AN AUTOCORRELATION METHOD FOR DETERMINING THE SUNSPOT ACTIVITY PERIOD

G. MARIŞ

Astronomical Observatory, Bukarest, Roumania

Abstract: An autocorrelation method was used to estimate the solar activity period assuming an autoregressive series of the data. An autoregression period of 10.99 years was found for the annual mean daily areas of sunspots and 10.98 years for the annual mean daily sunspot number, R.

Discussion and suggestions about the application of this method to the forecast of solar activity and the study of north-south sunspot asymmetry were also made.

Introduction

All the determinations of the sunspot period of activity carried out, according to various indices, have revealed values very appropriate to 11 years. With this end in view, several methods, have been used from simple determinations to the most sophisticated methods of mathematical statistics (harmonic analysis or autocorrelation calculus; Cole, 1973). A number of papers dealt either with the determination of the longer periods of the sunspot activity (Vasilev and Kandaurova, 1970; Vasilev, 1970) or the determinations of short fluctuations within the 11-year cycle (Kandaurova, 1971a; Ramanuja Rao, 1973).

The present paper attempts to establish the basic cycle duration of the sunspot activity, using the autocorrelation method, namely the study of correlograms.

Data and Method of Analysis

We have used the annual mean daily areas of the sunspots and the annual mean daily sunspot numbers, R, for the period 1874—1953, tabulated in the catalogue Sunspot and Geomagnetic Storm Data derived from Greenwich Observations, 1874—1954. In the second stage, we have used the sunspot number series, smoothed by subsťracting a running mean from it, over the period 1749—1960 (Vitinskii, 1963). For these series of data we have calculated the autocorrelation coefficients, defined as (Yule and Kendall, 1969):

$$r_k = \frac{n}{n-k} \frac{\sum_{t=1}^{n-k} u_t u_{t+k}}{\sum_{t=1}^{n} u_t^2},$$

where k is the lag ($k = m\Delta t$, m is the lag number and Δt is the time interval between successive observations) and n is the total number of data. The maximum lag number is chosen to be consistent with the requirement of stability and small variance of the estimates. In our case it is about 1/2 the total number of data. The correlograms, the diagram of these coefficients, the function of the k lag, for the series of sunspot areas is shown in Figure 1. The correlogram aspect of an attenuating harmonic curve represents a first indication that the initial series of data can be considered autoregressive, in a first approximation. For this type of series, a term can be represented in a three-term autoregressive pattern, as:

$$u_{t+2} = -\alpha u_{t+1} - \beta u_t + \varepsilon_t,$$

where α and β are constants and ε_t is the perturbation factor which interposes at the moment t, and whose influence is manifested on the following terms of the series. For this type of series, it is possible to establish the constants α and β, as well as an autoregressive period. This period is not, however, totally equal to the distance between the peaks, which could vary. These values are given by:

$$\alpha = \frac{-r_1(1-r_2)}{1-r_1^2};$$

$$\beta = -1 + \frac{1-r_2}{1-r_1^2};$$

$$P = \frac{2\pi}{\theta},$$

where $\cos\theta = \dfrac{-\alpha}{2\sqrt{\beta}}$.

As for the perturbation factor, this cannot be determined by its own nature, but its dispersion can be estimated in the whole series dispersion by:

$$D^2(u) = \frac{1+\beta}{(1-\beta)\,[(1+\beta)^2-\alpha^2]}\,D^2(\varepsilon).$$

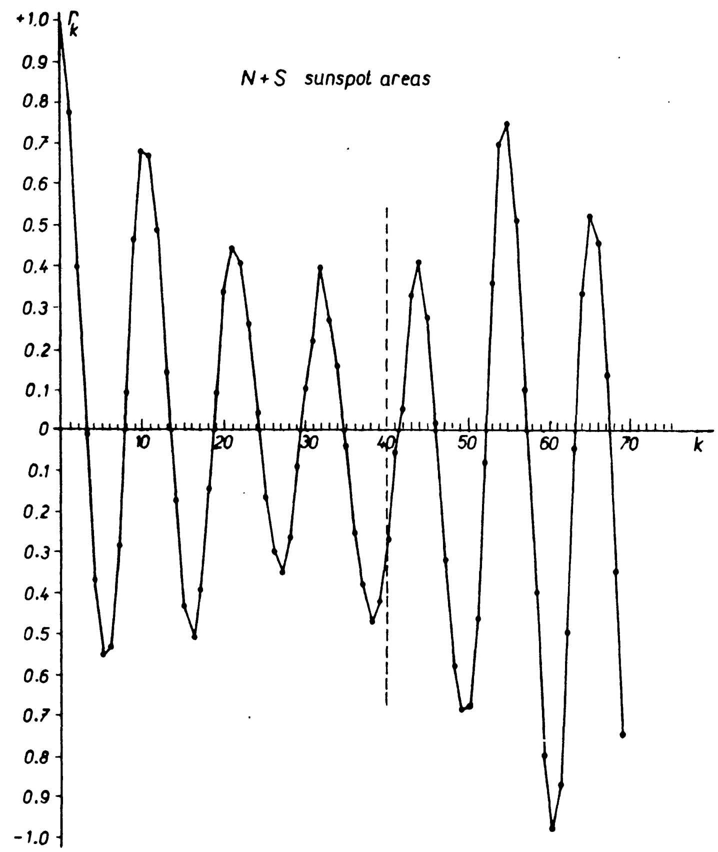

Fig. 1.

Table 1. The characteristic values of the solar index series

	α	β	P (years)	$D^2(\varepsilon)/D^2(u)$
Annual mean sunspot areas (1874—1953)	−1.1353	0.4555	10.99	0.30
R (1874—1953)	−1.2384	0.5435	10.95	0.25
R (smoothed) (1749—1960)	−1.4947	0.7904	10.98	0.11
$\dfrac{\mathrm{N-S}}{\mathrm{N+S}}$ sunspot areas (1874—1953)	−0.2949	0.0620	6.70	0.92

Results and Discussion

From the autocorrelation coefficient values, for each series in use, we have acquired the values presented in Table 1.The conditions that constants α and β should satisfy so that the series may oscillate within certain limits, complying with the autoregressive assumption, are deduced from the statistical theory: $0 \leqslant \beta \leqslant 1$ and $|\alpha| \leqslant 2$ (Yule and Kendall, 1969). These conditions were compiled with for all the analysed cases. The acquired autoregressive periods are generally close to each other and in conformity with the 11-year period of sunspot activity.

As regards the perturbation factor, an interesting fact is the situation in the two series of the sunspot numbers: using a longer data series, 211 values instead of 80, and, especially, applying a running mean over it, eliminates certain observational errors, and the perturbation factor is decreased to less than a half (0.11 instead of 0.25). This fact represents proof in that a good deal of the perturbation component is due to the observational errors, or to the processing. The difference between the sunspot number series and the sunspot area series for the same period (1874—1953) is not significant, in this regards. On the assumption of introducing certain observational errors and especially miscalculations, the sunspot areas, deduced by means of a longer calculation, can generate a larger error. The contribution of the perturbing component to the variation of the whole series of the solar activity indices, acquired by this method, is in good agreement with the value of 10—25%, acquired by other methods (Vasilev and Vitinskii, 1969; Vasilev, 1970; Kandaurova, 1971b).

The correlogram in fact offers initial information on whether a series of data may be considered periodical, or random. For example, the same method applied to the area N — S/N + S series, where N and S represent sunspot areas of the northern and southern hemispheres of the Sun, respectively, leads to the correlogram shown in Figure 2. The attempt to find an autoregressive period has led to the value of 6.70 years. The north-south asymmetry of solar activity, still unexplained, is in no case checked by such a period. A harmonic analysis is to be applied to this series for identifying the possible periods, because the phenomenon should be a superposition of several oscillations with different periods and amplitudes.

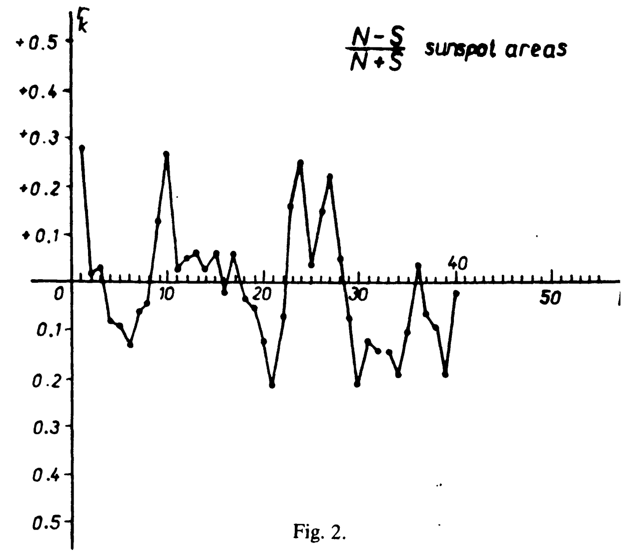

Fig. 2.

Table 2. The annual mean daily sunspot number predicted (using an autoregressive pattern) and observed

Year	R (predicted)	R (observed)
1959	135.7	152.0
1960	142.4	112.3
1970	73.8	104.5
1971	73.9	66.6
1972	93.2	68.9

The autoregressive pattern assumption, applied to the solar activity indices, allows for a short duration forecast, assuming that the perturbation factor is cumulative. An examination of the two last solar cycles has shown that for the second and third year after the maximum, the best forecast was obtained (Table 2). On the decreasing and increasing branches of the cycle, the acquired values differ more from the observed. In this pattern, it seems that the perturbation component reacts less in the maximum phase of cycle and much more in the other ones.

Although it does not provide good accuracy for a sunspot activity forecast, this method offers a good opportunity for the nalysis of the periodical series and, within a certain approximation, for finding an autoregressive period.

Acknowledgement

The author would like to express her sincere thanks to Prof. C. Popovici for his continuous interest and encouragement.

References

COLE, T. W. (1973): Solar Phys., *30*, 103.

KANDAUROVA, K. A. (1971a): Soln. Dann. Bull., *5*, 107.

KANDAUROVA, K. A. (1971b): Soln. Dann. Bull., *11*, 80.

RAMANUJA RAO, K. (1973): Solar Phys., *29*, 47.

VASILEV, O. B. and VITINSKII, Yu. I. (1969): Soln. Dann. Bull., *9*, 105.

VASILEV, O. B. (1970): Soln. Dann. Bull., *8*, 95.

VASILEV, O. B. and KANDAUROVA, K. A. (1970): Soln. Dann. Bull., *11*, 109.

VITINSKII, YU. I. (1963): Prognozy Solnechnoi Aktivnosti. Izd. Akad. Nauk USSR, Leningrad, p. 140.

YULE, G. U. and KENDALL, M. G. (1969): Introducere in Teoria Statisticii. Bucuresti.