiment and reduction procedures have been described by Hickey *et al.* (1988). The data shown here extend through October 1988.

The most precise measurements are those made with the Active Cavity Radiometer Irradiance Monitor (ACRIM) radiometer on the *SMM* spacecraft beginning in early 1980 and also continuing to the present, with a short interruption between late 1983 and early 1984 (Willson & Hudson 1988). The ACRIM data shown in figure 1 extend to late 1987. Three additional sets of radiometric measurements, from the *NOAA*-9, *NOAA*-10 and *ERBS* satellites, have been available since late 1984 (Mecherikunnel *et al.* 1988). The *ERBS* data to early 1988 are shown here. These measurements are only obtained a few times per month, but they provide useful checks of slow irradiance trends.

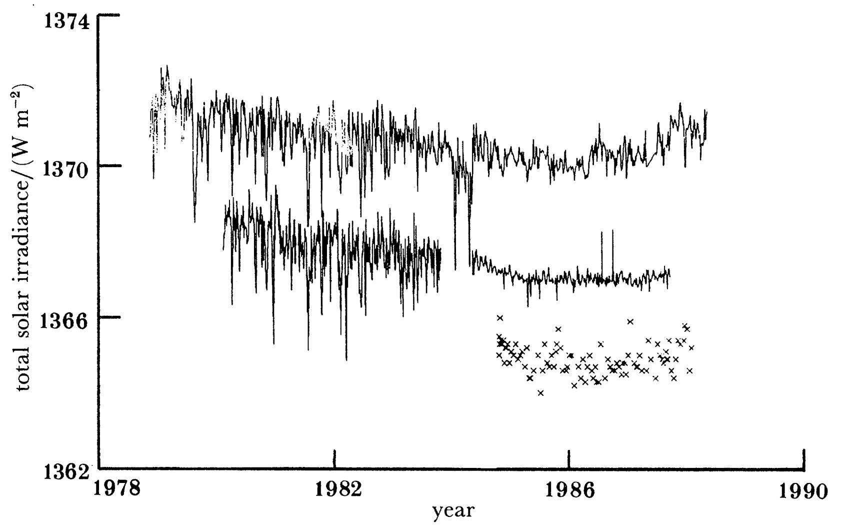


Figure 1. Total solar irradiance measurements from the *Nimbus*-7 (top), *SMM* (middle), and ERBS (bottom) satellites. The vertical offsets correspond to real differences of about 0.2 % in absolute calibration between each of the three radiometers.

The vertical offset of roughly 0.2 % between each of the three data-sets is caused by a real disagreement of absolute calibration between the three radiometers. However, the precision of each is much higher, as shown by the agreement in the short-term variations seen by the ERB and ACRIM.

A high-frequency variation on timescales of about 10 days, with peak-to-peak amplitude reaching about 0.3 % is one prominent feature of the irradiance data shown in figure 1. This is caused by dark spots and bright magnetic faculae rotating across the disc, as discussed below. The second is the gradual decline of the irradiance until early 1986, and then its gradual recovery. The decline observed by both the ERB and ACRIM between 1980 and 1986 amounts to somewhat less than 0.1 %. This variation is of greatest interest to climate studies, and its source is currently the object of intensive study.

3. Influence of spots and faculae

The large dips in total irradiance seen in figure 1 are easy to identify with the influence of spots. The irradiance variation expected from changes in the area of dark spots on the solar disc can be calculated from their photometric contrasts and from daily data on their coordinates and areas (Willson *et al.* 1981). As seen in figure 2 the calculated function accounts for the timing and depth of the larger dips rather well. The missing energy not radiated by spots is most probably stored in a small increase in the thermal and potential energy of the convecting layers outside the spot (Spruit 1982; Foukal *et al.* 1983). Thus it seems to represent a true luminosity decrease, not merely a reradiation into a different solid angle.

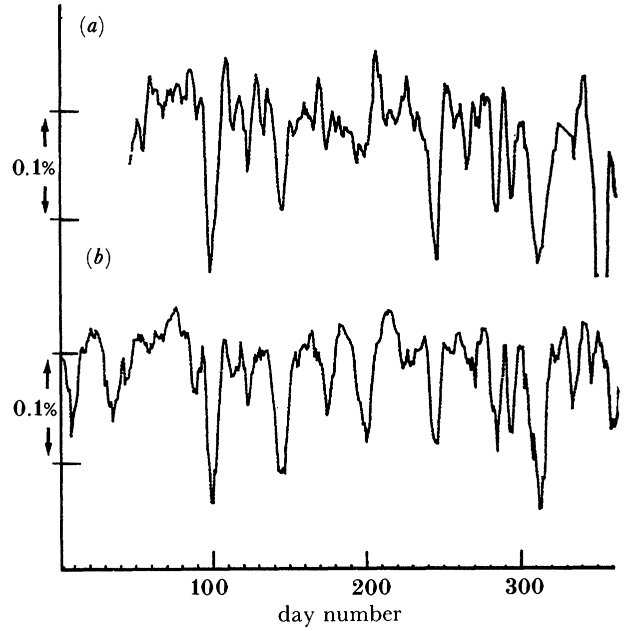


Figure 2. Comparison of the variations in measured solar irradiance for 1980 from (*a*) the *SMM* and the variations calculated from an empirical model (*b*) based on measured photometric contrasts of spots and their daily areas and disc coordinates. (From Foukal 1987.)

But this is not the only effect of solar magnetic activity on the total irradiance. If we subtract the calculated sunspot blocking function, P_s, from the measured irradiance, the residuals are still comparable in magnitude with the variations in P_s. Figure 3 illustrates such a subtraction for ACRIM data in 1980. We see that the residuals correlate well with variations in 205 nm ultraviolet (UV) flux. These UV variations are known to be caused primarily by rotation across the solar disc of bright magnetic faculae (see, for example, Lean 1987).

Figure 4, plate 1, shows the appearance of the solar photosphere and chromosphere at times of low and high magnetic activity. The photosphere is the level at which over 99 % of the solar

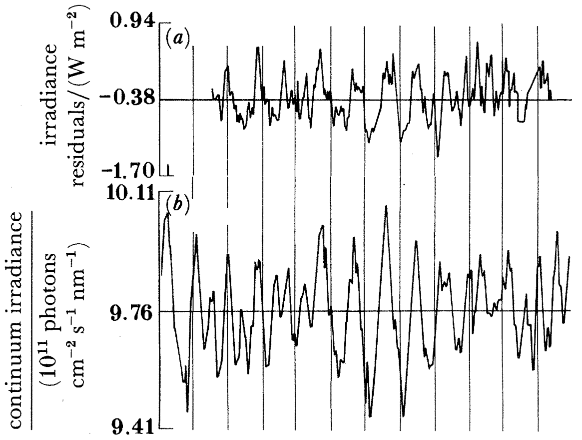


FIGURE 3. (*a*) Irradiance residuals found after removing the sunspot blocking effect from the ACRIM radiometry for 1980; (*b*) 205 nm UV continuum irradiance measured from the *Nimbus*-7 spacecraft. (From Foukal & Lean 1986.)

radiation originates. It is seen that as magnetic activity increases, more dark spots and also more bright magnetic faculae appear. The largest faculae are found around the active regions. But outside of these is a network of comparably bright, although smaller structures, that cover the whole disc. These are seen better in the chromosphere where they are brighter.

There is no real distinction between the bright faculae and the network except for the size. They both correspond to regions of intense, radial, small-scale magnetic fields. Their excess brightness in photospheric radiations is most likely caused by the relatively low plasma pressure in these intense magnetic flux tubes. This lowers the opacity and enables radiation to escape from deeper and hotter layers (Spruit 1976; Deinzer *et al.* 1984).

The contrast of faculae is quite low in broadband visible light, where most of their energy is emitted. This makes it relatively difficult to determine whether their signal can account for the amplitude of the irradiance residuals seen in figure 3. However, photometry of active regions (see, for example, Chapman & Meyer 1987) indicates that the relatively large areas covered by faculae in active regions enable them to make an irradiance contribution comparable with that of spots.

A clean test of faculae as the cause of the irradiance residuals is provided by their cross-correlation with the sunspot blocking functions P_s, and with the 205 nm flux (Foukal & Lean 1986). This cross-correlation shows that the residuals S-P_s are caused by solar structures that last typically at least three solar rotations, thus about two rotations longer than do spots. This is consistent with the well-known long lifetime of faculae, and argues against errors in the P_s function as the main cause of the residuals.

4. Detection and understanding of slower irradiance variations

Identification of the facular signal in the irradiance record seems to be helpful in understanding the nature of the slower variations over timescales from months to the full activity cycle. Thus Foukal & Lean (1988) found a good correlation between slow changes in the irradiance residuals, and in indices of facular area, for the four years between 1981 and 1984. The CaK plage index, one of the two used in that study, is based on the areas of only

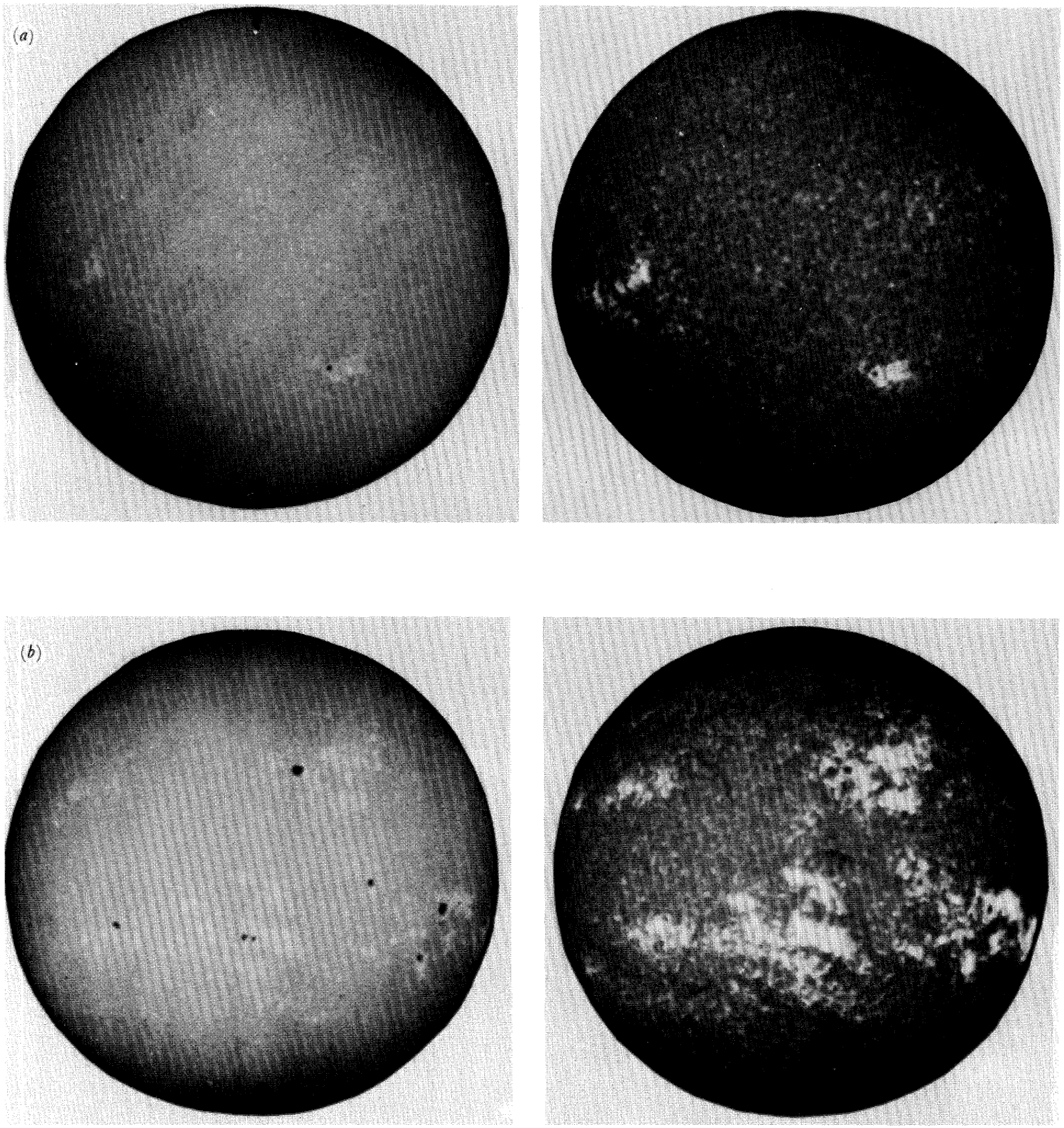


FIGURE 4. Images of the Sun at (*a*) low (10 May 1975) and (*b*) high (7 March 1979) magnetic activity levels. The images on the left are taken in the wing of the Ca II K absorption line, and refer to the highest photospheric levels. Those on the right are in the centre of that line, and show the overlying chromospheric layers. (From Lean 1987.)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

(*Facing p.* 594)

the large faculae in active regions. The second is the He I index that is global in the sense that it includes contributions from all the faculae, including those in the network.

The regressions constructed in that study were then used with daily values of the CaK and He I indices between 1981 and 1984, to reconstruct the irradiance residuals obtained for that period. Figure 5 shows the comparison of the reconstruction with the original data. It is seen that when the global index, He I, is used one can reconstruct not only the slow downtrend, but also the variations of 4–9 month timescale, remarkably well. When the CaK plage index is used as a basis for the reconstruction, the 4–9 month variations are reproduced about equally well, but the ability to reproduce the slow downtrend is lost.

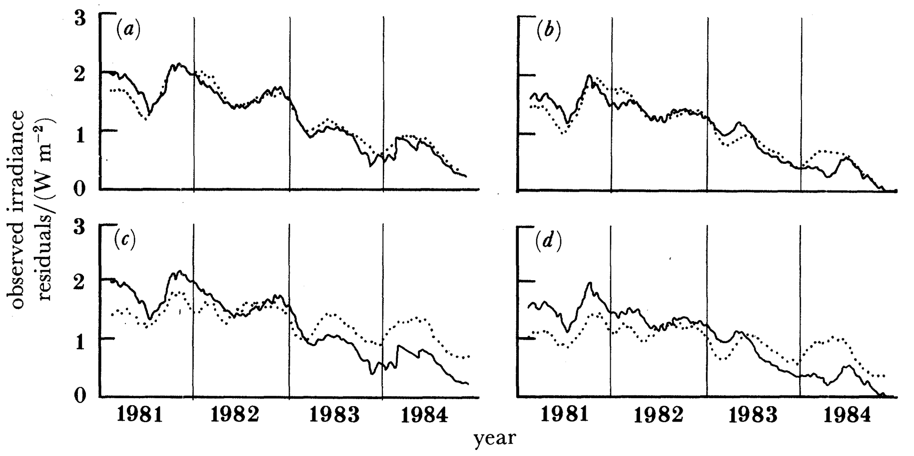


FIGURE 5. Comparison of the observed irradiance residuals (solid lines) for (*a*) and (*c*) the ACRIM data and for (*b*) and (*d*) the ERB data, with the irradiance residuals (dotted lines) calculated by using (*a*), (*b*) the He I and (*c*), (*d*) CaK indices. (From Foukal & Lean 1986.)

This result indicates that the 4–9 month irradiance variations are caused by something that is common to both the global index and the active region index, i.e. the large active regions. This timescale seems to be set by the tendency of extended complexes of activity to erupt and subside (Harvey 1984). The slow decline between 1980 and 1984 must be caused by a mechanism residing outside the active regions.

In principle, this mechanism could be a deep-seated alteration of solar convection or even of processes in the core as some have suggested (Kuhn *et al.* 1988; Gough 1988) and a contribution from such effects cannot be dismissed. But the finding that a single linear relation between facular area and total irradiance can account for both the 4–9 month variations and the slow downtrend indicates that most of the decline is caused by a solar-cycle variation in the surface density of the network faculae outside active regions.

A test that could distinguish between the two kinds of mechanism might be based on comparison of the solar cycle variations in effective temperature at relatively deep photospheric layers (where the enhancement of the network is smallest and effects due to changes in underlying convection should be largest) and at higher layers where the relative importance of these two effects should be reversed.

Data that might be used for such a test have been obtained (Livingston *et al.* 1988) by measurement of small changes in the strength of solar absorption lines in disc-integrated solar light between 1978 and 1988. The results (figure 6) show that lines formed around the

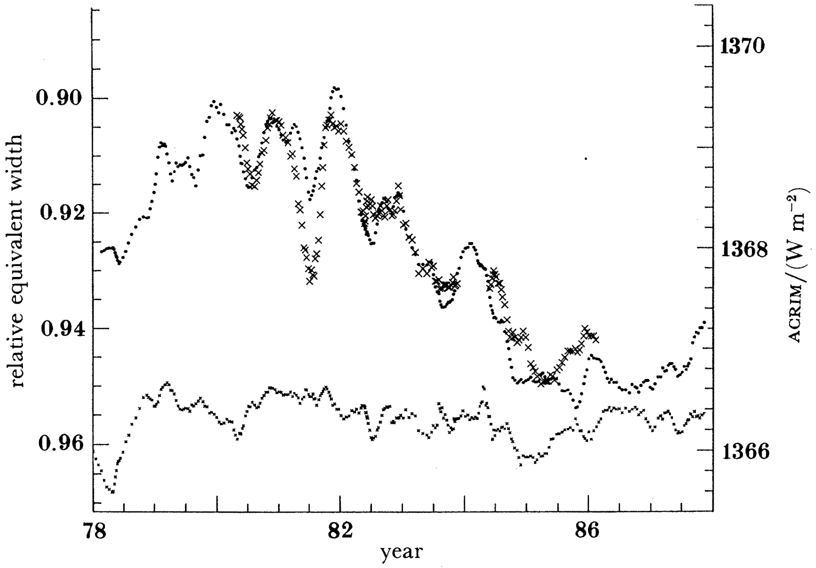


FIGURE 6. Irradiance residuals (crosses), and He I 1083 nm equivalent width (dots). At the bottom is the time series for C I 5380 treated in the same way. (From Livingston *et al.* 1988.)

temperature minimum and in the chromosphere follow the irradiance residuals well. The line C I 5380, formed deeper in the photosphere, shows no significant signal. This result would seem to rule out the deep-seated variations, but one must inquire into the sensitivity of the C I 5380 data to changes in T_{eff}. According to Livingston (personal communication) it turns out to be somewhat inadequate to rule out the 0.07% change in S found by the radiometry between 1980 and 1986.

An argument for the network as the source of total irradiance change over the solar cycle is that lines such as He I 10830 and Lyman α are formed in the non-thermally heated chromosphere and corona. The finding that they show the same relative amplitudes of the 4–9 month variations and of the slow decline between 1980 and 1986 as does the total irradiance, supports the dominant effect of the magnetic network, in which these lines are well known to be enhanced. It is difficult to understand why changes in deep convection or in nuclear processes should affect these radiations in such a similar way.

5. EMPIRICAL MODELLING OF TOTAL IRRADIANCE VARIATIONS CAUSED BY SOLAR ACTIVITY 1874–1988

If I suppose that the bright faculae in active regions and in the network, together with the spots, are the main cause of the irradiance variations observed so far, I can use our model to calculate these variations back as far as indices of spot and facular area are available. Daily values of spot coordinates and areas are on record from 1874. Good global indices of faculae are more difficult to obtain. The daily He I 10830 index begins in 1975. The 10.7 cm microwave emission, which arises primarily from enhanced bremmstrahlung radiation from the denser coronal atmosphere over the magnetic faculae, is available back to 1947.

The variation of S during cycle 21 as predicted from the daily He I values is shown in figure 7, along with the values of the sunspot and facular contributions plotted separately. The

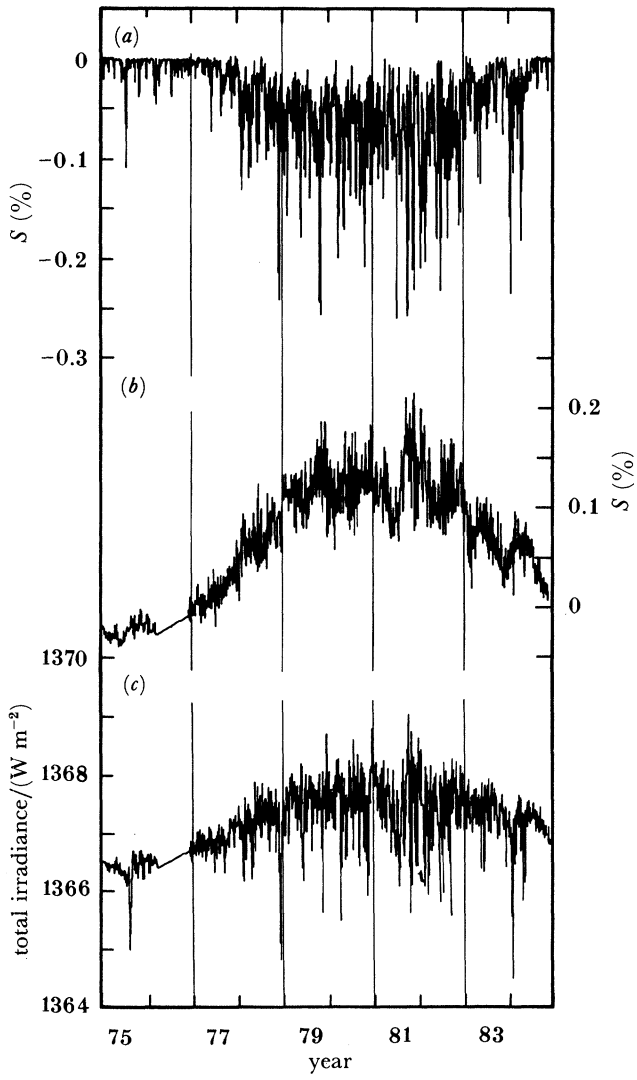


FIGURE 7. Solar cycle 21 behaviour of (*a*) the daily sunspot blocking function P_s, (*b*) the daily facular irradiance excess P_f from the He I index, and (*c*) the reconstructed total solar irradiance. (From Foukal & Lean 1986.)

main finding is that the expected irradiance increases from activity minimum in 1976 to maximum near 1980 by about 0.12 % (Foukal & Lean 1988).

A similar model, but based on the 10.7 cm microwave emissions has been used to estimate the irradiance variation between 1954 and 1984 (Lean & Foukal 1988). The results are shown in figure 8. It is seen that the Sun is consistently brighter at high activity levels. Of most interest is the finding that the calculated irradiance increase of about 0.1 % in cycle 21 was significantly larger than that calculated for cycle 19, which was the largest cycle in the history of reliable spot records (i.e. since about 1850).

This large brightening of the Sun during cycle 21 is caused by the relatively small amplitude of that cycle as measured in spot area compared with its amplitude in other indices, such as the Zurich sunspot number, R_Z, or the microwave flux. That is, during cycle 21 the mean area of a spot tended to be significantly smaller compared with the mean area for all

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

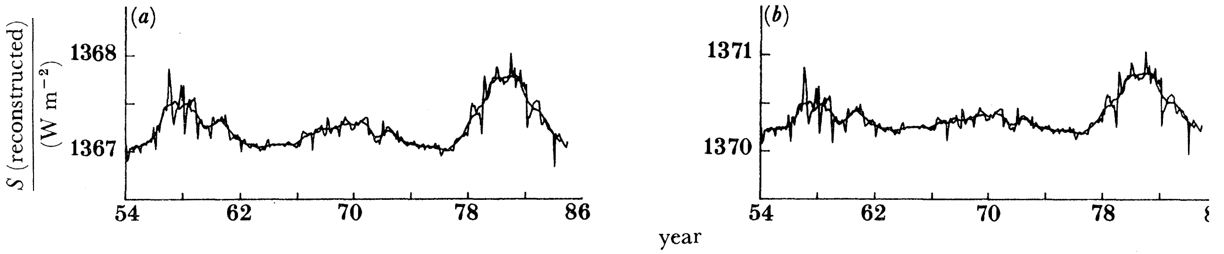


FIGURE 8. Behaviour of the total solar irradiance between 1954 and 1985 reconstructed with (*a*) the ACRIM data and (*b*) ERB data. The heavy line represents a 12-month running mean. The difference in the absolute scales in the two panels arises from a difference in the ACRIM and ERB absolute calibrations. (From Foukal & Lean 1988.)

cycles back to at least 1874. The reason for this aberrant behaviour, which has been recognized only recently (Foukal & Lean 1990) is not understood.

There is no reliable direct index of facular activity for calculation of the irradiance further back than cycle 19. The Greenwich photoheliograph data on faculae exhibit marked anticorrelations with other indices that call into question the reliability of that record, probably because of systematic errors caused by the difficulty of detecting white-light faculae near the limb.

However, our work (Foukal & Lean 1990) indicates that irradiance variations back to 1874 can be calculated by using the fact that variations in the sunspot number, R_Z, are highly correlated with microwave flux (i.e. facular area) variations. If we are interested only in the relative magnitudes of the cycles in facular activity, the spot number seems to give us a reasonable proxy index for faculae. By contrast, the sunspot area, which determines the sunspot irradiance effect, correlates much less well with $F_{10.7}$ since 1947.

A preliminary calculation of irradiance variation extending to 1874, based on P_s and R_Z, indicates a marked brightness increase for the peaks of the four cycles since 1950. This increase can be attributed in part to the general rise of solar activity beginning with cycle 18. But the most remarkable feature is the anomalously large irradiance increase around 1980 caused by the peculiar behaviour of spot areas during cycle 21.

6. SUMMARY AND CONCLUSIONS

Information on the Sun's luminosity behaviour, and our understanding of the mechanisms that seem to cause it, has progressed rapidly in the past 10 years. Relatively short-term, but large-amplitude, variations of total irradiance are caused by sunspots and faculae. Reasonable explanations have been put forward for the irradiance and luminosity variations caused by these large- and small-diameter flux tubes and these are found to be consistent with photometric observations, including recent near-infrared imaging of active regions at the deepest observable photospheric layers (Foukal *et al.* 1989). Still, the fundamental reasons for spot darkness and facular brightness cannot be considered to be well understood, and in particular the possibility of an energetic connection between these active region structures by non-thermal transports (Schatten *et al.* 1987) needs closer attention.

The most significant observational result is the finding that the Sun is somewhat brighter, rather than darker, at high magnetic activity levels. The peak-to-peak variation in cycle 21 seems to have been roughly 0.1 %. The most likely source of this variation is the predominant

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

irradiance contribution of bright faculae over that of dark spots. Although the time-integrated irradiance contributions of spots and faculae in active regions roughly balance, the enhancement and subsequent decay of radiation from the magnetic network seems to cause most of the general irradiance variation over the 11-year cycle. Deeper-seated influences on photospheric heat and non-thermal energy flow are certainly not ruled out, but more accurate observations are required to establish their importance.

It is interesting that this predominance of faculae in controlling solar luminosity variation may not have held for the early Sun. Radick *et al.* (1990) point out from stellar photometry that young, late-type dwarf stars of ages less than 10^9 years exhibit an anticorrelation of brightness with magnetic activity. It would seem that in such stars, the propensity to form large dark spots (rather than bright faculae) is greater than in the present Sun, leading to a net darkening at high activity levels.

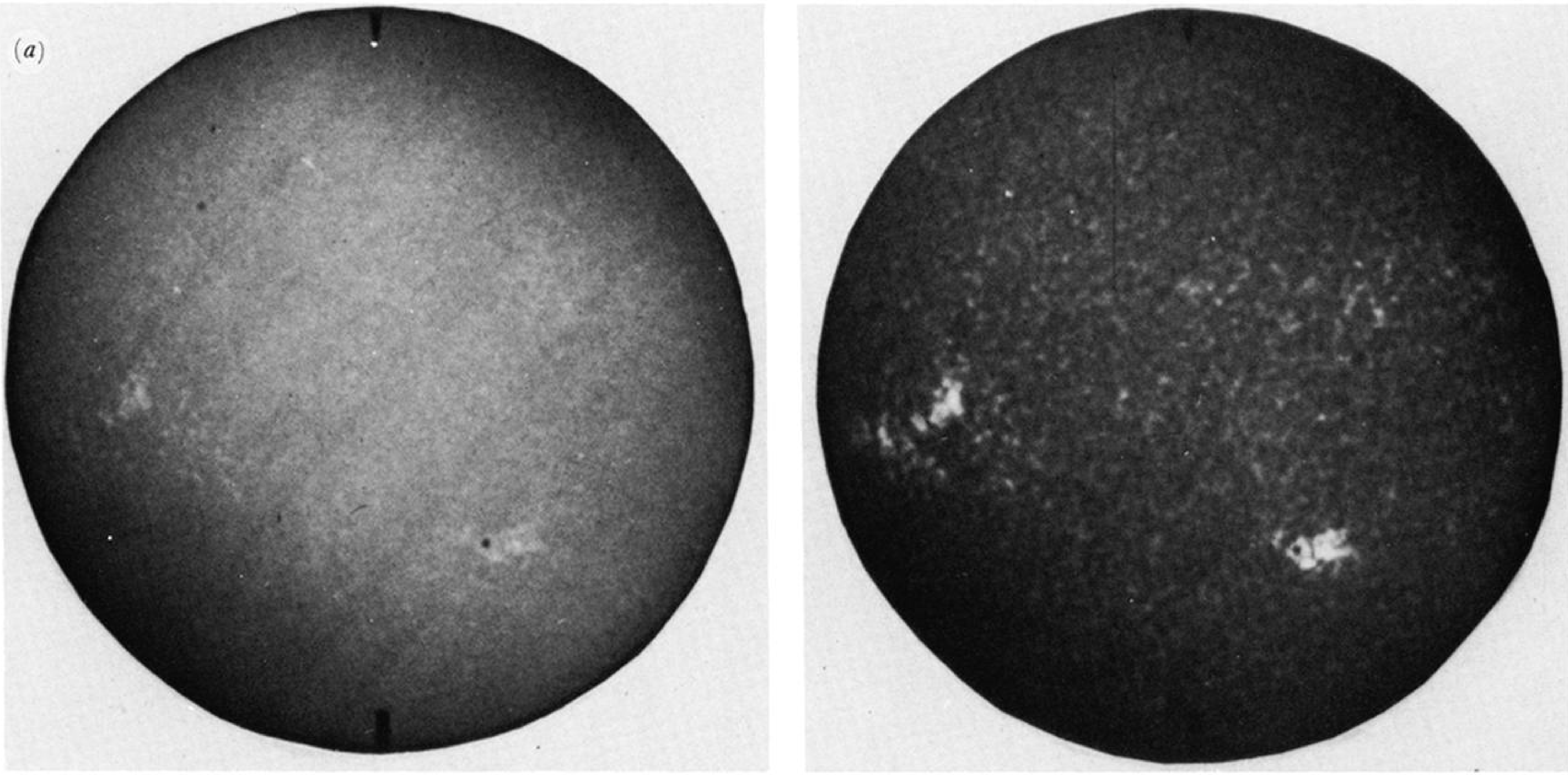
Regarding slow trends in solar irradiance, the most significant new evidence is that the irradiance increase measured in cycle 21 is likely to have been exceptionally large, exceeding even that of cycle 19, the largest sunspot cycle in the history of reliable records. The anomalously low area coverage of spots during cycle 21, compared to that of faculae, seems to have been an unusual occurrence with no precedent back to at least 1874.

More generally, the rise in irradiance signals associated with the last few cycles appears to produce a significant upward trend in solar irradiance beginning about 1950, and culminating in the exceptionally high irradiance levels associated with cycle 21, which peaked around 1980 (Foukal & Lean 1989). Although the mean irradiance increase averaged over these cycles is small, its duration for almost 40 years now may have begun to produce climatic consequences, that could be contributing to the general trend of global warming.

References

Baliunas, S. 1988 *Solar radiative output variation* (ed. P. Foukal), p. 230. (Available from the editor.)
Chapman, G. & Meyer, A. 1986 *Sol. Phys.* **103**, 21.
Deinzer, G., Henzler, G., Schussler, M. & Weisshaar, E. 1984 *Astr. Appl.* **139**, 435.
Foukal, P. 1987 *J. geophys. Res.* **92**, 801.
Foukal, P. & Lean, J. 1986 *Astrophys. J.* **302**, 826.
Foukal, P. & Lean, J. 1988 *Astrophys. J.* **328**, 347.
Foukal, P. & Lean, J. 1990 *Science, Wash.* (In the press.)
Foukal, P., Fowler, L. & Livshits, M. 1983 *Astrophys. J.* **267**, 863.
Foukal, P., Little, R. & Mooney, J. 1989 *Astrophys. J. Lett.* **336**, 33.
Gilliland, R. 1988 *Solar radiative output variation* (ed. P. Foukal), p. 289. (Available from the editor.)
Harvey, J. 1984 In *Solar irradiance variations on active region time scales* (ed. B. Labonte *et al.*), p. 197. NASA CP2310.
Hickey, J., Alton, B., Kyle, L. & Hoyt, D. 1988 *Space Sci. Rev.* **48**, 321–342.
Kuhn, J., Libbrecht, K. & Dicke, R. 1988 *Science, Wash.* **242**, 908.
Lean, J. 1987 *J. geophys. Res.* D **92**, 839.
Lean, J. & Foukal, P. 1988 *Science, Wash.* **240**, 906.
Livingston, W., Wallace, L. & White, O. 1988 *Science, Wash.* **240**, 1765.
Mecherikunnel, A., Lee, R., Kyle, L. & Major, E. 1988 *J. geophys. Res.* D **93**, 9503.
Radick, R., Lockwood, W. & Baliunas, S. 1990 (In preparation.)
Schatten, K., Mayr, H. & Omidvar, D. 1987 *J. geophys. Res.* **92**, 818.
Spruit, H. 1976 *Sol. Phys.* **50**, 269.
Spruit, H. 1982 *Astr. Appl.* **108**, 348, 356.
Spruit, H. 1988 *Solar radiative output variation* (ed. P. Foukal), p. 254. (Available from the editor.)
Willson, R., Gulkis, S., Janssen, M., Hudson, H. & Chapman, G. 1981 *Science, Wash.* **211**, 700.
Willson, R. & Hudson, H. 1988 *Nature, Lond.* **332**, 810.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES



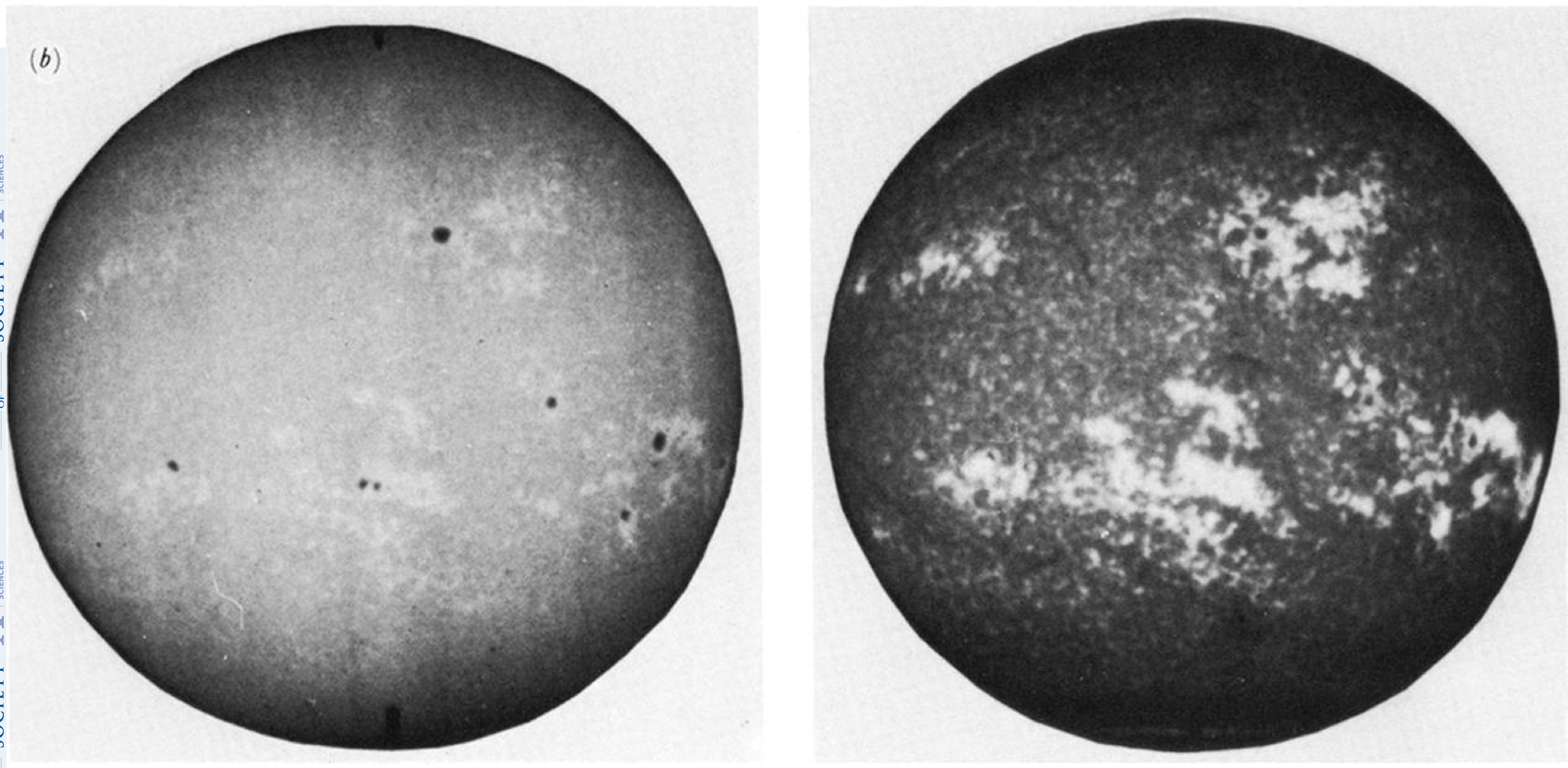


FIGURE 4. Images of the Sun at (*a*) low (10 May 1975) and (*b*) high (7 March 1979) magnetic activity levels. The images on the left are taken in the wing of the Ca II K absorption line, and refer to the highest photospheric levels. Those on the right are in the centre of that line, and show the overlying chromospheric layers. (From Lean 1987.)