Astrophysics and Space Science Proceedings 33

Manfred P. Leubner
Zoltán Vörös *Editors*

Multi-scale Dynamical Processes in Space and Astrophysical Plasmas

Springer

Stochastic Properties of Solar Activity Proxies

Kristoffer Rypdal and Martin Rypdal

Abstract The sunspot number (SSN), the total solar irradiance (TSI), a TSI reconstruction, and the solar flare index (SFI) are analyzed for long-range persistence. The SSN, TSI, and TSI reconstruction are almost certainly long-range persistent with most probable Hurst exponent $H \approx 0.7$. The SFI process, however, is either very weakly persistent ($H < 0.6$) or completely uncorrelated.

1 Introduction

The idea of long-range persistence in the solar records, on short as well as long times scales, is not new. Mandelbrot and Wallis [1] applied rescaled-range (R/S) analysis to monthly sunspot numbers (SSN) and found the characteristic bulge on the $\log(R/S)$ vs. $\log\tau$ curve for time scales τ around the period of the sunspot cycle, but a common slope corresponding to a Hurst exponent of $H \approx 0.9$ in the ranges $3 < \tau < 30$ months and $30 < \tau < 100$ years. Ruzmaikin et al. [2] obtained similar results for the SSN, and for analysis extended to the time range $100 < \tau < 3,000$ years by using a ^{14}C proxy for cosmic ray flux. Since such behavior of the R/S-curve is reproduced by adding a sinusoidal oscillation to a fractional Gaussian noise with Hurst exponent H, both groups concluded that the stochastic component of the SSN is a strongly persistent fractional noise in the time range from months to millennia. Ogurtsov [3] combined SSN and SSN reconstructions with ^{14}C proxies to obtain H in the range 0.9–1.0 in the time range $25 < \tau < 3,000$ years by

K. Rypdal (✉)
Department of Physics and Technology, University of Tromsø, Tromsø, Norway
e-mail: kristoffer.rypdal@uit.no

M. Rypdal
Department of Mathematics and Statistics, University of Tromsø, Tromsø, Norway
e-mail: martin.rypdal@uit.no

Rypdal, K. and Rypdal, M.: *Stochastic Properties of Solar Activity Proxies*. Astrophys Space Sci Proc. **33**, 227–232 (2012)
DOI 10.1007/978-3-642-30442-2_25,

means of R/S analysis and first and second order detrended fluctuation analysis (DFA(1) and DFA(2)). The statistical significance of the long-memory persistence hypothesis on time scales longer than the sunspot period has been questioned by Oliver and Ballester [4], using the so-called scale of fluctuation approach.

Renewed interest in long-range persistence has emerged from the debate on the connection between solar variability and climate change. One segment of this debate has focused on the detection of correlations and/or common statistical signatures between proxies of solar activity and climate signals on time scales of the sunspot period and shorter. Since the correlation appears to be quite weak, some works have drawn the attention to a possible "complexity-linking" which is proposed to be discernible by identification of a common long-range memory process represented by a common Hurst exponent H [5, 6]. This view was criticized by us in a recent Letter [7] where we point out that trends, like the sunspot cycle in solar signals, will create spuriously high Hurst exponents. For instance [6] (their Fig. 3B) obtain $H \approx 0.95$ for solar as well as global temperature signals. After such trends are accounted for, both solar signals and climate signals exhibit considerably lower Hurst exponents for the stochastic component of the signals.

2 Estimating Hurst Exponents from Data

The solar signals contain distinct periodicities, the most prominent being the 11-year solar cycle, and these trends will distort the result of the variogram or rescaled-range analysis. However, the detrended fluctuation analysis (DFA) [8], performs quite well on these data. The product of this analysis is the n'th order DFA fluctuation function $F^{(n)}(\Delta t)$, where Δt is the time scale [8]. If a fractional Brownian motion (fBm) with Hurst exponent $H \approx 0.5$ is superposed on a sinusoidal signal of comparable amplitude, variogram analysis will give an estimated Hurst exponent $H \approx 1$, while a third order DFA analysis (DFA(3)) will give $F^{(n)}(\Delta t) \propto \Delta t^H$ with H close to the Hurst exponent for the fBm. Thus, there are good reasons to assume that given sufficiently long data series DFA(3) would give an accurate estimate of the Hurst exponent for the underlying detrended stochastic process.

In Fig. 1a we have plotted the solar flare index (SFI), SSN and a total solar irradiance (TSI) reconstruction over the last four sunspot cycles, and an instrumental TSI (PMOD) composite for the last two cycles (all as daily averages). The thick smooth curves are moving averages. In order to extract the stochastic component the smoothed signal should be subtracted, but the signals are still strongly non-stationary in the amplitudes of the rapid fluctuations. The variance of the fluctuations around the local mean $y(t)$ is roughly proportional to $y(t)$, i.e. $\mathrm{Var}[x(t)|y(t) = y] \propto y$. This means that the mean- and amplitude-detrended signals $z(t) \equiv (x(t) - y(t))/\sqrt{y(t)}$, which is plotted in Fig. 1b, are approximately stationary. The variograms (squared DFA(0) fluctuation function) of the raw signals yield Hurst exponents in the range $0.88 < H < 0.97$ and correspond to the results

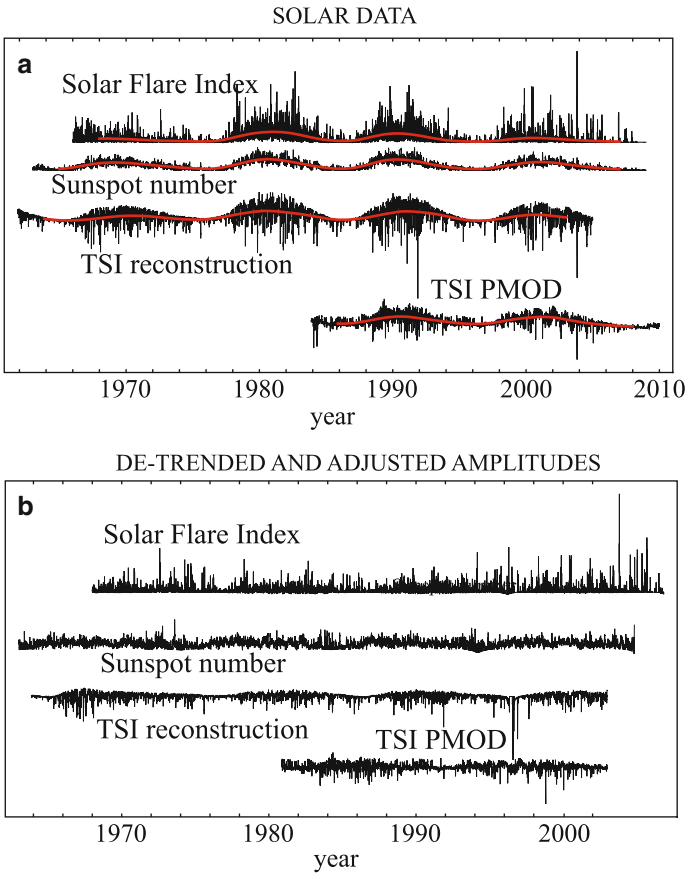

Fig. 1 (**a**) Sunspot number (SN), solar flare index (SFI), total solar irradiance (TSI) reconstruction, and the TSI PMOD composite. Smooth curves are gaussian moving averages with a 1-year standard deviation. (**b**) In this figure we have subtracted the smoothed signals from the original times series in order to de-trend the time series and then divided the detrended signals by the square root of the smoothed signals in (**a**). Prior to this amplitude adjustment, the origins for each of the signals are shifted to the minimal values of the smooth signals in (**a**)

obtained in Fig. 3B of [6]. The DFA(3) analysis, however, yields $0.55 < H < 0.67$. Results similar to the latter is obtained by computing variograms of the detrended signal in Fig. 1b.

3 A Stochastic Model for Assessing Uncertainty

The limited length of the observed data records makes it difficult to compute error bars in the estimates of Hurst exponents directly from the data. What we need to know is the PDF of estimated values $\hat{H}$ in an imaginary ensemble of realizations

of data sets of the same length as the observed record. Such an ensemble can be generated synthetically from a model that is assumed to have the same statistical properties as the observed data, including an hypothesized value of H that can be varied. From this ensemble one can construct a conditional probability density $p(\hat{H}|H)$ for obtaining an estimated value $\hat{H}$ for the Hurst exponent, given that the "true" exponent is H. Then, by means of Bayes' theorem we can obtain the conditional PDF $p(H|\hat{H})$, which gives us the probability of having a "true" Hurst exponent H provided we have estimated a value $\hat{H}$ from the observational data. The width of $p(H|\hat{H})$ gives us the error bar of our estimate.

We shall model $z(t)$ as a self-similar (in general non-Gaussian) process with self-similarity exponent H. By denoting this fractional noise process by $w_H(t)$, we can write $x(t) = y(t) + \sigma\sqrt{y(t)}\,w_H(t)$, where $y(t)$ is a 11-year-periodic oscillation with amplitude A, and $w_H(t)$ is a fractional non-Gaussian noise with unit variance. We assume that the observed signal is a realization of the stochastic model where $w_H(t)$ is a fractional noise with the measured distribution and with Hurst exponent $H \in [0, 1]$. We construct a uniform (all values of $H \in I \equiv [0, 1]$ are equally probable) sample space $\mathscr{S}$ of realizations of these processes. For each realization with a given H we measure (estimate) a value $\hat{H}$. In general $\hat{H} \neq H$. We can then, for instance, compute the conditional probability density $p(\hat{H}|H)$ from an ensemble of numerical solutions to the stochastic model where for each realization H is chosen randomly in the interval I. The conditional PDF $p(H|\hat{H})$ can also be computed directly from the ensemble or from Bayes' theorem: $p(H|\hat{H}) = p(\hat{H}|H)p(H)/p(\hat{H}$. Here $p(H)$ is by construction of $\mathscr{S}$ uniform on the unit interval I and hence $p(H) = 1$. However, $p(\hat{H})$ is not necessarily uniform and must be computed from the ensemble.

4 Results

In Fig. 2a (crosses) we show the mean estimated $\hat{H}$ computed by the DFA(3) method from ensembles of synthetic realizations of the SFI. $\hat{H}$ is estimated from DFA(3) applied to the synthetic $x(t)$ and the deviation from the straight dashed line indicates the systematic bias of the DFA(3) method. When the process is persistent ($H > 0.5$) the method performs quite well, with a slight overestimation of H when the process is strongly persistent ($H \to 1$). The circles (partly hidden by the crosses) are computed from DFA(3) applied to the fractional noise directly. The fact that the two curves fall on top of each other shows that the DFA(3) method removes the influence of the periodic trend. There is also a statistical spread in the estimated $\hat{H}$ indicated by the error bars. The spread in estimated $\hat{H}$ depends on the length of the synthetic data record which we have chosen equal to the observational record. The conditional PDF $p(H|\hat{H})$ can be computed from the same ensemble. In Fig. 2b the conditional cumulative distribution $P(H|\hat{H}) \equiv \int_{-\infty}^{H} p(H'|\hat{H})dH'$ is computed from the conditional PDF $p(\hat{H}|H)$ by means of the uniform sample space

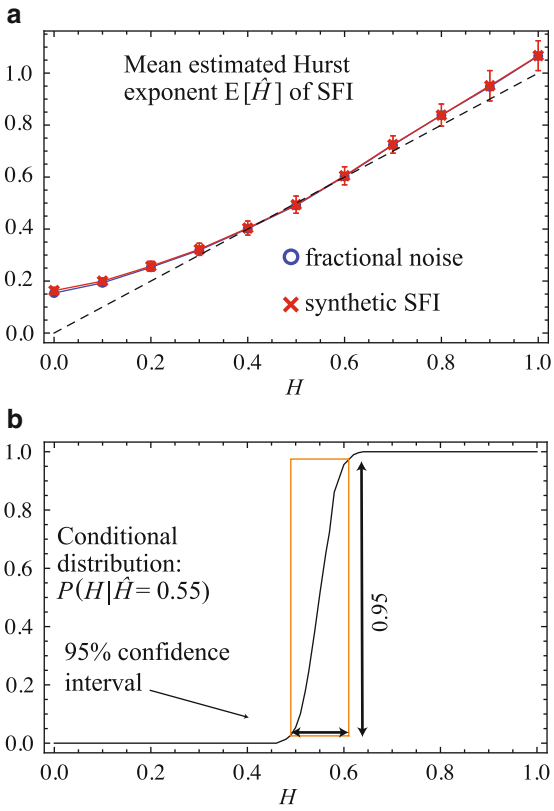

Fig. 2 (**a**) Ensemble mean $E[\hat{H}]$ of estimated Hurst exponent for SFI computed by the DFA(3) method. The crosses are computed from an ensemble of realizations generated from the stochastic model where the fractional noise $w_H(t)$ used in the model has Hurst exponent H and the PDF computed from the SFI time series. The circles (partly hidden by the crosses) are computed from DFA(3) applied to the fractional noise directly. The Gaussian statistical spread in the estimated $\hat{H}$ is given by the error bars. (**b**) The conditional cumulative distribution $P(H|\hat{H}) \equiv \int_{-\infty}^{H} p(H'|\hat{H})dH'$ computed from the conditional PDF $p(\hat{H}|H)$ by means of the uniform sample space and Bayes' theorem for $\hat{H} = 0.55$

Table 1 1. row: Hurst exponent $\hat{H}$ as estimated by DFA(3) from the observed time series. 2. row: The most probable Hurst exponents $\int H\, p(H|\hat{H})\, dH$ computed from the uniform sample space. 3. row: The 95% confidence intervals computed from the uniform sample space

	TSI reconstruction	TSI (PMOD)	Sunspot number	Solar flare index
$\hat{H}$	0.61	0.71	0.78	0.55
$\int H\, p(H\|\hat{H})\, dH$	0.62	0.70	0.73	0.54
95% confidence	$0.57 < H < 0.69$	$0.62 < H < 0.81$	$0.68 < H < 0.76$	$0.49 < H < 0.61$

and Bayes' theorem for $\hat{H} = 0.55$. From this distribution it is easy to compute the conditional mean $\mathbb{E}(H|\hat{H})$ (which is the best estimate for H given the observation $\hat{H}$) and the 95% confidence interval for this estimate. These are given in Table 1, and does not rule out the possibility that the SFI stochastic process is uncorrelated ($H = 0.5$). Note that $\mathbb{E}(H|\hat{H} = 0.55) = 0.54$, i.e. that the best estimate for H and the observation $\hat{H}$ are slightly different.

The mean estimated Hurst exponents and their 95% confidence intervals are summarized for all four proxies in Table 1. Except for the SFI this analysis indicates that there is a significant persistence in all solar activity proxies analyzed, but their

Hurst exponents are considerably lower than obtained by most previous authors. The largest H is found in the sunspot number, with the most probable value $H = 0.73$, and with 95% confidence $H < 0.76$.

References

1. Mandelbrot, B. B., and J. R. Wallis (1969), Global dependence in geophysical records, *Water Resor. Res.*, *5*, 321.
2. Ruzmaikin, A., J. Feynman, and P. Robinson (1994), Long-term persistence of solar activity, *Solar Phys*, *148*, 395.
3. Ogurtsov, M. G. (2004), New evidence for long-term persistence in the Sun's activity, *Solar Physics*, *220*, 93.
4. Oliver, R., and J. L. Ballester (1998), Is there memory in solar activity?, *Physical Review*, *58*, 5650.
5. Scafetta, N., and B. J. West (2003), Solar Flare Intermittency and the Earths Temperature Anomalies, *Phys. Rev. Lett.*, *90*, 248701.
6. Scafetta, N. and B. J. West (2005), Multiscaling Comparative Analysis of Time Series and Geophysical Phenomena, *Complexity*, *10*, 51.
7. Rypdal, M., and K. Rypdal, (2010), Testing Hypotheses about Sun-Climate Complexity Linking, *Phys. Rev. Lett.*, *104*, 128501.
8. Peng C-K, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger (1994), Mosaic organization of DNA nucleotides, *Phys. Rev. E*, *49*, 1685.