

NONLINEAR PREDICTION OF SOLAR CYCLE 24

A. KILCIK¹, C. N. K. ANDERSON², J. P. ROZELOT³, H. YE², G. SUGIHARA², AND A. OZGUC⁴

¹ Department of Physics, Akdeniz University, Faculty of Arts and Sciences, 07058 Antalya, Turkey; alikilcik@akdeniz.edu.tr

² Scripps Institution of Oceanography, University of California, San Diego 92103, USA

³ Université de Nice Sophia-Antipolis CNRS-OCA, Fizeau Department, Av. Copernic, 06130 Grasse, France

⁴ Kandilli Observatory and E. R. I., Bogazici University, Cengelkoy, Istanbul, Turkey

Received 2008 September 1; accepted 2008 December 2; published 2009 March 5

ABSTRACT

Sunspot activity is highly variable and challenging to forecast. Yet forecasts are important, since peak activity has profound effects on major geophysical phenomena including space weather (satellite drag, telecommunications outages) and has even been correlated speculatively with changes in global weather patterns. This paper investigates trends in sunspot activity, using new techniques for decadal-scale prediction of the present solar cycle (cycle 24). First, Hurst exponent H analysis is used to investigate the autocorrelation structure of the putative dynamics; then the Sugihara–May algorithm is used to predict the ascension time and the maximum intensity of the current sunspot cycle. Here we report $H = 0.86$ for the complete sunspot number data set (1700–2007) and $H = 0.88$ for the reliable sunspot data set (1848–2007). Using the Sugihara–May algorithm analysis, we forecast that cycle 24 will reach its maximum in 2012 December at approximately 87 sunspot units.

Key words: Sun: activity – sunspots – methods: data analysis – methods: statistical

1. INTRODUCTION

Solar radiation is far from constant. These changes can be seen in many solar activity indicators, such as sunspot number, sunspot area, total solar irradiance, solar flares, the recurrence index of geomagnetic disturbances (Kane 2008), dynamo models (Hiremath 2001), and solar cycle length (Kane 2008). However, sunspot number is the most common solar activity indicator, having been recorded since 1700, and used as an indicator for solar cycle prediction since 1913 (Kimura 1913; Currie 1973; DeMeyer 1981; Kane 2007). Any change in solar activity presents challenges in solar physics (understanding solar cycle mechanisms, prediction of solar events such as flares, CME, etc.) or in the field of space weather (satellite drag, telecommunication outages, etc.). Therefore solar cycle prediction is of vital importance now, and will be even more important in the future. Despite general knowledge of solar cycles, reliable forecasting of sunspot numbers remains problematic. Several statistical methods have already been applied, but due to the nonlinear nature of the time series, linear approaches are expected to fail. Other techniques have been employed (e.g., precursors; Kane 2007) that offer theoretical advantages. Because of the wide variety of techniques used, the present solar cycle (Cycle 24) has been predicted to have a maximum sunspot number as low as 50 or as high as 180 (Obridko & Shelting 2008). In this study, we predict the length and intensity of cycle 24 using two statistical/dynamical methods: the Hurst exponent and the Sugihara–May algorithm.

2. DATA AND METHODS

The monthly ISSN (International SunSpot Number) data were used from 1847 to 2007. Though data is available back to 1700, according to Kane (2008) “the quality of the data is considered as poor during 1700–1748, questionable during 1749–1817, good during 1818–1847, and reliable since 1848.” The maximum of intensity of each cycle and the time from the beginning of the cycle to its maximum (hereafter “ascension time”) was derived for cycles 1 (1755 February) to 23 (2000 March). All the data sets used in this study were taken from the National

Geophysical Data Center (NGDC).⁵ Because standard methods are expected to fail to make accurate predictions, we apply two unusual techniques to this data to make our predictions. One is the Hurst Exponent, or Rescaled Range Analysis. This method was proposed by Hurst (1951) for an experimental study of long-term information storage in time series data. This technique has had wide application in many research fields, including finance (Grech & Mazur 2004; Qian & Rasheed 2004), astronomy (Komm 1995; Rozelot 1995, 2008; Ruzmaikin et al. 1994), climate (Rangarajan & Sant 2004), and others. Hurst analysis is a simple and robust way to analyze randomness in a data set. The parameter H measures the persistence of structures in the time series, indicating whether the data represent a pure random walk or have underlying trends. Another way to state this is that a random process with an underlying trend has some degree of autocorrelation. When the autocorrelation has a very long (or mathematically infinite) decay, the process is referred to as a long memory process. The value of H varies from 0 (indicating antipersistent brown noise) to 0.5 (random white noise), 1.0 (indicating a strong, smooth trend). Hurst found that the rescaled range series (R/S) over a time window of width t is described as a power law:

$$(R/S)_t = c^* t^H, \quad (1)$$

where c^* is a constant and H is the Hurst exponent. To estimate the value of the Hurst exponent, R/S is plotted versus t on log–log axes. The slope of the linear regression gives the value of the Hurst exponent. More details about R/S analysis, see Qian & Rasheed (2004). The second technique is the Sugihara–May Algorithm. This algorithm compares a library of (known) past patterns to patterns seen later in the real time series. It does this by reconstructing an attractor from the library, locating the “present” point on the attractor, and tracing that point forward along the attractor’s trajectory. The dimension of the attractor is determined by the complexity of the process that generates it, and the number of points involved in tracing

⁵ <http://www.ngdc.noaa.gov/stp/SOLAR/ftpsunspotnumber.html>

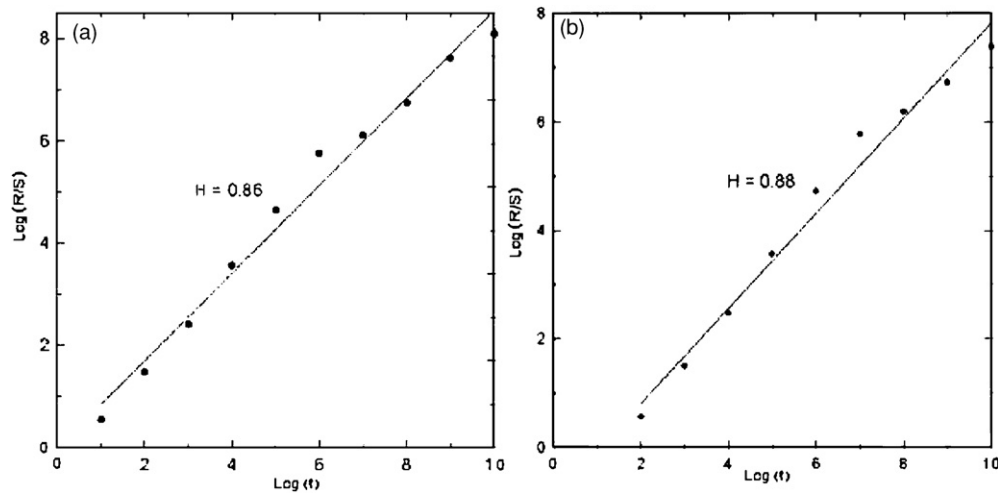


Figure 1. Hurst exponent analysis results deduced from monthly ISSN data sets; (a) from 1700 to 2007, (b) from 1848 to 2007. The slight departure from linearity is a signature of the cyclicity.

forward determines the nonlinearity of the system. A similar chaos technique was firstly applied to the historical sunspot data by Kurths & Ruzmaikin (1990) to determine the nonlinearity of the data set and to predict the following solar cycle. For a chaotic time series, the accuracy of nonlinear forecast falls off as prediction time increases (Sugihara & May 1990). This is a two-step procedure. First, simplex projection identifies the best embedding dimension, which is then used in an S-map procedure to check the nonlinearity (Sugihara & May 1990; Sugihara 1994). A time series X of length N , is embedded in D -dimensional Euclidean space to create a “landscape” x of $N - D - 1$ vectors, where $x_t: (X_{t-D+1}, X_{t-D+2}, \dots, X_t)$. The first n of these x_t vectors are associated with output values X_{t+T_p} . Then forecasts of the $(N - n)$ remaining input vectors, i.e. predictions, are made by

$$\hat{y}_{t+T_p} = \frac{\sum_{k=1}^{D+1} X_{k+T_p} \exp(-d_k)}{\sum_{k=1}^{D+1} \exp(-d_k)} \quad (2)$$

$$d_k = \|y_t - x_k\|, \quad (3)$$

where the summation is taken over the $D + 1$ closest neighbors in D -dimensional Euclidean space. Ideally, these closest neighbors form the vertices of the smallest simplex containing the predictee. In this context, this predictive approach is a local approximation similar to kernel density estimation. To find an appropriate value of D , various values of D are applied to find the value that minimizes prediction error (i.e., an embedding that minimizes the singularities or indeterminate crossing points of trajectories in the putative attractor). The optimal D is then used in an S-map to build weighted (local) linear predictions. The rate of decay of weight given to each point is set by a θ parameter, and it describes the degree of chaoticity of a data set. The forecast improvement with local weighting indicates that the underlying dynamics are nonlinear. Hence $\theta = 0$ performs best for linear time series (i.e., $\theta = 0$ is the global linear solution), and when $\theta > 0$ performs the best, the time series is nonlinear (Miyano et al. 2000). We investigate the significance of the historical relationship between the maximum sunspot number and the ascension time with Pearson’s test.

3. ANALYSIS AND RESULTS

Hurst exponent analysis was applied to the monthly ISSN data separated in two time intervals: 1700–2007 and

1848–2007. $H = 0.86 \pm 0.008$, and $H = 0.88 \pm 0.009$ for these data sets respectively (Figure 1). The bulge between six and eight in panels (a) and (b) of Figure 1 is due to the main contribution of the 11-year cycle (Ruzmaikin et al. 1994). To investigate the complexity, nonlinearity and predictability of the data, the Sugihara–May algorithm was applied as described above. From this analysis, it was discovered that the data are high-dimensional, i.e. governed by many orthogonal processes such as solar rotation and inner magnetic activity (Figure 2). In the second step of this procedure, the S-Map analysis (Figure 3), a weak nonlinear signature was detected in eight dimensions, indicating weak chaotic behavior. We also tested the prediction capability of our method by predicting the second half of the given data set (1927–2008) using only data from 1848–1926. The model produced very skillful predictions, showing a correlation of 0.94 with the observed values (Figure 4). We therefore attempted to perform long-range forecasts by sequentially increasing T_p (the time length into the future that one tries to predict from the last data point). These results are mostly an echo of the previous cycle. This technique predicts that the current solar cycle will reach a maximum on 2012 December, peaking at a sunspot number of 87.4 (Figure 5). A similar method to our prediction analysis was used by Kurths & Ruzmaikin (1990) for the solar cycle 22. They found that the maximum sunspot number would be about 150. Because the observed maximum of that cycle was 158, a rather good prediction estimate, it instills more confidence in the technique. For historical cycles, we further investigated the link between the maximum sunspot number and the ascension time. We found that the two variables have a strong negative correlation ($r = 0.82$, $df = 21$, and $p < 0.001$; Figure 6). Our prediction above (87.4 sunspots in 5.1 years from 2007 December) is plausible given the historical relationship between the two values, as cycles with an ascending phase between 4.5 and 5.5 years have a maximum in the 64.2 to 131.6 sunspot range.

4. DISCUSSION

Our results show that the Hurst exponents of the ISSN data for the periods of time 1700–2007 and 1848–2007 are 0.86 ± 0.008 and 0.88 ± 0.009 , respectively. This is in agreement with Mandelbrot & Wallis (1969) who first applied this type of analysis to the monthly sunspot number, and found $H = 0.86$.

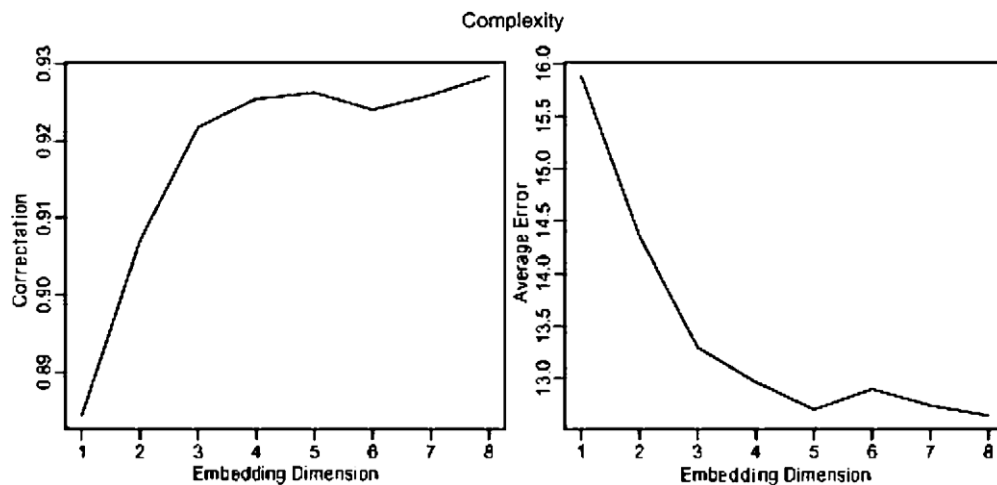


Figure 2. Complexity of the ISSN data. These figures show that the best embedding dimension is moderate ($D = 5-8$).

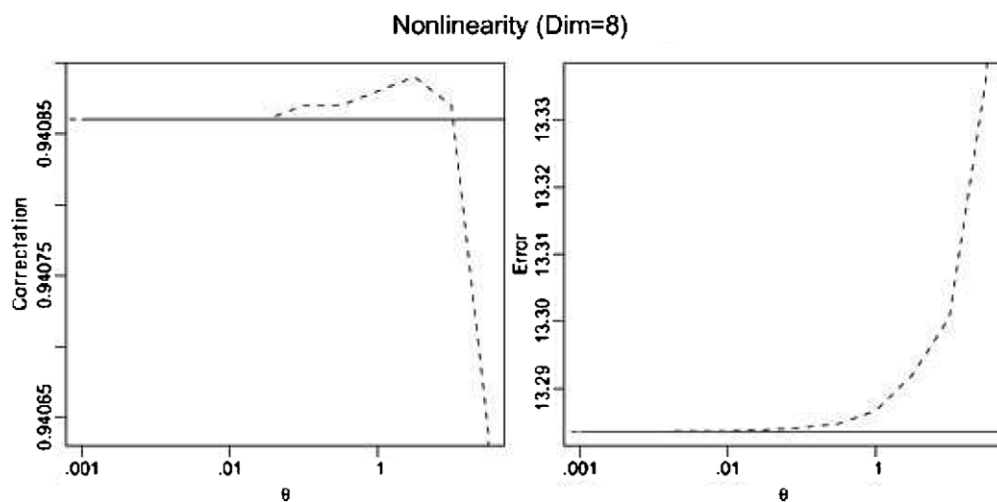


Figure 3. For the best embedding dimension of ($D = 8$) as deduced from Figure 2, the data set is most predictable when using a nonlinearity tuning $\theta > 0$, of about 0.18. For linear data sets this value is equal to zero. Qualitatively similar results are obtained with $D = 5$.

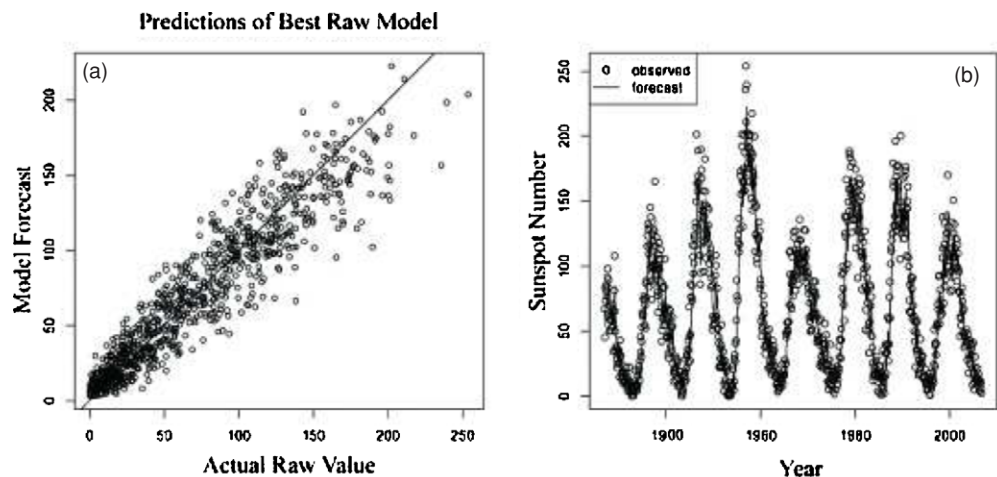


Figure 4. (a) Correlation/error between observed and real data. (b) Comparison of observed and predicted ISSN values for the used data period (1848–2007).

Similarly Ruzmaikin et al. (1994) analyzed the ^{14}C radiocarbon data as a proxy of solar activity, and found $H = 0.84$. In analyzing daily Doppler solar differential rotation coefficients A and B measured at Mount Wilson, USA, Komm (1995) found H to be 0.83 and 0.86, respectively. The slightly higher exponent

obtained for the recent dataset indicates that more reliable sunspot data exhibits a slightly stronger autocorrelation (or tendency to trend), and thereby has more statistical predictability than the full data set (1700–2007). Thus, it is likely that errors or uncertainty within the unreliable data affected its inherent

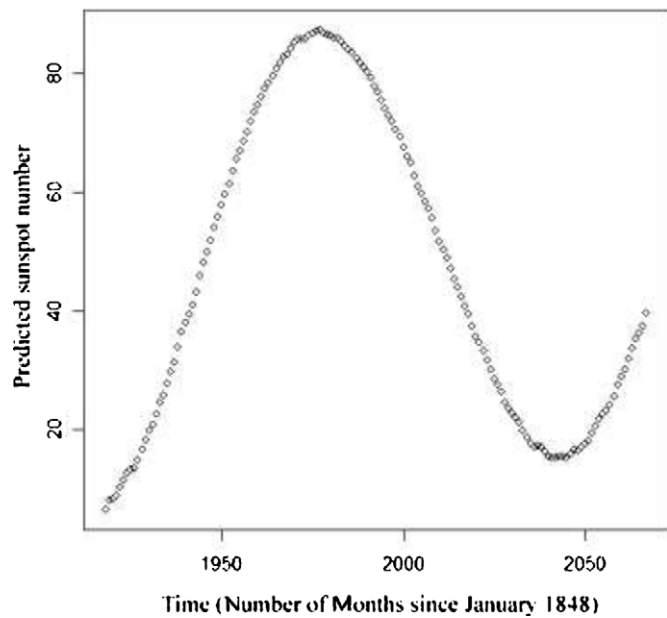


Figure 5. Monthly ISSN data predictions by using Sugihara and May Algorithms. The first value is taken as 1918 which describes also 2008 January.

predictability. Our prediction of the ascension time compares with values already available in the literature, which range from 2009 December to 2014 December (Maris & Oncica 2006; Tsurulnik et al. 1997). Obridko & Shelting (2008) reported that the second half of 2010 or the first quarter of 2011 would be the most reliable estimates for the maximum of cycle 24. Kane (2008) suggested 2011 October or 2012 August, while Maris & Oncica (2006) found that cycle maximum will occur as early as 2009 December. Our prediction of Cycle 24’s intensity also bears comparison with other forecasts. As early as 1983, Chistyakov (1983) claimed that cycle 23 and 24 would be low and cycle 24 would be lower than cycle 23. Duhau (2003) predicted a maximum of 87.5 ± 23.5 , and Wang et al. (2002) estimated a peak between 83.2 and 119.4. Maris et al. (2004) reported that cycle 24 has to be low by analyzing the number of flares. Javaraiah (2007) obtained a value of 74 from the sunspot group data and Hiremath (2007) obtained a value of 116 by using a harmonic oscillator solar cycle model. Kane (2007) predicted that the sunspot number would be 129.7 ± 16.3 for the present Cycle, but later (2008) revised this estimate to either 140 ± 20 in 2011 October or 90 ± 10 in 2012 August. It must be emphasized that the observational result for the start of cycle 24 (2008 April),

is in rather good agreement with Kane’s prediction. Thus, we can think that Kane’s estimates for the time at which cycle 24 will reach its maximum could be reliable. Using neuronal prediction, Maris & Oncica (2006) predicted 145 sunspots at the peak in 2009 December. The precursor models, which use data from the declining phase of cycle $N - 1$ to predict height and timing of cycle N , were used by Obridko & Shelting (2008). They reported that “the precursor models based on the polar field or H_α data often yield lower values of the current cycle,” which vary from 70 ± 10 and 120 ± 40 . From the dynamo model Choudhuri et al. (2007) have reported that cycle 24 would be 35% lower than cycle 23. (Dikpati & Gilman 2006) disagreed—using a flux transport dynamo model they argued that cycle 24 will be 30%–50% higher than cycle 23. Hathaway & Wilson (2006) gave a value of 160 based on geomagnetic activity at the minimum. From the index of the global magnetic field, Obridko & Shelting (2008) forecast that cycle 24 would be of “medium high, the same or somewhat higher than cycle 23.” Finally, Kitiashvili & Kosovichev (2008) using a nonlinear dynamo model as described by Kleorin & Ruzmaikin (2008) which takes into account dynamics of the turbulent magnetic helicity, predict that the next sunspot cycle will be significantly weaker (by $\sim 30\%$) than the previous cycle, continuing the trend of low solar activity.

By means of the two methods utilized here, the Rescaled Range Analysis and the Sugihara–May algorithm, we were able to deduce from the significant trends of the ISSN data both the cycle duration and its maximum intensity. Unlike previous analysis our forecasting results are based on models that were tested out of sample to have a high degree of forecast skill. Our conclusions therefore are (1) the reliable monthly mean sunspot data during 1848–2007 (Kane 2008) yield a slightly higher Hurst exponent than do all the historical observational data; (2) H , being greater than 0.5, shows that the sunspots series are highly persistent (exhibit momentum or trending); (3) concerning cycle 24, the maximum intensity of 87.4 will be reached in 2012 December; (4) according to this forecast, the current solar cycle will have a magnitude far lower than any other since 1890–1910.

All data sets used in this study are taken from NGDC web page. One of the authors (A.K.) is very thankful to the SWYA School organizing committee for providing financial support and to the lecturers for their valuable comments. This work, which is a small part of the PhD thesis of A.K., was supported by the Scientific and Technical Council of Turkey by the project of 107T878.

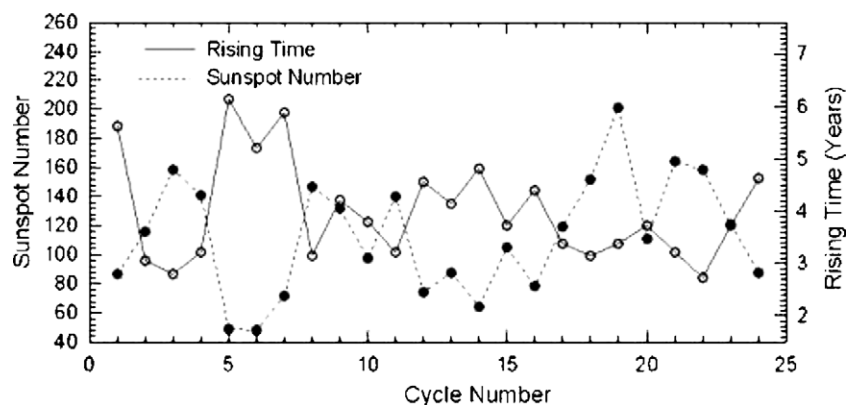


Figure 6. Comparisons of sunspot numbers and ascension time of historical cycles ($r = 0.82$, $df = 21$, and $p < 0.001$). The two last values describe our prediction.

REFERENCES

- Chistyakov, V. F. 1983, *Soln. Dannye.*, **1**, 97
- Choudhuri, A. R., Chatterjee, P., & Jiang, J. 2007, *Phys. Rev. Lett.*, **98**, 131103
- Currie, R. G. 1973, *Ap&SS*, **20**, 509
- DeMeyer, F. 1981, *Sol. Phys.*, **70**, 259
- Dikpati, M., & Gilman, P. A. 2006, *ApJ*, **649**, 498
- Duhau, S. 2003, *Sol. Phys.*, **213**, 203
- Grech, D., & Mazur, Z. 2004, *Physica A*, **336**, 133
- Hathaway, D. H., & Wilson, R. M. 2006, *Geophys. Res. Lett.*, **33**, L18101
- Hiremath, K. M. 2001, *Bull. Astron. Soc. India*, **29**, 169
- Hiremath, K. M. 2007, *Ap&SS*, **314**, 45
- Hurst, H. E. 1951, *Trans. Amer. Soc. Civ. Eng.*, **116**, 770
- Javaraiah, J. 2007, *MNRAS*, **377**, L34
- Kane, R. P. 2007a, *Sol. Phys.*, **246**, 471
- Kane, R. P. 2007b, *Sol. Phys.*, **246**, 487
- Kane, R. P. 2008, *Sol. Phys.*, **248**, 203
- Kimura, H. 1913, *MNRAS*, **73**, 543
- Kitiashvili, I., & Kosovichev, A. G. 2008, *ApJ*, **688**, L49
- Komm, R. W. 1995, *Sol. Phys.*, **156**, 17
- Kleeorin, N. I., & Ruzmaikin, A. A. 2008, *Magneto-hydrodynamics*, **1982**, 18, 116–122.
- Kurths, J., & Ruzmaikin, A. A. 1990, *Sol. Phys.*, **126**, 407
- Mandelbrot, B., & Wallis, J. 1969, *Water Resour. Res.*, **5**, 321
- Maris, G., & Oncica, A. 2006, *Sun Geosph.*, **1**, 8
- Maris, G., Popescu, M. D., & Besliu, D. 2004, in *IAU Symp. 223, Multi-wavelength Investigations of Solar Activity*, ed. A. V. Stepanov, E. E. Benevolenskaya, & A. G. Kosovichev (Cambridge: Cambridge Univ. Press), 127
- Miyano, T., Kimoto, S., Shibuta, H., Nakashima, K., Ikenaga, Y., & Aihara, K. 2000, *Physica D*, **135**, 305
- Obridko, V. N., & Shelting, B. D. 2008, *Sol. Phys.*, **248**, 191
- Qian, B., & Rasheed, K. 2004, in *IASTED Conf. Ser., The IASTED conference on Financial Engineering and Applications (San Francisco, CA: IASTED)*, 203
- Rangarajan, G., & Sant, D. A. 2004, *Chaos, Solitons and Fractals*, **19**, 285
- Rozelot, J. P. 1995, *A&A*, **297**, L45
- Rozelot, J. P. 2008, in *Lecture Notes in Physics 729*, ed. D. Benest, C. Froeschle, & E. Lega (New York: Springer) 391–405
- Ruzmaikin, A., Feynman, J., & Robinson, P. 1994, *Sol. Phys.*, **149**, 395
- Sugihara, G. 1994, *Philos. T. Roy. Soc. A*, **348**, 477
- Sugihara, G., & May, R. M. 1990, *Nature*, **344**, 734
- Wang, J. L., Gong, J. C., Liu, S. Q., Le, G. M., & Sun, J. L. 2002, *CJAA*, **2**, 557
- Tsirulnik, L. B., Kuznetsova, T. B., & Oraevsky, V. N. 1997, *Adv. Space Res.*, **20**, 2369