Astron. Astrophys. 345, 1006–1010 (1999)

ASTRONOMY
AND
ASTROPHYSICS

The shape of the sunspot cycle described by sunspot areas

Kejun Li

Yunnan Observatory, The Chinese Academy of Sciences, P.R. China (ssg@public.km.yn.cn)

Received 21 October 1998 / Accepted 13 January 1999

Abstract. The temporal behavior of the sunspot cycles 12–22, as described by quarterly mean values of sunspot areas, can be represented by a single function containing only two parameters: the starting time and the amplitude of a cycle. The parameters are determined for the more recent 11 sunspot cycles and examined for any predictable behavior. As a sunspot cycle progresses, the amplitude parameter can be better determined to within 10% at 4.3 years into its cycle, and the starting time can be well determined to within ±0.3 year at 4.6 years into the cycle, thereby providing a good prediction both for the timing and size of sunspot maximum and for the behavior of the remaining 5–10 years of the cycle. Characteristics of the sunspot cycle shape are investigated in this paper. The shape of the sunspot area cycle is generally asymmetric, taking less time rising from minimum to maximum than reaching the next minimum from the maximum, and cycles with larger amplitude are more asymmetric and take less time to reach maximum, the same as the so-called Waldmeier effect of the cycles described by the sunspot relative numbers.

Key words: Sun: sunspots – Sun: activity

1. Introduction

Sunspots are among the solar active phenomena, which were earliest observed. Sporadic naked-eye observations exist in chinese dynastic histories since 165 BC (Wittmann & Xu, 1987). The sunspot number is dominated by an 11-year periodicity, which was first reported by Wolf from his examination of the Zurich observatory sunspot records. Earlier, Schwabe had announced an apparent systematic fluctuation with a shorter periodicity (see Meadows, 1970; Shove, 1983; Hathaway, et al. 1994). At present, it is known that sunspot cycles are not strictly periodic, and periods of inactivity like the Maunder minimum from 1645 to 1715 and the Spörer minimum in the 15th century (Maunder, 1922; Eddy, 1976; Ding Youji, 1978) are known to have occurred. For details, see the paper by Polygiannakis, et al.(1996), and references herein Several different approaches have been attempted for understanding and, if possible, predicting evolution of the sunspot cycle by studying the spectral, statistical, morphological and chaotic properties of the sunspot index, area or latitudes (Vigouroux & Delache, 1994; Leftus, 1994; Hong, 1990, 1994; Elling & Schwentek, 1992; Polygiannakis, et al., 1996; Feynman & Gabriel, 1990). In this paper, we describe our efforts to find a simpler function that might be used to reproduce the shape of the sunspot area cycles 12–22, simpler in terms of having fewer free parameters. We examine the relationships between the various parameters and assess the potential for an early determination of the relevant parameters for use in predicting future solar activity. The analysis reveals that, indeed, the temporal behavior of the sunspot area can be adequately described by a simple function of only two parameters for each cycle and these parameters can be determined fairly early in the cycle. Characteristics of the sunspot cycle shape are investigated in this paper too.

2. Basis function and cycle shape parameters

For our analysis, we use quarterly averages of the sunspot areas. They extend from the year 1875 to 1996 and consist of two parts having different reliability. The first part comes from the Greenwich Photoheliographic Results, extending from 1875 to 1971, but the graph given by White & Trotter (1976) was used as we have no primary digital data. Five points divided one year of the abscissa into four equidistance parts in the enlarged graph, then we regarded the average of coordinate values of adjacent two points as quarterly average of the corresponding time. The rest comes from sunspot observations of Yunnan Observatory (Hong & Wang, 1988). The Greenwich Photoheliographic Results are available until the year 1976, a comparison of the first part with the second was done for the overlapping period (1971 to 1976). Deviation between the two parts is less than 18%, and for most values of sunspot areas it is less than 12%. These quarterly averages are plotted in Fig. 1 for the entire interval (1875–1996).

Inspection of Fig. 1 reveals that individual cycles show a wide range of temporal behavior. For example, some cycles (like 12–16, and 20) are small in amplitude (about half the size of cycle 19), while others are considerably larger. Most cycles show substantial asymmetry with the rise to maximum being faster than the fall to minimum. Although the average cycle length (minimum to minimum) is about 11 years, individual cycles vary in length from about 9 to 14 years. All these properties

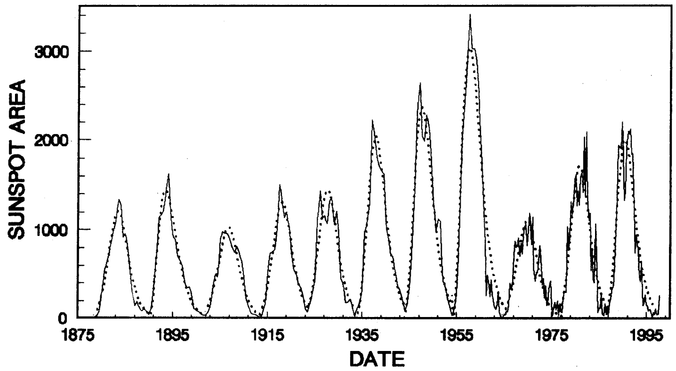

Fig. 1. Quarterly averages of the sunspot areas for 1875 to 1996. This illustrates the cyclic behavior of the sunspot areas and shows the variations in size and shape of each sunspot cycle. A comparison of the sunspot data and the two-parameter functional fit for the cycles 12–22 is given too (details are given late in the text). The thin line represents the data, the thick dotted line represents the two-parameter fit for each cycle.

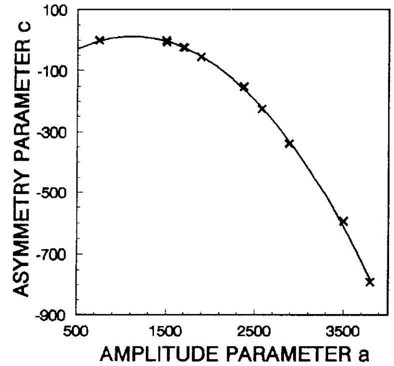

Fig. 2. Parameter c plotted as a function of parameter a (amplitude) for all the cycles 12–22. The functional fit given by Eq. (2) is represented by the thick line through the data points marked by symbol $\times$.

are similar to the characteristics of the solar cycles described by the international sunspot relative numbers. Initiated by the work of Hathaway, et al. (1994), a function of the form

$$f(t) = \frac{a(t-t_0)^3}{e^{(t-t_0)/b} - c} \quad (1)$$

is proposed to describe the shape of the sunspot area cycles, where parameter a represents the amplitude, b is related to the time from minimum to maximum, c gives the asymmetry of the cycle, t_0 denotes the starting time. Here we use a linear term of $(t-t_0)/b$ in the exponential function of the equation instead of the quadratic term used by Hathaway, et al. (1994). This function reproduces both the rise and decay portions of the sunspot cycle. We determined the best-fit parameters for each cycle using a nonlinear least-squares fitting algorithm in which all four parameters can vary. In attempting to fit each cycle with the four-parameter basic function we found that the process initially gave parameters with large uncertainties and the parameter b could be fixed at a constant value for the 11 complete cycles. Although slight variations of the values of the parameter b occur from cycle to cycle, they appear to be unrelated to the other parameters and the results are consistent with taking a simple value of $b = 1.128(years)$ for all cycles. This effectively reduces the number of parameters by one and also stabilizes the fitting procedure by removing one degree of freedom. Eq. (1) becomes a function of three parameters. Then we fix the parameter $b(= 1.128)$ to reproduce the shape of the sunspot area cycle again. Table 1 gives summary of best-fit parametric values for the three-parameter function. Examination of the best-fit parametric values for each cycle shows that parameters a and c are related. Fig. 2 shows parameter c as a function of parameter a for all 11 cycles. A strong relationship between a and c is clearly seen, and the best fit to this relationship is given by a two-order polynomial expression of the form

$$c = -1.104 \times 10^{-4} a^2 + 0.24666a - 123.593 \quad (2)$$

The correlation coefficient $r = -0.9372$. This effectively reduces again the number of parameters by one, leaving only two. Then the two-parameter fit is attempted to be done for all 11 cycles. Table 1 also lists the best-fit parametric values for the two-parameter function. Fig. 1 shows a comparison of the sunspot area data and the two-parameter functional fit for the years 1875 to 1996 (cycles 12–22). A measure of the goodness-of-fit is given in the fifth and eighth columns respectively for the three-parameter functional fit and the two-parameter functional fit, which is defined by $\sigma = \sqrt{\sum_{i=1}^{N}(A_i - f_i)^2/N}$, where A_i is the quarterly-averaged value of the sunspot area, f_i the functional fit value, and N the number of quarters in the cycle. The three-parameter function usually gives a slightly better fit than the two-parameter function, and the best fit of the above three kinds is given by the four-parameter function. These differences are not, however, considered significant. The more important result is that sunspot area data can be adequately fit with a relatively simple function of only two parameters.

3. Forcasting potential

The analysis in Sect. 2 implies that the temporal behavior of a sunspot cycle is governed by two parameters, amplitude and the starting time of the cycle. These parameters would be useful in forecasting future activity if they could be determined early in the cycle or found to vary predictably from one cycle to the next. Given a series of different values of the starting time t_0 (about 300 values around the observational minimum time of a cycle) and a series of different values of the amplitude parameter a (about 3000 values around the observational maximum amplitude), for each cycle we calculate the errors of the sunspot area data and the two-parameter function values. Each set of parametric values, consisting of one value of t_0 and one value of a, gives one error value, and about 900000 error values are obtained for each cycle. Then we chose the parametric values corresponding to the minimum error as the best-fit result. Fig. 3 shows determinations of the amplitude parameter a at 3-month time intervals into each cycle. The ratio of the estimated values a to its final value determined at the end of the cycle is plotted

Table 1. Summary of the cycle shape parametric values

cycle	Three-parameter fit				Two-parameter fit					
	$a \times 10^{-4}$	c	$t_0(years)$	σ	$a \times 10^{-4}$	$t_0(years)$	σ	A_{max}	$T(years)$	$(t_m - t_0)(years)$
12	0.2883	−338.65	1876.81	11.75	0.2941	1876.72	12.12	1190.955	13.20	6.6891
13	0.0744	0.08328	1890.03	11.78	0.0746	1889.91	12.01	1425.159	9.50	3.4346
14	0.3799	−791.14	1899.41	13.95	0.3776	1899.43	13.95	1025.134	12.20	7.3440
15	0.2574	−225.13	1911.73	14.36	0.2642	1911.61	14.64	1289.784	9.93	6.3562
16	0.2372	−150.59	1921.65	27.07	0.2340	1921.53	27.75	1420.279	11.54	5.9686
17	0.1688	−21.762	1933.14	18.27	0.1720	1933.07	18.46	2024.681	10.50	4.6660
18	0.1501	−5.8790	1943.49	26.42	0.1572	1943.57	26.89	2461.595	10.70	3.9924
19	0.1495	0.93512	1954.31	29.31	0.1500	1954.27	29.47	2973.208	8.30	3.2899
20	0.3499	−594.27	1962.68	21.74	0.3432	1962.57	21.98	1084.141	12.87	7.0954
21	0.1887	−54.266	1975.56	29.84	0.1960	1975.44	29.59	1690.669	10.13	5.3072
22	0.1699	−23.598	1985.68	28.60	0.1744	1985.56	28.64	1978.869	10.85	4.7457

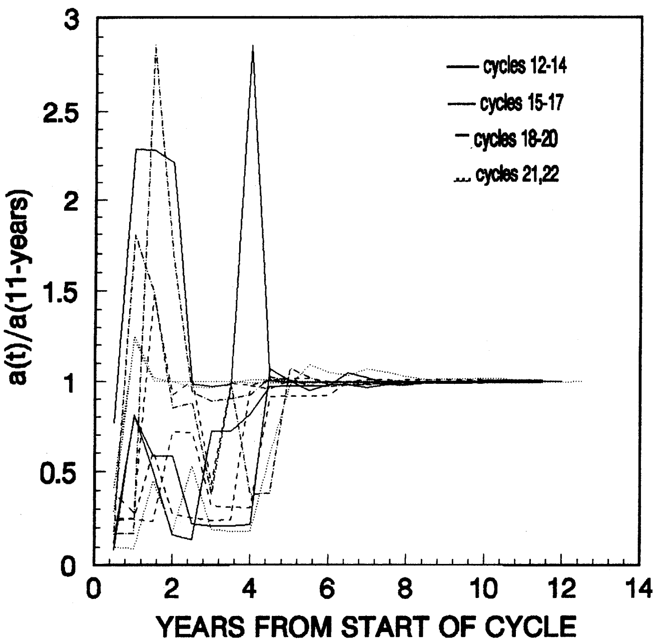

Fig. 3. Determinations of the amplitude parameter a at 3-month intervals into each cycle. The ratio of the estimated values of a to its final value determined at the end of the cycle is plotted for each cycle.

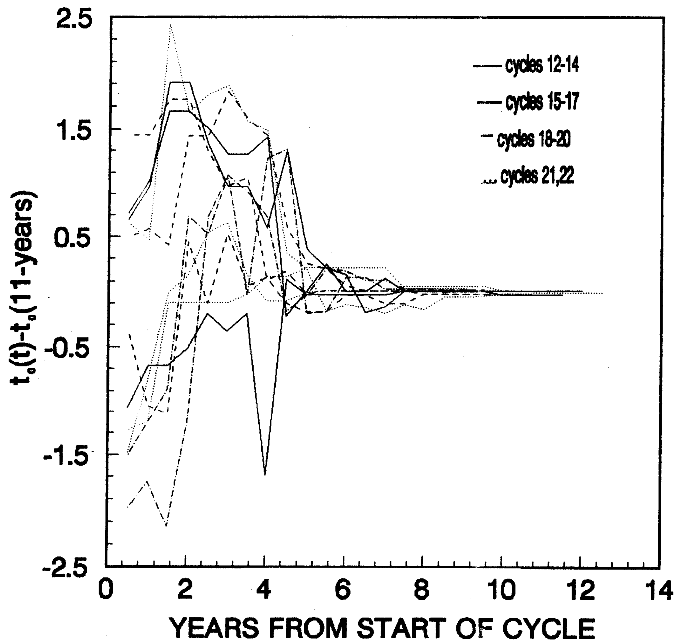

Fig. 4. Determinations of the starting time parameter t_0 at 3-month intervals into each cycle. The difference of the estimated values of t_0 and its final value determined at the end of the cycle is plotted for each cycle.

in the figure for each cycle. The value for parameter a can, in fact, be quite accurately determined with the first 4–5 years following the start of the cycle, (as the sunspot cycle progresses, the amplitude can be determined to within 10% at 4.3 years into its cycle). Fig. 4 shows determinations of the starting time parameter t_0 at 3-month intervals into each cycle. The difference between the estimated values t_0 and its final value determined at the end of the cycle is plotted in the figure for each cycle. The value for the starting time t_0 can be well determined with the first 5–6 years following the start of the cycle, about one year longer than for the amplitude parameter, and can be determined to within ±0.3 year at the first 4.5 years into the cycle. This suggests that fitting the behavior of the solar cycle by means of the two-parameter function during the rising phase gives a prediction of the behavior of solar activity over the remaining 5 to 10 years of the cycle. What needs emphasis is that the two parameters a and t_0 are determined at the same times.

4. Features of the sunspot cycle

The time of the sunspot area maximum, t_m, was given explicitly by the value of $t = t_m$ for which $\frac{df(t)}{dt} = 0$, then we have

$$3e^x - 3c - xe^x = 0 \tag{3}$$

$$t_m = t_0 + 1.128x \tag{4}$$

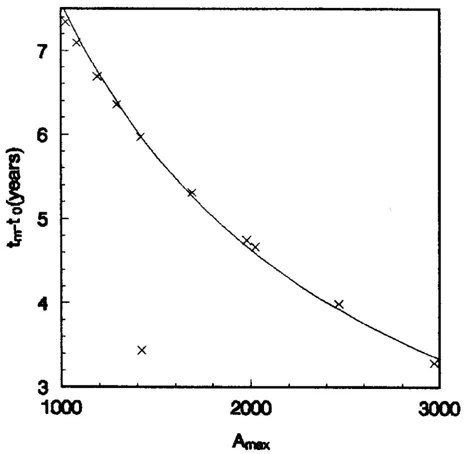

Fig. 5. Relationship of the rising time $(t_m - t_0)$ and the maximum amplitude A_{max}. The thin curve is the regression line.

$(t_m - t_0)$ is the rising time of a cycle. The value of the sunspot area at maximum, A_{max} is given by

$$A_{max} = \frac{1.435ax^3}{e^x - c} \tag{5}$$

Using Eqs. (3) to (5), we calculate values of A_{max} and t_m of each cycle, which are listed in Table 1 as well. Here, we define the cycle length T of a solar cycle as the difference between the starting points of two successive cycles, and the value of T of each cycle is listed in the table too. The starting time of the 23rd solar cycle, now known to be 1998 May, is regarded as the parametric value of the starting time t_0 of the cycle.

The relationship of the rising time $(t_m - t_0)$ and the maximum amplitude A_{max} is shown in Fig. 5. The thin line in the figure is the regression line of the form

$$t_m - t_0 = \frac{1}{8.36515 \times 10^{-5}x + 0.04846} \tag{6}$$

The figure shows that the shorter the rising time of a cycle is, the larger the maximum amplitude of the cycle is. In other words, cycles with large amplitude take less time to rise to maximum, similar to the so-called Waldmeier effect of the sunspot number. A weak correlation between cycle length and the rising time of solar cycles is found and shown in Fig. 6. The linear regression coefficient $r = 0.8081$, and the line is

$$T = 6.462 + 0.826(t_m - t_0) \tag{7}$$

The figure indicates that cycles that take a long time to get maximum amplitude tend to run a long time to get ended. The mean value of the rising time is 5.253 years for the 11 cycles, the average of the cycle length is 10.883 years. So on an average, the rising time is less than the falling time of a cycle. We define asymmetry of a cycle as $asymmetry = \frac{T-(t_m-t_0)}{t_m-t_0}$. Relation between $asymmetry$ and A_{max} is shown in Fig. 7. There appears to be a upward trend with $A_{max} = 528.327 + 1031.99 \times asymmetry$. It shows that cycles with larger amplitude are more asymmetric.

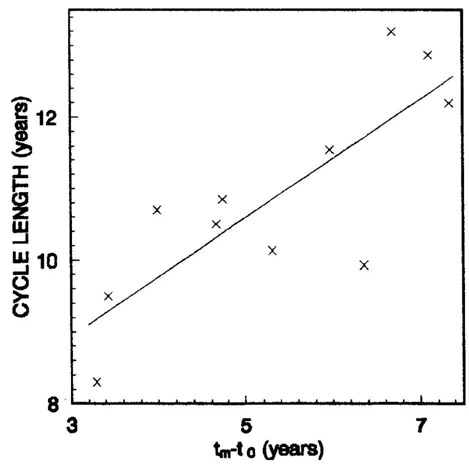

Fig. 6. Correction between cycle length and the rising time of the solar cycles 12–22. The linear regression line is shown by the thin line.

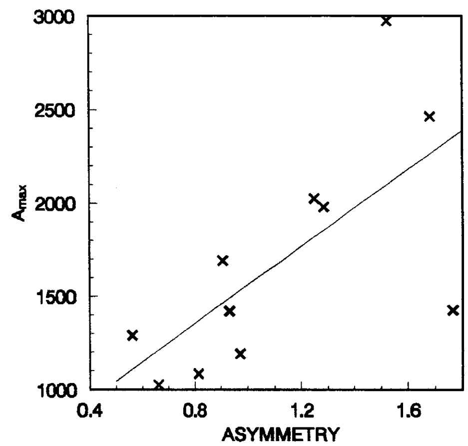

Fig. 7. Relation between the maximum amplitude A_{max} and $asymmetry$ defined in the text. The thin line, i.e. the linear regression line shows a upward trend with a linear equation given in the text.

5. Conclusions

Data of the quarterly mean values of the sunspot area from the year 1875 to 1996 (cycles 12–22) are used to discuss the shape of the solar cycles in this paper. First, we describe our efforts to find a simpler function that might be used to reproduce the shape of the sunspot cycles, simpler in terms of having fewer free parameter. Our study clearly shows that the shape of the sunspot cycle can be adequately described using a simple function with only two parameters, the starting time and the amplitude of a cycle. This function is derived from a 4-parameter functional fit of the quarterly mean sunspot areas. Then we find that the values of the amplitude and the starting time are well determined early in the cycle. As a sunspot cycle progresses, the amplitude parameter can be better determined to within 10% at 4.3 years into its cycle, and the starting time can be well determined to within ±0.3 year at 4.6 years into its cycle. This determination of the parameters provides an early estimate for the temporal behavior of the sunspot cycle over its remaining years including sunspot maximum. Characteristics of the sunspot cycle shape are investigated in this paper too. The shape of the sunspot area cycle is generally asymmetric, taking less time rising to maximum from minimum than reaching the next minimum from the maximum. Probably the most significant relation among the cycle shape characteristics is the Waldmeier effect: cycles with larger amplitude are more asymmetric and take less time to reach maximum. This effect, however, is modified for the sunspot area cycles, comparing with the cycles of the sunspot numbers. It is well known that the reliability of sunspot numbers is larger than that of sunspot areas because the former is easier to measure. In a recent paper by Pettauer & Brandt (1994) errors of sunspot areas determined even with sophisticated methods are given between 8.5% and 10%. The difference of the Waldmeier effects for the sunspot area cycles and for the sunspot relative number cycles is perhaps caused by the bigger uncertainty of the sunspot area

Acknowledgements. The author gratefully acknowledges the useful help in the map-making work from Dr. Zhang Xiaoyu of Yunnan Observatory. A part of the data used in this paper is provided by the staff of the sunspot observation group of the Yunnan Observatory, the author is deeply indebted to their support. The author is grateful to Dr. Wöhl for his useful comments on this paper.

References

Ding Youji, 1978, Bulletin of Chinese Sciences 2, 35
Eddy J.A., 1976, Sci 192, 1189
Elling W., Schwentek H., 1992, Solar Phys. 137, 155
Feynman J., Gabriel S.B., 1990, Solar Phys. 127, 393
Hathaway D.H., Wilson R.M., Reichman E.J., 1994, Solar Phys. 151, 177
Hong Qinfang, Wang Hongzao, 1988, Sunspot relative numbers and sunspot areas from 1961 to 1986. Library of Yunnan Observatory
Hong Qinfang, 1990, Publication of the Yunnan Observatory 36, 9
Hong Qinfang, 1994, Publication of the Yunnan Observatory 55, 6
Leftus V., 1994, Solar Phys. 149, 405
Maunder E.W., 1922, Brit. Astron. Assoc. J. 32, 140
Meadows A.J., 1970, Early Solar Physics. Pergamon Press, New York, 95
Pettauer T., Brandt P.N., 1997, Solar Phys. 175, 197
Polygiannakis J.M., Moussas X., Sonett C.P., 1996, Solar Phys. 163, 193
Shove D.J., 1983, Sunspot Cycles. Hutchinson Ross Publ., Stroudsberg, Pennsylvania, 81
Vigouroux A., Pelache P., 1994, Solar Phys. 152, 267
Wittmann A.D., Xu Z.T., 1987, A&AS 70, 83
White O.R., Trotter D.E., 1976, ApJS 33, 391