Rev. Mexicana Astron. Astrof., **6**, 359-362 (1981)

SOLAR VARIABILITY

S. Sofia

Laboratory for Planetary Atmospheres
Goddard Space Flight Center, NASA

RESUMEN

Las mediciones directas llevadas a cabo con detectores espaciales han mostrado variaciones de la constante solar de 0.2 por ciento, en escalas de tiempo de días a decenas de días. Se argumenta que estos cambios no reflejan necesariamente cambios en la luminosidad solar, y que en general, mediante mediciones directas no se han podido establecer (ni excluir) cambios de la luminosidad solar en escalas de tiempo mayores. Sin embargo, por técnicas indirectas, en particular mediciones del radio solar, se establece que han ocurrido variaciones en la luminosidad solar de hasta ~ 0.7 por ciento en un lapso de decenas o centenas de años.

ABSTRACT

Direct measurements carried out by space-borne detectors have shown variations of the solar constant at the 0.2 percent level, having timescales of days to tens of days. It is argued that these changes do not necessarily reflect variations of the solar luminosity, and that in general, direct measurements have not yet been able to establish (or exclude) longer timescale solar luminosity changes. However, indirect techniques, radius measurements in particular, indicate that solar luminosity variations of up to ~ 0.7 percent have occurred within tens to hundreds of years.

Key words: SUN – VARIABILITY

I. INTRODUCTION

By analogy with stellar concepts, I shall assume in this paper that *solar variability* means the occurrence of detectable changes of the energy output of the Sun on non-evolutionary timescales. In particular, I will consider frequency-integrated energy, so that the relevant quantity is the solar luminosity, $L_\odot$.

The solar luminosity and its variation can be measured in two ways. Direct techniques measure the amount of solar energy striking normally one cm^2 of a detector at earth per second (also known as the *solar constant* or solar *irradiance*), S, and then inferring $L_\odot$ from the relation

$$L_\odot = 4\pi d^2 S \quad (1)$$

where d is the Earth-Sun distance. The second way is by indirect techniques, which measure the solar physical parameters that determine the luminosity as given, for example, by the Stefan-Boltzmann law.

We shall show in this paper that direct techniques have shown that the solar energy output varies on timescales < 1 year, whereas indirect measurements have shown variations having a magnitude and duration large enough to have climatic consequences.

II. DIRECT MEASUREMENTS

Direct measurements of the solar constant with ground-based instruments have been carried out throughout most of this century (Abbot 1966; Angione 1981). Variations at the level of 1-2 percent were detected in these measurements. Because of the very large corrections which must be applied to ground-based observations in order to account for the (variable) atmospheric absorption, the reality of these variations is not widely accepted.

With the advent of the space program, extensive solar constant measurements from instruments placed above the Earth's atmosphere have been carried out in the last two decades. Again, variations at the level of ~ 0.5 percent have been detected. However, because of instrumental stability and calibration difficulties, these measurements are not beyond question, and they are consistent with no real change at all (cf. Frohlich 1981).

For timescales of a few months, however, the stability problems are not very serious. Independent measurements of S made from two spacecrafts. During 5 months in 1980 showed (in excellent agreement with each other) that variations at the level $\lesssim 0.2$ percent occurred with timescales of days. The measurements were made with *active cavity radiometers* flying on board Nimbus 7 (Hickey 1981) and the Solar Maximum Mission (Willson 1981).

The relationship (1) between S and $L_\odot$ is only valid if

(a) no absorption takes place between the Sun and the Earth, and (b) the solar radiation is isotropic. All available evidence supports the validity of (a). However, the larger dips in S are always observed when the Zurich sunspot number is at a maximum, and indeed a large portion of the variability can be explained in terms of sunspot effects (Willson *et al.* 1981). This means that the assumption of isotropy of the radiation is not valid, since an observer from a direction where the sunspots were not visible would not see the dip. Moreover, since the sunspots may only temporarily hamper the diffusion of some radiation, (say by temporarily storing the energy and then releasing it when the spot is gone) these observations do not allow us to infer variations in $L_\odot$ on timescales of years. Consequently, while direct measurements show solar variability in short timescales, they have not been able to prove or reject long-term variability at a meaningful level.

III. INDIRECT MEASUREMENTS

Indirect measurements of variations in the solar luminosity are based on the Stefan-Boltzmann law,

$$L_\odot = 4\pi\sigma R_\odot^2 T^4 \qquad (2)$$

If the Sun were a black-body, a determination of its radius and temperature would determine $L_\odot$. However, since the Sun is not a black-body, T in equation (2) must be understood as an *effective temperature*. By definition, then, T can only be determined from measurements of $L_\odot$ and $R_\odot$ and so this technique *cannot* be used as a means of determining $L_\odot$. However,

$$\frac{dL_\odot}{L_\odot} = 2\,\frac{dR_\odot}{R_\odot} = 4\,\frac{dT}{T}\ , \qquad (3)$$

and *variations* of the luminosity can be followed as long as it is possible to measure any temperature variation which is *related* to a variation of the effective temperature, as well as variations of the radius derived by self-consistent methods.

In general, $dT/T \gg dR_\odot/R_\odot$, and consequently, if we could obtain reliable measurements of dT/T, the radius variation could be safely ignored. Livingston (1978) by monitoring the strength of the λ5380.3 line of C I (a weak line formed in approximately the same level of the photosphere as the nearby continuum) thought that he was in fact monitoring variations in the effective temperature. However, more recently, Livingston (1981) upon monitoring the strength of other weak lines with different temperature dependences, found that different lines require temperature changes that differ both in magnitude and sign. It now appears that the interpretation of these spectroscopic temperature changes in terms of variations of the effective temperature of the Sun is strongly model-dependent, and current models do not provide reliable interpretations.

Finally, the suggestion has been made that the change $dL_\odot$ is ultimately caused by a change $dR_\odot$, which also causes a change dT (Sofia *et al.* 1979). In other words,

$$\frac{dT}{T} = f\left(\frac{dR_\odot}{R_\odot}\right)\ ,$$

and hence

$$\frac{dL_\odot}{L_\odot} = \phi\left(\frac{dR_\odot}{R_\odot}\right)\ , \qquad (4)$$

where f and ϕ are unknown functions. A physical explanation of this relationship is given, for example, in Sofia and Endal (1980).

For small changes,

$$\frac{dL_\odot}{L_\odot} \approx W\,\frac{dR_\odot}{R_\odot}\ , \qquad (5)$$

where W is a coefficient which varies slowly as a function of the time elapsed since the perturbation occurred.

The values of W for timescales of 10^4 years can be obtained quite generally on the basis of virial theorem arguments (Sofia and Endal 1980). However, for shorter timescales (years to hundreds of years) W must be obtained by numerical modeling. Unlike the temperature modeling, though, the modeling required to obtain W is based on structure and evolution codes whose physics and numerical behavior are well understood. Extensive (and independent) calculations have lead to the following values of W (for timescales of years to hundreds of years) according to the location of the perturbation within the Sun:

a) For perturbations in the shallow layers, for example, perturbations of the mixing length of convection (α-type perturbations), $W < 10^{-3}$ (Sofia and Chan 1981; Sweigart 1981; Twigg and Endal 1981).

b) For perturbations in the deep layers of the convective envelope (β-mechanisms), $W \cong 0.1$ (Sofia and Chan 1981).

c) For perturbations in the radiative interior of the Sun, $W \cong -0.2$ (Gough 1981).

IV. SOLAR RADIUS MEASUREMENTS

Measurements of the Solar radius may be a very efficient means of determining solar variability. This would not be the case if the α-mechanism were the one operating in the Sun, since essentially no radial changes would accompany the luminosity changes. On the other hand, if changes in radius are detected, then the α-mechanism cannot be the one that operates in the Sun. In this case, either the β mechanism, or an interior perturbation operates. With the interior perturbation, it is difficult to visualize scenarios capable of producing changes that occur more frequently than million of years apart. Since, as will be discussed later, changes in radius occur on timescales of tens of years, the β-mechanism seems favored. An example of a β-mechanism is given by the variable magnetic pressure at the base of the convective zone produced by the solar dynamo during the solar cycle. Thus, it would appear that $W \cong 0.1$. This result is important on two accounts. First, by a variety of techniques it is possible to reconstruct the behavior of the solar radius for the past two centuries or more. Second, the SCLERA telescope (Hill 1981) is capable of measuring $dR_\odot/R_\odot$ with an accuracy of $\sim 10^{-6}$. This means that $dL_\odot/L_\odot$ can be monitored with an accuracy of 10^{-5}, far better than any current direct or indirect technique.

V. MEASUREMENTS OF THE SOLAR RADIUS

In the past, variations of the solar radius have been made by three different techniques. The first technique involves the use of meridian circles (Sofia *et al.* 1979; Eddy and Boornazian 1979). While Sofia *et al.* derived a radius change $\leqslant 0\farcs25$ for the last century, Eddy and Boornazian claimed to have detected a secular radius decrease of 1″ per century. The second technique makes use of timings of transits of Mercury in front of the solar disk. Shapiro (1980) and Parkinson *et al.* (1980) used this information to successfully disprove the large secular contraction claimed by Eddy and Boornazian for the past two-and a half centuries. However, the technique is plagued by timing uncertainties produced by observational difficulties, and the results are only accurate to $\sim \pm 1''$. The third technique derives solar radius information from timings of the duration of total solar eclipses. Two versions have been proposed for this technique. One (Dunham *et al.* 1980) makes use of observations obtained near the edge of the path of totality. The other (Parkinson *et al.* 1980) uses central timings. The edge timings are a more reliable means of monitoring the solar size on two accounts. First, timing errors, inevitable in the early observations, have much more serious consequences in the central timings than at the edge. For example, 1 sec. error in the duration of totality at the path center will typically cause 0″.3 error in the diameter determination. However, sufficiently near the edge, errors of a few seconds produce negligible diameter errors (cf. Dunham *et al.* 1980, Fig. 1). Second, because of the geometry of lunar librations, edge, second, and third contacts, which occur in the polar region of the Sun and Moon, have a high probability of occurring at the same lunar features. The lunar profile (as given by the Watts tables) is not known more accurately than ±0″.2. Since eclipses where contacts occur in the same features give results independent of the exact depth of the lunar valleys, they avoid the error mentioned above as happened, for example, in the 1925 and 1979 eclipses. However, the likelihood of finding central contacts occurring at the same lunar features is negligible, and so all central timing results contain the lunar profile errors. We have applied our technique to several total solar eclipses, and the results are shown in Table 1.

TABLE 1

CORRECTIONS TO THE STANDARD SOLAR RADIUS DERIVED FROM PATH-EDGE OBSERVATIONS

Year	$\Delta R_\odot$[a]	*Error*
1715	+0″.52	±0″.2
1925	+0.62	0.08
1976	−0.23	0.14
1979	−0.08	0.09
1980	−0.03	0.04

a. $R_\odot$ = 959″.63

For the future, path-edge timing observations must be continued (the lunar theory is not adequate to predict the path center well enough so that one edge observation will suffice). The limit of accuracy of this technique (dictated by the fact that the Sun's edge is not infinitely sharp) is ±0″.05. This limitation, combined with the fact that total solar eclipses are rare, makes the eclipse timing techniques adequate to carry out a finely tuned monitoring of the solar cycle. New techniques (as developed by Hill, at SCLERA, and as currently being developed at the High Altitude Observatory) have potential accuracies of ±0″.001, and measurements can be carried out daily. If this potential accuracy is borne out by practical experience, the new techniques are the logical way to monitor $R_\odot$ in the future. However, since these new techniques are very complex, they may contain unknown systematic errors of magnitude far larger than the ±0″.001 random errors. Consequently, in the near future the eclipse timing observations must be also continued to check the validity of the new techniques.

VI. SUMMARY AND CONCLUSIONS

Space-born instruments have detected variations of the solar constant at the level of ~0.2 percent, with a timescale of days to tens of days. From these direct measurements we are unable at the present time to determine whether or not variations on timescales of years to tens of years, which are of important climatic consequences and require structural changes in the Sun, occur. However, measurements of the solar radius do indicate a structural change of up to 0."7. If the physical origin of this change is located at the base of the convection zone, then W ~ 0.1, and this would imply variations of the solar luminosity of up to 0.7 percent. If these changes are found to be long-lived, as we would expect such deep seated changes to be, they undoubtedly have profound effects on the variations of the Earth's climate.

I would like to thank K. Schatten for a critical reading of the manuscript.

REFERENCES

Abbot, C. G. 1966, *Smithsonian Misc. Collect.*,448, 7.
Angione, R. J. 1981, in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Dunham, D. W., Sofia, S., Fiala, A. D., Herald, D. and Muller, P. M. 1980, *Science,* **210**, 1243.
Eddy, J. A., and Boornazian, A. A. 1979, *Bull. A.A.S.,* **11**, 437.
Frohlich, C. 1981, in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Gough, D. 1981, in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Hickey, J. 1981, in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Hill, H. 1981, in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Livingston, W. C. 1978, *Nature* **272**, 340.
Livingston, W.C. 1981, in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Parkinson, J. H., Morrison, L.V., and Stephenson, F. R. 1980, *Nature* **288**, 548.
Shapiro, I. I. 1980, *Science*, **208**, 51.
Sofia, S., O'Keefe, J., Lesh, J.R., and Endal, A.S. 1979, *Science,* **204**, 1306
Sofia, S., and Endal, A. S. 1980, in *Proceedings Conf. Ancient Sun*, eds. R. O. Pepin, J. A. Eddy and R. B. Merrill p. 139.
Sofia, S., and Chan, K. L. 1981, *Solar Physics*, submitted for publication.
Sweigart, A. 1981, in *Proceedings of the Workshop on Solar Constant Variation*, ed. S. Sofia, (Greenbelt: NASA), in press.
Twigg, L. and Endal, A.S. 1981 in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Willson, R. C. 1981 in *Proceedings of the Workshop on Solar Constant Variations*, ed. S. Sofia, (Greenbelt: NASA), in press.
Willson, R. C., Gulkis, S., Janssen, M., Hudson, H. S., and Chapman, G.A. 1981, *Science*, in press.

DISCUSSION

Ferrín: El diámetro lineal del Sol que Ud. determina, depende del diámetro angular de la Luna, y éste a su vez de la distancia Tierra-Luna. ¿La indeterminación en esta distancia cuánto afecta el valor del radio solar obtenido?
Sofia: La distancia Tierra-Luna desde luego afecta nuestro método de determinación del diámetro solar. Sin embargo esto no constituye un problema porque: 1) la distancia actual entre la Tierra y la Luna, determinada por la reflexión de la luz de laser en los retrorreflectores dejados en la Luna por uno de los vuelos Apolo, es conocida con una exactitud de decenas de cm. 2) Debido a las observaciones de la posición de la Luna llevadas a cabo durante varios siglos, el cambio relativo de la distancia lunar puede establecerse con una exactitud de pocos metros en, digamos, 200 años.
Mendoza, C.: ¿Qué es más importante: determinar las variaciones de la luminosidad solar o del radio?
Sofia: Ambos factores son importantes en su propio contexto. Por ejemplo, para quienes estudian clima, la luminosidad es más importante que el radio. Para los físicos solares, en cambio, ambos factores son esenciales, debido a que la comprensión del mecanismo físico de las variaciones solamente puede lograrse con base en ambos parámetros.
Pérez-Peraza: Existe una gran multitud de pulsaciones en el Sol con diferentes escalas temporales, ¿cuál es la semejanza o diferencia entre esas pulsaciones y las variaciones radiales de las que se está hablando ahora?
Sofia: La diferencia entre las pulsaciones y los cambios estructurales es una de amplitud. Los cambios estructurales tienen amplitudes varias órdenes de magnitud mayores que las pulsaciones.
Mendoza, E.: ¿Son independientes los resultados de la constante solar obtenidos fuera de la atmósfera?
Sofia: Sólo parcialmente. Por ejemplo, algunos efectos que introducen errores están en la vecindad inmediata del satélite (por ejemplo su posición con respecto a los cinturones de Van Allen). Los satélites desde los que se han determinado están en órbitas distintas, por lo que la coincidencia de cambios en dos satélites indican que no se debe a este tipo de error. Sin embargo, puede haber otros efectos (partículas de origen solar, por ejemplo) que afecten igualmente a ambos satélites, y por ende que las mediciones no sean independientes.

Sabatino Sofia: Laboratory for Planetary Atmospheres, Goddard Space Flight Center, NASA, Greenbelt, MD 20771, EUA.