A&A 379, 611–615 (2001)
DOI: 10.1051/0004-6361:20011277

Astronomy & Astrophysics

Research Note

Nonlinear time series analysis of northern and southern solar hemisphere daily sunspot numbers in search of short-term chaotic behavior

N. Jevtić[1], J. S. Schweitzer[1], and C. J. Cellucci[2]

[1] Physics Department, University of Connecticut, 2152 Hillside Rd U-3046, Storrs, CT 06269-3046, USA
e-mail: nj@phys.uconn.edu
[2] Physics Department, Ursinus College, PO Box 1000, Collegeville, PA 19426-1000, USA
e-mail: ccellucci@ursinus.edu

Received 19 April 2001 / Accepted 11 September 2001

Abstract. Daily sunspot number data for the northern and southern solar hemispheres from 22 cycle maximum to the current cycle 23 maximum are analyzed using nonlinear time series methods in an effort to identify chaotic behavior on short time scales as indicated by the recent helioseismic observations. We find that separate analyses of north and south solar hemisphere data provide more consistent results. Different behavior at maxima and minimum are observed. The maximum sections have a region that bears the marks of self-similarity whereas the minimum shows no such behavior. However, it is the minimum data that determine the results for the correlation dimension over one cycle. Thus, over many cycles, minima mask the potentially self-similar sections at maximum; access to information on whether the solar cycle is chaotic long-term is severely limited, if not precluded. To make sure the results were not affected by noise, the method of artifactual correlations was used to investigate the noise content of the data, indicating that both the north and south data sets were relatively free of additive noise.

Key words. methods: data analysis – Sun: activity, sunspots

1. Introduction

The recent detection of velocity shear at the solar tachocline with a period of about 16 months and the possibility of a small scale surface dynamo (Howe 2000) justify a renewed look at solar activity indicators on shorter time scales.

Sunspots remain to this day the best-known manifestation of solar magnetic activity and its cycle. The sunspot number, $R_z = s + 10 * g$, is reported daily and is the sum of s, the number of single spots, and ten times the number of groups, g.

The daily sunspot numbers for the northern and southern solar hemispheres are seen to exhibit different trends over the short term in Fig. 1 for the period from the maximum of cycle 22 to the current maximum of cycle 23 (de Toma 2001). This difference is a natural consequence in $\alpha\Omega$ flux-tranport dynamo models driven by a tachocline α-effect with advective transport of magnetic flux by meridional circulation (Dikpati 2001). Thus for these non-local models with meridional circulation, where the flow upwells and separates at the equator, hemispherical sunspot numbers do couple the active degrees of freedom and should be treated separately.

Nonlinear time series analysis is used to investigate whether the signature of chaotic behavior can be observed in north and south hemisphere daily sunspot number data. An additional reason to treat them separately is that for slightly phase shifted records such as the northern and southern solar hemisphere sunspot numbers, the usefulness of these methods is reduced because summing results in a loss of information.

For the purposes of this analysis, the continuous and self-consistent sunspot number data set is superior to sunspot area data that is currently available from the various observatories. These data currently suffer from significant gaps in the record making them inapplicable for techniques requiring uniform sampling. In addition, data from different sources are frequently difficult to reconcile.

The data are divided into three sections: the 22 cycle post-maximum descending section (22D), the 22-to-23 intercycle minimum (MIN) and the cycle 23 ascending

Send offprint requests to: J. S. Schweitzer,
e-mail: schweitz@phys.uconn.edu

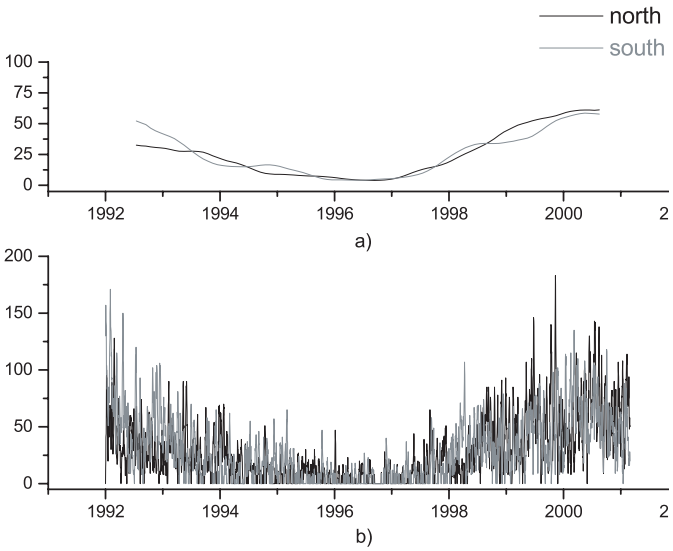

Fig. 1. Northern and southern solar hemisphere with differences observable both at minimum and maximum **a)** smoothed monthly **b)** daily sunspot numbers.

section with the current maximum (23A). Time-delay reconstruction methods are used here in a non-invariant manner as comparison tools.

To evaluate one solar maximum in its entirety we looked at the whole-sun 22 cycle maximum from a historical record starting with cycle 9. This is then compared to the 22-to-23 cycle minimum and both are compared with the results for the historical whole-sun daily sunspot number record.

Early dynamo models (Weiss 1990) addressed time scales of 10–1000 years. This prompted searches for long-term chaotic behavior on averaged monthly and annual whole-sun sunspot number data that were at times also smoothed (Carbonell 1994). These results have to be viewed with caution due to the averaging and smoothing used (Kantz & Schreiber 1997). Not surprisingly, claims of a wide range of correlation dimension values were made depending on the pre-processing of the data. Moreover, the absolute values reported are suspect for reasons ranging from small data set size to unknown noise content and filtering. Many authors found no indication of chaotic behavior at all. One of the positive results for the whole sun daily data spanning a hundred years (Wolf et al. 1985) yielded an embedding dimension of 8 and a fractal dimension of the order of 5.

2. Data

The northern and southern solar hemisphere sunspot number for almost a whole solar cycle from Jan. 1, 1992 to Feb. 28, 2001, a total of 3346 points each, are currently available from the Royal Belgian Observatory (SIDC RWC Brussels 2001), (Fig. 1). These data are integer, and have limited resolution. From the standpoint of nonlinear time series analysis the data sets are also of limited length. Furthermore, since there is no information available between the values of 0 and 1, the dynamic range accessible from these data is threshold limited.

3. Methodology

3.1. Phase space reconstruction

Nonlinear time series analysis is based on observing a system in phase space. From the measured time series we obtain a "reconstructed" phase space representation using time delay coordinates. For truly chaotic systems, if the observable links all the active degrees of freedom, this phase space representation may be used to obtain valid information about the dynamics of the source of the signal. For potentially chaotic systems, if a clear signature of approximate determinism is found and the observed variable couples the active degrees of freedom of the system, parameters defined on such representations may be used as non-invariant probes of the dynamics and, with caution, for setting limits on the dynamics at the source.

For an optimum time delay, τ, and an embedding dimension m the observed scalar time series which was uniformly sampled with interval t

$$s(t), s(2t), s(3t), s(4t) \ldots \tag{1}$$

a time delay representation is obtained of the form:

$$\overrightarrow{a}(t) = \overrightarrow{a}(t, t+\tau, t+2\tau, \ldots t+(m-1)\tau). \tag{2}$$

To determine the optimal time delay τ we use the phase space analog of the autocorrelation function, an information-theoretic concept, the average mutual information (AMI). The optimal delay for the phase space reconstruction is chosen at its first minimum. The product of this delay $\tau_{\min}$ with the value of $AMI_{\tau\min}$ on sums smaller than the data set length is singularly useful when used in a moving average manner to test for internal consistency of subsets of the data.

The autocorrelation function for these data itself has a shallow flat minimum from 8 to 12 days whereas the AMI delay was 7 days for the historical ∼55 000 point data set. Since the reliability of the choice increases with length of data set the value of 7 was used as a reference. The analyses were, however, performed for a range of delays: 3, 7, 11, 23, 31, 51 and 101 days. Though the curves presented are for $t = 7$ days, for all the delays considered, including the range defined by the broad flat minimum of the autocorraletion function, the reported behavior was observed.

The optimal dimension of the phase space is obtained using the minimum of false nearest neighbors. However, the embedding dimension is hard to determine reliably on short noisy data sets. Thus, we use it only to set a lower bound in terms of the smallest dimension in which to expect the signature of potential chaotic behavior in the correlation sum and local dimension plots.

3.2. Dynamic measures

A phase space portrait may be used to probe system dynamics. One such dynamic measure defined on this object

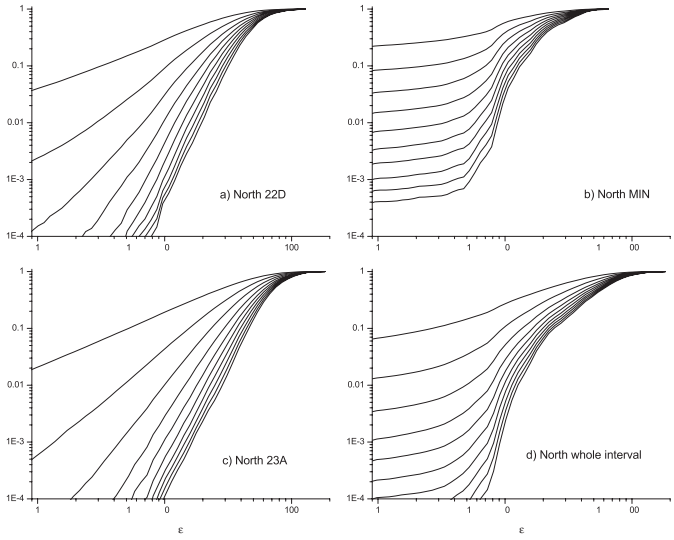

Fig. 2. Correlation sums $C(\epsilon)$ for the northern solar hemisphere **a)** 22D **b)** MIN **c)** 23A **d)** one cycle. The straight sections in 22D and 23A indicate self-similarity.

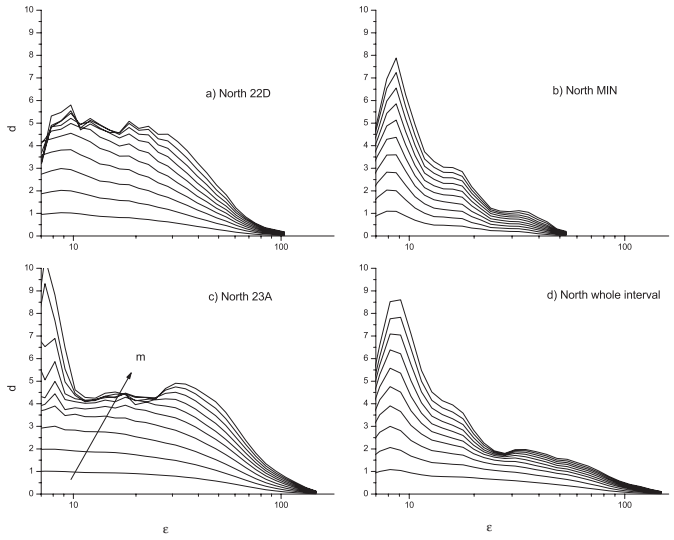

Fig. 3. Northern solar hemisphere $d(\epsilon)$ **a)** 22D **b)** MIN **c)** 23A **d)** one cycle. The flat sections in 22D and 23A indicate self-similarity.

is the correlation sum $C(\epsilon, N)$ (Kantz & Schreiber 1997; Schreiber 1999). It is the fraction of all possible point pairs that are closer than a certain distance ϵ in a particular norm

$$C(\epsilon, N) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \Theta(\epsilon - \|x_i - x_j\|). \qquad (3)$$

Here Θ is the step function. We also define:

$$d(\epsilon, N) = \frac{\partial(\ln C(\epsilon, N))}{\partial(\ln \epsilon)}. \qquad (4)$$

If $C(\epsilon, N)$ (Eq. (3)) has a constant slope with ϵ i.e. its slope $d(\epsilon, N)$ (Eq. (4)) (Kantz & Schreiber 1997) has a constant value over a range of ϵ for dimensions larger than the embedding dimension, the potential for self-similarity exists. This latter measure (Eq. (4)) for any N may be called the "local dimension" $d(\epsilon)$ since the true correlation dimension is obtained for N going to infinity and ϵ going to zero, conditions that can never be met for real data.

The accepted rule is that to reliably determine $d(\epsilon)$ we need at least $10^{\frac{d(\epsilon)}{2}}$ points (Eckman & Ruelle 1992). Since the size of our data subsets is only 1024 points, estimates of $d(\epsilon)$ have 6.02 as the upper bound. In view of the unknown absolute noise content, $d(\epsilon)$ that are obtained from these sums in the range of 4–5 are considered marginal and are used in a relative sense only as comparative dynamic measures.

For the coarsely digitized sunspot number data a possible solution is to add random noise on a (−0.5, 0.5) scale (Kantz & Schreiber 1997). In this manner the data are shaken off a coarse grid though the slope is not changed, i.e. the local dimension is not affected. In addition, since $d(\epsilon)$ is a slope, judicious averaging of the results for $C(\epsilon)$ and $d(\epsilon)$ can enhance our ability to identify regions of constant $d(\epsilon)$. Unless otherwise noted, all the graphs presented are for sunspot number data with (–0.5, 0.5) white noise added and with three point adjacent averaging of $C(\epsilon)$ and $d(\epsilon)$.

4. Results

Though correlation sums and local dimensions are not invariant, it is justified to compare them for data that are obtained with the same methodology and instrumentation such as the three subsets of the single cycle sunspot record and the north and south data. The boundaries of the 22D, MIN and 23A subsets of 1024 points were appropriately shifted in order to bracket the minimum properly. Most of the calculations were performed using the TISEAN package (Hegger et al. 1999).

4.1. Behavior at maxima and minimum

The plots for $C(\epsilon)$ for 22D, MIN and 23A and the entire interval for the north are given in Fig. 2. Near 22 cycle maximum, the sums have parallel straight sections in phase space dimensions >6 for values of $\epsilon > 7$. These correspond to horizontal sections in $d(\epsilon)$ at ∼4.5 over a range of ϵ of 8 to 30 (about half a decade) (Figs. 3a and c). At minimum, $C(\epsilon)$, (Fig. 2b) the graph flattens out at small ϵ in an S-like manner with a section parallel to the ϵ axis implying periodicity. The corresponding flat-like portion of $d(\epsilon)$ disappears. The MIN data $d(\epsilon)$ plot (Fig. 3b) has just a noise-like bump at large ϵ, which is present in all the plots, and, at lower ϵ, for increasing embedding dimension, the slopes of the curves increase rapidly. The sums (Fig. 2c) and local dimensions for $23A$ (Fig. 3c) revert to those of $22D$. In the $C(\epsilon)$ plots for the entire interval (Fig. 2d), the straight parallel portions are no longer present with the attendent loss of flatness in $d(\epsilon)$ (Fig. 3d) due to the inclusion of the MIN data.

For the south (Fig. 4) and for the whole-sun (Fig. 5) data, the results of the analysis of the 22D, MIN and 23A

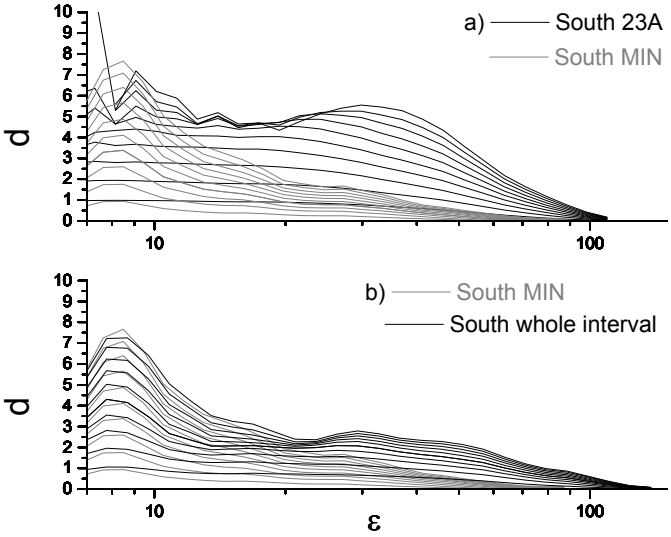

Fig. 4. Comparison of $d(\epsilon)$ for the southern solar hemisphere **a)** 23A and MIN **b)** one cycle and MIN. A flat section exists for 23A but not for MIN or whole interval.

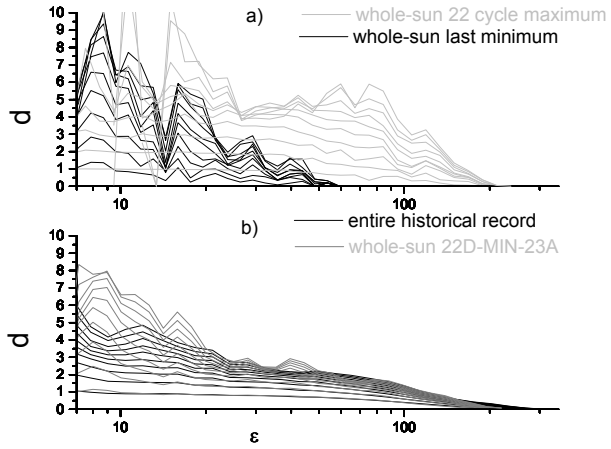

Fig. 5. a) Whole-sun $d(\epsilon)$ for 22 cycle maximum and the 22-to-23 cycle minimum showing their very different bahavior. **b)** Whole-sun $d(\epsilon)$ for the historical record since cycle 9 with 22-to-23 cycle 1992 to 2001 data with no regions of self similarity

sections are analogous. For the north and south data, $d(\epsilon)$ for 22D and 23A are strikingly similar though for the whole-sun the "flat" sections in $d(\epsilon)$ are not as flat or as long (most of these are shown in the figures). For the north, south and whole-sun the minimum $d(\epsilon)$ curves increase towards lower ϵ with no plateaus (Fig. 3b for north, Figs. 4a and b for south, Fig. 5a for the whole interval). When the 22D-MIN-23A interval (∼solar cycle) is analyzed as a whole, the north, south and whole-sun $C(\epsilon)$ curves lose their straight parallel sections and resemble the plots for the minimum (Figs. 3b and 4a). The $d(\epsilon)$ plots for the entire cycle for the south daily sunspot data set are compared to that of the minimum in Fig. 4b. Of note is that the minimum section and the whole interval south sunspot number data set $d(\epsilon)$ track in the range of values of ϵ where they both exist and neither exhibits scaling.

For all the delays considered the above behavior persisted. The maxima and the minimum differed in the above manner. For all, the minimum response resulted in loss of information on self-similarity on cycle time scales.

4.2. Whole-sun longer term data

The entire 22 cycle maximum and the 22-to-23 cycle minimum from the ∼55 000 data point whole-sun historical record were also analyzed. A comparison of the whole-sun data is given in Fig. 5. The effect of the summing of the north and south data is a less identifiable region of constant $d(\epsilon)$.

Even more striking, if we compare our shorter 22-to-23 cycle data set for the whole-sun with the historical record (Fig. 5b), though the former is definitely less well behaved, the historical record results look very similar to $d(\epsilon)$ of the 22-to-23 cycle minimum section (Fig. 5a) if one allows for the fact that there is more than an order of magnitude more data in the historical record.

4.3. Relative noise analysis

To safeguard that the results are not due to noise, a relative noise analysis using the method of artifactual correlations (Cellucci et al. 1997) was conducted. The data set is used to obtain a phase randomized surrogate. A quantitative metric D is used to test the north, south and whole-sun sunspot data sets. This analysis shows all three data sets to be clean, without additive noise, based on the amount of noise that must be added before the D metric approaches zero. The curves are not smooth because the time series are short. Figure 6a shows the results for the north, south and whole-sun data sets.

Within the limitations of the analysis due to the small size of the data sets, the north and south data show a similar response, while the whole-sun data tends to produce a lower value of D as a function of signal-to-noise ratio. This qualitative behavior is similar to what we have obtained earlier where the whole-sun data did not have as well-defined a behavior as a function of the phase space length scale, ϵ, as did the north and the south data. Similarly, the trends seen in Fig. 6b show very similar behavior for the 22D and 23A south data, with a noticeable difference for the MIN data, again consistent with what we observed previously (see Fig. 4a).

5. Conclusions

The nonlinear time series analysis of north and south sunspot number data for one solar cycle show that the maxima and minimum differ significantly. Around solar maximum we find signatures of chaotic-like behavior during both 22D and 23A. No such signature can be seen during the minimum. However, the sections of the data record during solar minimum are shown to determine the results of the analysis on scales of a solar cycle and they also

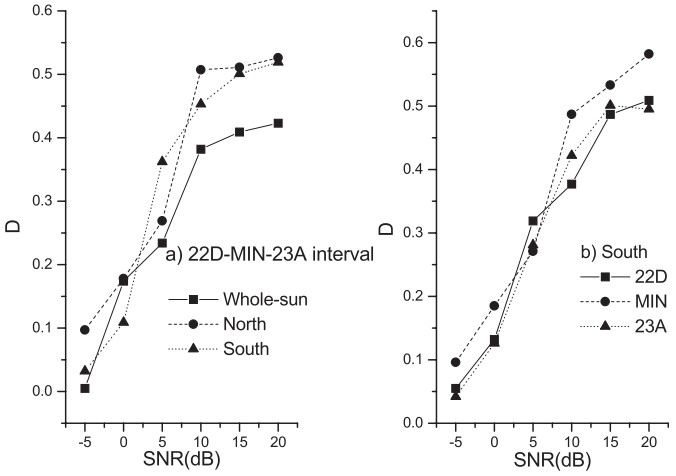

Fig. 6. a) Relative noise estimate D metric for whole interval north, south and whole-sun. **b)** Comparison of relative noise estimate D metric for the south 22A, MIN and 23D.

appear to determine the results of the historical record analysis.

When compared to the results of the same analysis for the whole sun where the hemispherical numbers are summed, the north and south results are markedly more obvious.

The analysis is threshold limited. Information is lost on smaller phase space scales because sunspot numbers go from 1 to 0 with no information in-between which is in effect thresholding. To obtain information about the self-similarity of the system over many phase space length scales seems impossible with frequent zero-one sections.

The embedding dimension and $d(\epsilon)$, on the basis of the data for the maxima, are at least 6 and 4–5, respectively. These numbers are in reasonable agreement with the results reported earlier (Wolf et al. 1985) though we do not see a region of potential self-similarity for the entire multi-cycle historical record. An increase of complexity for sunspot number data at solar maximum has also been reported using the wavelet method (Sello 2000).

The observed possible indications of self-similarity for 22D and 23A for both the north and south data support dynamo chaotic action on time scales of less than 3 years as reported for velocity shear at the tachocline.

Therefore, a nonlinear time series analysis is strongly indicated of other indicators of solar activity such as sunspot area that do not suffer from the same dynamic range threshold and statistics issues, preferably for which data can be viewed separately for the two solar hemispheres. Since the efficacy of the nonlinear time series analysis used improves with size of data set, more reliable results would be obtained with either data sets covering longer intervals or with more points from sampling at shorter intervals.

The hope is that with this better understanding of the limitations inherent in the sunspot data, future developments will permit more meaningful information to be obtained from the historical sunspot data record, the only long-term data we have.

Acknowledgements. This work was partially supported by University of Connecticut Research Foundation Grants.

We would like to thank Marijan Kostrun of the University of Connecticut Physics Department and Prof. Nalini Ravishanker and Prof. Igor Perisic from the Statistics Department at the University of Connecticut for many helpful discussions and suggestions.

References

Abarbanel, H. D. I. 1996, Analysis of Observed Chaotic Data (Springer-Verlag)
Carbonell, M., Oliver, R., & Ballester, J. L. 1994, A&A, 290, 983
Cellucci, C. J., Albano, A. M., et al. 1997, Chaos, 7, 414
Dikpati, M., & Gilman, P. A. 2001, ApJ, 559, 428
de Toma, G., White, O. R., & Harvey, K. L. 2000, ApJ, 529, 1101D
Eckman, J. P., & Ruelle, D. 1992, Physica, 56D, 185
Gao, J., & Zheng, Z. 1993, Phys. Lett. A, 181, 153
Hegger, R., Kantz, H., & Schreiber, T., The TISEAN package, 1999, CHAOS, 9, 413
Howe, R., Christense-Dalsgaard, J., et al. 2000, Science, 287, 2456 (in Reports)
Kantz, H., & Schreiber, T. 1997, Nonlinear Time Series Analysis (Cambridge University Press)
Li, K. J., Yun, H. S. & Gu, X. M. 2001 ApJ, 554, 115L
Schreiber, T. 1999, Phys. Rep., 308, 1
Sello, S. 2000, A&A, 363, 311
SIDC RWC Brussels, World Data Center for the Sunspot Index, Royal Observatory of Belguim, Bruxelles
Weiss, N. O. 1990, Phil. Trans. of Royal Soc. Lond. A, 330, 617
Wolf, A., Swift, J. B., Swinney, H. L., & Vastano, J. A. 1985, Physica, 16D, 285