Nonlin. Processes Geophys., 21, 1059–1073, 2014
www.nonlin-processes-geophys.net/21/1059/2014/
doi:10.5194/npg-21-1059-2014

Nonlinear Processes in Geophysics
Open Access

Nonlinear fluctuation analysis for a set of 41 magnetic clouds measured by the Advanced Composition Explorer (ACE) spacecraft

A. Ojeda González[1,3,4], W. D. Gonzalez[1], O. Mendes[1], M. O. Domingues[2], and R. R. Rosa[2]

[1]DGE/CEA/National Institute for Space Research, INPE 12227-010 São José dos Campos, SP, Brazil
[2]LAC/CTE/National Institute for Space Research, INPE 12227-010 São José dos Campos, SP, Brazil
[3]Laboratory of Physics and Astronomy, IP & D/University of Paraíba Valley, UNIVAP 12244-000 São José dos Campos, SP, Brazil
[4]Junior Postdoctoral Fellow – CNPq, Brasilia-DF, Brazil

Correspondence to: A. Ojeda González (ojeda.gonzalez.a@gmail.com)

Received: 24 February 2014 – Published in Nonlin. Processes Geophys. Discuss.: 11 April 2014
Revised: 17 September 2014 – Accepted: 18 September 2014 – Published: 30 October 2014

Abstract. The statistical distribution of values in the signal and the autocorrelations (interpreted as the memory or persistence) between values are attributes of a time series. The autocorrelation function values are positive in a time series with persistence, while they are negative in a time series with anti-persistence. The persistence of values with respect to each other can be strong, weak, or nonexistent. A strong correlation implies a "memory" of previous values in the time series. The long-range persistence in time series could be studied using semivariograms, rescaled range, detrended fluctuation analysis and Fourier spectral analysis, respectively. In this work, persistence analysis is to study interplanetary magnetic field (IMF) time series. We use data from the IMF components with a time resolution of 16 s. Time intervals corresponding to distinct processes around 41 magnetic clouds (MCs) in the period between March 1998 and December 2003 were selected. In this exploratory study, the purpose of this selection is to deal with the cases presenting the three periods: plasma sheath, MC, and post-MC. We calculated one exponent of persistence (e.g., α, β, Hu, Ha) over the previous three time intervals. The persistence exponent values increased inside cloud regions, and it was possible to select the following threshold values: $\langle \alpha_{(j)} \rangle = 1.392$, $\langle Ha_{(j)} \rangle = 0.327$, and $\langle Hu_{(j)} \rangle = 0.875$. These values are useful as another test to evaluate the quality of the identification. If the cloud is well structured, then the persistence exponent values exceed thresholds. In 80.5 % of the cases studied, these tools were able to separate the region of the cloud from neighboring regions. The Hausdorff exponent (Ha) provides the best results.

1 Introduction

Coronal mass ejections (CMEs) are massive expulsions of magnetized plasma from the solar atmosphere (see, e.g., Dasso et al., 2005, and references therein). As a consequence of this ejection, CMEs can form confined magnetic structures with both extremes of the magnetic field lines connected to the solar surface, extending far away from the Sun into the solar wind (SW). Solar Ejecta – also known as interplanetary coronal mass ejections (ICMEs) – are the interplanetary manifestation of CME events (see, e.g., Dasso et al., 2005, and references therein). The important subset of ICMEs, known as interplanetary magnetic clouds (MCs), a term introduced by Burlaga et al. (1981), is characterized fundamentally by enhanced magnetic field strengths with respect to solar wind ambient values (Klein and Burlaga, 1982; Burlaga, 1991). A comprehensive study about the properties of MCs at 1 astronomical unit (AU) was approached by Ojeda et al. (2013, 2014) and Klausner et al. (2014).

The test for independence and search for correlations in a time series can be carried out using an analytical tool from nonlinear dynamics: the estimation of the Hurst exponent

(Hurst et al., 1965). Mandelbrot and Wallis (1969) first used it to study a series of a monthly sunspot of 200 years. It had a Hurst exponent (with rescaled range – R/S analysis) significantly larger than 0.5. On others papers such as Ruzmaikin et al. (1994), they showed that solar activity has long-term persistence when exploring time series of ^{14}C (Carbone-14). Calzadilla and Lazo (2001) and Wei et al. (2004) studied time series of D_{st} geomagnetic index, which showed chaotic properties in association with self-affine fractals. The D_{st} index can be viewed as a self-affine fractal dynamic process as a result of SW–magnetosphere interactions. In fact, the behavior of the D_{st} index, with a Hurst exponent $Hu \approx 0.5$ (power-law exponent $\beta \approx 2$) at high frequency, is similar to that of Brownian motion. Therefore, perhaps the dynamical invariants of some physical parameters of the solar wind, specifically the MCs, may have spectral characteristics similar to Brownian motion.

Price and Newman (2001) analyzed the behavior of a solar wind data set – interplanetary magnetic fields (IMFs) and solar wind speed – with a 1 min resolution from September 1978 to July 1979, using the ISEE-3 spacecraft. They showed the time series, the power spectrum and the R/S analysis for the IMF B_z component for the month of March 1979. The B_z time series was self-similar for all time scales, highly coherent for time scales less than 1 day, and only slightly coherent for time scales greater than 1 day. In addition, they found self-similarity and coherence properties when calculating β-power spectrum values to $v\,B_z$, geomagnetic auroral electrojet (AE) index, and the horizontal (H) component of the Earth's magnetic field. Tsurutani et al. (1990) studied the nonlinear response of AE to the IMF B_s driver. For this, the similarities between the power spectrums of the two measurements are analyzed. Sharma and Veeramani (2011) analyzed long-range correlations, using detrended fluctuation analysis (DFA) based on autocorrelations functions in auroral electrojet index lower (AL) data for the period 1978–1988.

This paper is a detailed study of persistence in magnetic clouds. The manuscript is divided in five sections. A review about persistence analysis is presented in Sect. 2. Section 3 presents the data set and the analyzed periods. Section 4 presents the methodology implemented. In Sect. 5, the results are discussed. In Sect. 6, the conclusions are shown.

2 Persistence in time series

In this work, persistence analysis is used to study IMF time series. The purpose throughout this section is to review the physical–mathematical concepts of these tools.

The main attributes of a time series include the statistical distribution of values in the signal and the autocorrelations (interpreted as the memory or persistence) between values. Positive values of autocorrelation function, $r_k = C_k / C_0$, indicate persistence, while negative values indicate anti-persistence. For example, in a Gaussian white noise, if each time series value is independent of other values, then the correlation and persistence are 0. Time series of Brownian motion is derived from a running sum of a Gaussian white-noise sequence. If the values in a time series of a Brownian motion are well correlated, then this time series exhibits long-range persistence. In summary, the persistence can be grouped in three categories: strong, weak, or nonexistent.

The word "memory" is the common term used to explain and understand the persistence concept in a time series. The values in the time series could be considered "intelligent entities" that have "knowledge" or memory of the existence of other "individuals" (values). The ideal case of maximum persistence is when each value has memory of all previous values of the time series. Thus, a strong correlation implies a memory of previous values in the time series. Persistence is a mathematical number to measure how good the "mean memory" is in a time series. The long-range persistence in a time series could be studied using semivariograms, rescaled range, detrended fluctuation analysis, Fourier spectral analysis, or wavelet variance analysis, respectively (e.g., Malamud and Turcotte, 1999).

A statistically self-similar fractal can be defined with the function $f(rx, ry)$ (with the scaling factor r) in two-dimensional xy space. This fractal is, by definition, isotropic, and the previous function is statistically similar to $f(x, y)$. It is quantified by the fractal relation $N_i \sim r_i^{-D}$, where the number of objects, N_i, and the characteristic linear dimension, r_i, are related by a power law, and the constant exponent, D, is the fractal dimension (Turcotte, 1997).

A statistically self-affine fractal can be defined with the function $f(rx, r^{Ha} y)$ (generally not isotropic) in two-dimensional xy space, where Ha is called the Hausdorff exponent. The previous function is statistically similar to $f(x, y)$ (Mandelbrot, 1983; Voss, 1985b) and the relationship between Ha and D is $Ha = 2 - D$ (e.g., Malamud and Turcotte, 1999). If $Ha = 1$, then the self-affine fractal is self-similar at the same time. Brownian motion is an example of the self-affine time series.

The power spectrum (Priestley, 1981), a measure of long-range persistence and anti-persistence, is used frequently in the analysis of geophysical time series (e.g., Pelletier and Turcotte, 1999). The periodogram is a plot of power-spectral density (PSD) of a signal $S(f)$ vs. frequency f, and it is an estimate of the spectral density of a signal. For a time series that is self-affine – $S(f) \sim f^{-\beta}$ (e.g., Voss, 1985a) –, the slope of the best-fit straight line from $\log(S(f))$ vs. $\log(f)$ is a constant called β-power spectrum exponent. The relationship between β, Ha, and D was obtained by Voss (1986):

$$\beta = 2Ha + 1 = 5 - 2D. \tag{1}$$

In the paper by Malamud and Turcotte (1999), validation intervals for a self-affine fractal were derived: $0 \leq Ha \leq 1$,

$1 \le D \le 2$, and $1 \le \beta \le 3$. Then, in a time series of a Brownian motion, the exponent values are $Ha = 1/2$, $D = 3/2$, $\beta = 2$ while a white noise has $\beta = 0$. Hausdorff exponent is only applicable for self-affine time series with validation intervals from $0 \le Ha \le 1$. However, β is a measure of the strength of persistence valid for all β, not just $1 \le \beta \le 3$ (Malamud and Turcotte, 1999). An anti-persistent time series has $\beta < 0$ and persistent time series has $\beta > 0$, respectively.

Mandelbrot and Ness (1968) developed a method to study a self-affine time series, the semivariogram, γ_k, scale with k, the lag, such that $\gamma_k \sim k^{2Ha}$, that is:

$$\gamma_k = 2^{-1}(N-k)\sum_{n=1}^{N-k}(y_{n+k}-y_n)^2. \quad (2)$$

For the uncorrelated Gaussian white noise ($\beta = 0$), the semivariogram is about $\gamma_k = 1$ (the same as the variance: $V_a = 1$). For $\beta = 1$, 2, and 3, good correlations are obtained by Malamud and Turcotte (1999, p. 40) with the expression $\gamma_k \sim k^{2Ha}$.

Following Malamud and Turcotte (1999), another alternative method to measure the persistence in time series was developed by Hurst (1951) and Hurst et al. (1965). They studied the Nile River flow as a time series to introduce the concept of rescaled range (R/S) method used to calculate the scaling exponent (Hurst exponent), Hu, to give quantitative measure of the persistence of a signal. Hurst (1951) and Hurst et al. (1965) found empirically the power-law relation:

$$\left[\frac{R(\tau)}{S(\tau)}\right]_{\text{av}} = \left(\frac{\tau}{2}\right)^{Hu}, \quad (3)$$

where the successive subintervals τ vary over all N values in the time series, y_n. The running sum, y_m, is

$$y_m = \sum_{n=1}^{m}\left(y_n - \overline{y}_N\right). \quad (4)$$

The range is defined by $R_N = (y_m)_{\max} - (y_m)_{\min}$ with $S_N = \sigma_N$, where $\overline{y}_N$ and σ_N are the mean and standard deviation of all N values in the time series, y_n. The R/S analysis is a statistical method to analyze long records of natural phenomena (Vanouplines, 1995).

Tapiero and Vallois (1996) found that $0.5 < Hu \le 1.0$ implies persistence and that $0 \le Hu < 0.5$ implies anti-persistence. This would imply that (Tapiero and Vallois, 1996; Malamud and Turcotte, 1999):

$$\beta = 2Hu - 1 = 2Ha + 1. \quad (5)$$

Equation (5) only has a small validation region (see Malamud and Turcotte, 1999, Figs. 17 and 25). This result should be considered when one exponent is derived from another.

Other technique (called detrended fluctuation analysis – DFA) to study persistence in time series was introduced by Peng et al. (1994). This tool could also be used to study persistence on IMF time series.

The fluctuation function $F(L)$ is constructed over the whole signal at a range of different window size L, where $F(L) \sim L^{\alpha}$. The obtained exponent, α, is similar to the Hurst exponent, but it also may be applied to nonstationary signals, this is a great advantage. DFA measures scaling exponents from nonstationary time series for determining the statistical self-affinity of an underlying dynamical nonlinear process (e.g., Veronese et al., 2011). It is useful to characterize temporal patterns that appear to be due to long-range memory stochastic processes (Veronese et al., 2011).

Bryce and Sprague (2012) reported that DFA asymptotically provides good results for stationary time series, which is a characteristic of several techniques of time series analysis; nonstationarity remains the biggest problem in a time series analysis. However, DFA is a commonly used technique, in the context of persistence analysis, to work with nonstationarity time series. Furthermore, they found a little problem when DFA is applied in time series with nonlinear trends, and other limitation in the partitioning scheme of the DFA for short data sets is reported. The weak point in the previous work was that they do not offer a clear solution to the reported limitations. And it is not included in this study. For a detailed description of this step-by-step method, see Peng et al. (1994), Little et al. (2006), Baroni et al. (2010), and Veronese et al. (2011).

Based on the Wiener–Khinchin theorem (Kay and Marple, 1981), it is possible to show that the two exponents β (from PSD) and α (from DFA) are related by

$$\beta = 2\alpha - 1. \quad (6)$$

For fractional Brownian motion, we have $1 \le \beta \le 3$, and then $1 \le \alpha \le 2$. The exponent of the fluctuations can be classified according to a dynamic range values (Kantelhardt et al., 2002; Bashan et al., 2008; Zheng et al., 2008):

- $\alpha < \frac{1}{2}$: anti-correlated, anti-persistence signal.
- $\alpha \cong \frac{1}{2}$: uncorrelated, white noise, no memory.
- $\alpha > \frac{1}{2}$: long-range persistence.
- $\alpha \cong 1$: $1/f$ noise or pink noise.
- $\alpha > 1$: nonstationary, random walk-like, unbounded.
- $\alpha \cong \frac{3}{2}$: Brownian noise or red noise.

Polynomial of different order could be used during computational implementation of the DFA method. For example, DFAn uses polynomial fits of order n (Buldyrev et al., 1995). DFA1 (used in this work) only removes constant trends in the time series, and it is equivalent to Hurst R/S analysis. The effect of trends on DFA was studied in Hu et al. (2001), and the relation to the power spectrum method is presented in Heneghan and McDarby (2000). Veronese et al. (2011) showed that the DFA method is especially useful for short records of stochastic and nonlinear processes.

Table 1. Solar wind data studied (from Huttunen et al., 2005).

No.	Year	Shock, UT	MC start, UT	MC stop, UT	Post-MC, UT
01	1998	6 Jan, 13:19	7 Jan, 03:00	8 Jan, 09:00	9 Jan, 15:00
02		3 Feb, 13:09	4 Feb, 05:00	5 Feb, 14:00	6 Feb, 23:00
03		4 Mar, 11:03	4 Mar, 15:00	5 Mar, 21:00	7 Mar, 03:00
04		1 May, 21:11	2 May, 12:00	3 May, 17:00	4 May, 22:00
05		13 Jun, 18:25	14 Jun, 02:00	14 Jun, 24:00	15 Jun, 22:00
06		19 Aug, 05:30	20 Aug, 08:00	21 Aug, 18:00	23 Aug, 04:00
07		24 Sep, 23:15	25 Sep, 08:00	26 Sep, 12:00	27 Sep, 16:00
08		18 Oct, 19:00	19 Oct, 04:00	20 Oct, 06:00	21 Oct, 08:00
09		8 Nov, 04:20	8 Nov, 23:00	10 Nov, 01:00	12 Nov, 02:00
10		13 Nov, 00:53	13 Nov, 04:00	14 Nov, 06:00	15 Nov, 08:00
11	1999	18 Feb, 02:08	18 Feb, 14:00	19 Feb, 11:00	20 Feb, 08:00
12		16 Apr, 10:47	16 Apr, 20:00	17 Apr, 18:00	18 Apr, 16:00
13		8 Aug, 17:45	9 Aug, 10:00	10 Aug, 14:00	11 Aug, 18:00
14	2000	11 Feb, 23:23	12 Feb, 12:00	12 Feb, 24:00	13 Feb, 12:00
15		20 Feb, 20:57	21 Feb, 14:00	22 Feb, 12:00	23 Feb, 10:00
16		11 Jul, 11:22	11 Jul, 23:00	13 Jul, 02:00	14 Jul, 05:00
17		13 Jul, 09:11	13 Jul, 15:00	13 Jul, 24:00	14 Jul, 09:00
18		15 Jul, 14:18	15 Jul, 19:00	16 Jul, 12:00	17 Jul, 05:00
19		28 Jul, 05:53	28 Jul, 18:00	29 Jul, 10:00	30 Jul, 02:00
20		10 Aug, 04:07	10 Aug, 20:00	11 Aug, 08:00	11 Aug, 20:00
21		11 Aug, 18:19	12 Aug, 05:00	13 Aug, 02:00	13 Aug, 23:00
22		17 Sep, 17:00	17 Sep, 23:00	18 Sep, 14:00	19 Sep, 05:00
23		2 Oct, 23:58	3 Oct, 15:00	4 Oct, 14:00	5 Oct, 13:00
24		12 Oct, 21:36	13 Oct, 17:00	14 Oct, 13:00	15 Oct, 09:00
25		28 Oct, 09:01	28 Oct, 24:00	29 Oct, 23:00	30 Oct, 22:00
26		6 Nov, 09:08	6 Nov, 22:00	7 Nov, 15:00	8 Nov, 08:00
27	2001	19 Mar, 10:12	19 Mar, 22:00	21 Mar, 23:00	23 Mar, 24:00
28		27 Mar, 17:02	27 Mar, 22:00	28 Mar, 05:00	28 Mar, 12:00
29		11 Apr, 15:18	12 Apr, 10:00	13 Apr, 06:00	14 Apr, 02:00
30		21 Apr, 15:06	21 Apr, 23:00	22 Apr, 24:00	24 Apr, 01:00
31		28 Apr, 04:31	28 Apr, 24:00	29 Apr, 13:00	30 Apr, 02:00
32		27 May, 14:17	28 May, 11:00	29 May, 06:00	30 May, 01:00
33		31 Oct, 12:53	31 Oct, 22:00	2 Nov, 04:00	3 Nov, 10:00
34	2002	23 Mar, 10:53	24 Mar, 10:00	25 Mar, 12:00	26 Mar, 14:00
35		17 Apr, 10:20	17 Apr, 24:00	19 Apr, 01:00	20 Apr, 02:00
36		18 May, 19:44	19 May, 04:00	19 May, 22:00	20 May, 16:00
37		1 Aug, 23:10	2 Aug, 06:00	2 Aug, 22:00	3 Aug, 14:00
38		30 Sep, 07:55	30 Sep, 23:00	1 Oct, 15:00	2 Oct, 07:00
39	2003	20 Mar, 04:20	20 Mar, 13:00	20 Mar, 22:00	21 Mar, 07:00
40		17 Aug, 13:41	18 Aug, 06:00	19 Aug, 11:00	20 Aug, 16:00
41		20 Nov, 07:27	20 Nov, 11:00	21 Nov, 01:00	22 Nov, 15:00

The four techniques explained previously are used in this work. Some models were tested to successfully reconstruct the magnetic structure of MCs (Dasso et al., 2005; Ojeda et al., 2013), which implies that a memory exists in the time series of IMF. We hypothesize that the magnetic field inside these structures has greater persistence than ambient solar wind. If the previous hypothesis is true, then the persistence exponent could be transformed in an auxiliary tool to study MCs. We decided to test the four techniques because there is only a small validation region between them (see Malamud and Turcotte, 1999, Figs. 17 and 25). The ideal is to use as many techniques as possible to measure the persistence and compare them.

3 IMF data set

In this work, we use data from the IMF-components (Advanced Composition Explorer (ACE) spacecraft/Magnetic Field Experiment (MAG); Lepping et al., 1995) with time resolution of 16 s and using geocentric solar magnetospheric coordinate system (GSM). We work with 41 of 80 events (73 MCs and 7 cloud candidates) identified by Huttunen et al. (2005). These events in chronological order are shown in Table 1 ((see more details in Ojeda et al., 2013, 2014), where the same data set was studied with other techniques: the spatio–temporal entropy and discrete wavelet transform). The columns from the left to the right give a numeration of the events, year, shock time (UT), MC start time (UT), MC end time (UT), and the end time (UT) of the third region, respectively. In this exploratory study, the purpose of this selection is to deal with the cases presenting the three periods (clear pre-MC or plasma sheath, MC, and post-MC).

4 Methodology

To calculate the persistence exponents, the following computational programs are used:

1. If we installed GNU/Octave, then a hurst(x) function is created; for example, in /usr/share/octave/3.0.1/m/signal/. The function is used to calculate the Hurst exponent (*Hu*).

2. Following the work of Malamud and Turcotte (1999), we did a program in GNU/Octave to calculate the Hausdorff exponent (see Appendix A).

3. A program using GNU/Octave by McSharry and Malamud (2010) is implemented to calculate the β exponent.

4. A fast Matlab implementation[1] of the DFA algorithm by Little et al. (2006) is performed.

The behavior of the persistence in time series of the IMF components, measured by the ACE spacecraft with a time resolution of 16 s, is explored. We study the persistence between time series corresponding to sheaths, MCs, and a quiet SW after the MC (post-MC) with equivalent time duration to it. We calculate one exponent of persistence (e.g., α, β, *Hu*, *Ha*) over each of three time intervals corresponding to distinct processes. For example, the persistence in the case numbered as 1 in Table 1 is studied. The interval from 6 January 13:19 UT to 7 January 02:59 UT was classified as the sheath region. In the sheath, the persistence exponents to B_x component are calculated. These values are $\alpha = 1.27$, $\beta = 1.71$, $Hu = 0.86$, and $Ha = 0.31$, respectively.

[1] http://www.maxlittle.net/software/

Table 2. We calculate the persistence in the IMF components using four different methods: β exponent of power spectrum, α exponent of DFA, Hurst of R/S analysis, and Hausdorff *Ha* exponent of semivariogram, respectively. The interval from 6 January 13:19 UT to 7 January 02:59 UT 1998 was classified as sheath. The intervals 7 January 03:00 UT to 8 January 09:00 UT and from 8 January 09:01 UT to 9 January 15:00 UT were classified as MC and solar wind after the MC, respectively. Dates are shown in Table 1, event no. 1.

Event no. 1	α	β	*Hu*	*Ha*
B_x:				
Sheath	1.27	1.71	0.86	0.31
MC	1.41	1.60	0.89	0.31
Post-MC	1.31	1.70	0.87	0.31
B_y:				
Sheath	1.34	1.68	0.87	0.27
MC	1.52	1.55	0.91	0.42
Post-MC	1.37	1.65	0.88	0.31
B_z:				
Sheath	1.39	1.65	0.85	0.31
M	1.45	1.75	0.90	0.36
Post-MC	1.23	1.64	0.86	0.23
Mean Values:	$\langle\alpha_{(j)}\rangle \pm \sigma$	$\langle\beta_{(j)}\rangle \pm \sigma$	$\langle Hu_{(j)}\rangle \pm \sigma$	$\langle Ha_{(j)}\rangle \pm \sigma$
Sheath	1.33 ± 0.06	1.68 ± 0.03	0.86 ± 0.01	0.30 ± 0.02
MC	1.46 ± 0.06	1.64 ± 0.11	0.90 ± 0.01	0.37 ± 0.05
Post-MC	1.30 ± 0.07	1.66 ± 0.04	0.87 ± 0.01	0.28 ± 0.05

The interval from 7 January 03:00 UT to 8 January 09:00 UT is the MC region. The post-MC region was selected from 8 January 09:01 UT to 9 January 15:00 UT. The persistence exponents are shown in Table 2, rows 4 and 5, respectively.

The previous methodology is extended for the other two components; i.e., B_y and B_z, respectively. The results are shown in Table 2, rows 6–13 .

MCs exhibit flux-rope characteristics: a large-scale winding of a closed magnetic structure that is nearly force-free. It is possible to see anisotropy of magnetic field fluctuations in an average interplanetary MC at 1 AU (Narock and Lepping, 2007; Ojeda et al., 2013, 2014). We do not expect to find the same behavior in all three components by the existence of anisotropy. The anisotropic behavior, in our opinion, is caused by the geometry of flux-rope and the axis inclination angle. We are interested in a single value to characterize the persistence in the IMF. For this reason, a mean persistence value using the three IMF components was calculated at each time. It is the only form that we found to quantify the persistence in all structures and to minimize the anisotropy in the calculation. The mathematical expressions can be generalized in the following equations:

$$\langle\beta_{(j)}\rangle = \frac{1}{3}\sum_{i=1}^{3}\beta_{(j)}^{(i)}; \tag{7}$$

$$\langle \alpha_{(j)} \rangle = \frac{1}{3}\sum_{i=1}^{3} \alpha_{(j)}^{(i)}; \tag{8}$$

$$\langle Hu_{(j)} \rangle = \frac{1}{3}\sum_{i=1}^{3} Hu_{(j)}^{(i)}; \tag{9}$$

$$\langle Ha_{(j)} \rangle = \frac{1}{3}\sum_{i=1}^{3} Ha_{(j)}^{(i)}. \tag{10}$$

The angle brackets $\langle \cdots \rangle$ denote an average of the IMF components ($i = 1, 2, 3 = B_x, B_y, B_z$), and the standard deviation is also calculated. Each of the three regions are represented in one j value: $j = 1 \equiv$ sheath, $j = 2 \equiv$ MC, $j = 3 \equiv$ post-MC.

In Table 2, the average and standard deviation values for all persistence exponents are shown. In Table 2, as we thought, the persistence values increase inside the MC. This increase, according to the hypothesis raised at the end of Sect. 2, was expected. The previous idea is not always true when using the spectral-power β exponent. However, one of the main problems in using a discrete Fourier transform is spectral variance and leakage (Priestley, 1981; Percival and Walden, 1993). This shows a range of uncertainty in the values of β. The other problem is the nonstationarity of the IMF components. The previous study was generalized for a group of 41 events shown in Table 1, and will be discussed in next section.

5 Results and discussion

Initially, the persistence analysis was done to establish a preliminary categorization of the periods in the SW related to the MC occurrences.

Persistence analysis on the IMF variation

The methodology that uses the persistence exponents (see Sect. 4) is applied to 41 events. Using Eq. (7), the $\langle \beta_{(j)} \rangle$ values are calculated. To make a comparison between all events, it is necessary to build a histogram. Figure 1a is a histogram built from a frequency table of $\langle \beta_{(j)} \rangle$ values plotted in Fig. 2a. The $\langle \beta_{(j)} \rangle$ values for the sheath, MC, and post-MC regions were plotted as gray, black, and white bars, respectively. The bars have a uniform distribution from $1.5 < \langle \beta_{(j)} \rangle < 1.8$. For $\langle \beta_{(j)} \rangle < 1.5$, there are 7 of 41 sheath, 2 of 41 MCs, and 15 of 41 post-MC events, while for $\langle \beta_{(j)} \rangle > 1.8$ there are 3 of 41 sheath, 9 of 41 MCs, and 3 of 41 post-MC events, respectively. As previously stated, the $\langle \beta_{(j)} \rangle$ exponent is not suitable to measure the persistence in the data set used in this work. Nevertheless, the largest values of $\langle \beta_{(j)} \rangle$ were found in the MCs.

Figure 1b has the same format as Fig. 1a, but for the $\langle \alpha_{(j)} \rangle$ exponent. For $\langle \alpha_{(j)} \rangle > 1.4$, we have 6 of 41 sheath, 29 of 41 MCs, and 3/41 post-MC events, respectively. Thus, we have many MCs with the large alpha values. For $1.0 < \langle \alpha_{(j)} \rangle < 1.3$, the number of events by regions is 21 of 41 in the sheath, 3 of 41 in the MC, and 23 of 41 in the post-MC. In MC events, the separation of the $\langle \alpha_{(j)} \rangle$ values to the right corner is an interesting result. In Fig. 1c and d, approximately 30 of 41 MC events have the large values of the persistence exponents. One difficulty in studying the persistence is the time series extension (Veronese et al., 2011).

The $\langle \beta_{(j)} \rangle$ values for the 41 events (Sheath, MCs, and post-MCs) are shown in Fig. 2a. The three intervals of time for each event are plotted as "□", "⊗", and "△" symbols, respectively. The error bar represents the standard deviation for each value. It shows the power spectral density (PSD) scaling exponent $\langle \beta_{(j)} \rangle$ as a self-affine fractal ($1 < \langle \beta_{(j)} \rangle < 2$), but there is no pattern that allows the separation of MCs from the other two cases; a total of 18/41 events exist in which the clouds do not have the larger values. We understand that, in nonstationary time series, the Fourier transform is not suitable, because the core functions of the transform is composed of sines and cosines.

For short time series, DFA can detect the correlation length more accurately than the PSD scaling exponent (β) (Veronese et al., 2011). The alpha exponent value is not affected by spectral variance and leakage, and it is possible to use in nonstationary time series. Figure 2b has the same format as Fig. 2a, but was built for $\langle \alpha_{(j)} \rangle$ exponent using the Eq. (8). The results show $\langle \alpha_{(j)} \rangle$ values from 1.00 to 1.60; i.e., long-range persistence and some MCs with typical values of a Brownian noise ($\langle \alpha_{(j)} \rangle \cong 1.50$).

In 38 of the 41 events, the alpha ($\langle \alpha_{(j)} \rangle$) value in the MC ("⊗") is larger than the one in the sheath ("□"), respectively. It is noteworthy that there are some exceptions, such as events 5, 20, and 25 in Table 1. However, in the context of the present analysis, we did not investigate each of these cases in detail, because they are only a few of the 41 time series. However, this is a study to be carried out later, because they are important to redefine the boundaries of the clouds.

The Hurst exponent was presented in Sect. 4 as an useful methodology to study MCs. Using Eq. (9), the $\langle Hu_{(j)} \rangle$ exponents in the three regions are calculated. Figure 2c has the same format as Fig. 2a and b, respectively, except for the $\langle Hu_{(j)} \rangle$ exponent. Similar to Fig. 2b, the $\langle Hu_{(j)} \rangle$ exponents have larger values in the MCs. Nevertheless, 4 of 41 MCs (events 11, 19, 28, and 30) do not have large $\langle Hu_{(j)} \rangle$ exponents in the MC region. None of these cases coincide with the three events (5, 20, and 25) when the alpha exponent is used. This causes a certain degree of distrust in the identification of these clouds, but also suggests that all techniques must be used together to increase the confidence level of the results. Nevertheless, for 34 of 41 events, both exponents have large values in the cloud region.

The last tool we use is the Hausdorff exponent (Ha). To calculate the mean Hausdorff exponents, Eq. (10) is used. In Fig. 2d, the $\langle Ha_{(j)} \rangle$ exponents have largest values in the MC regions; only 2 of 41 MCs (events 10 and 28) do not have the largest $\langle Ha_{(j)} \rangle$ exponents. Thus, this tool provides the best results.

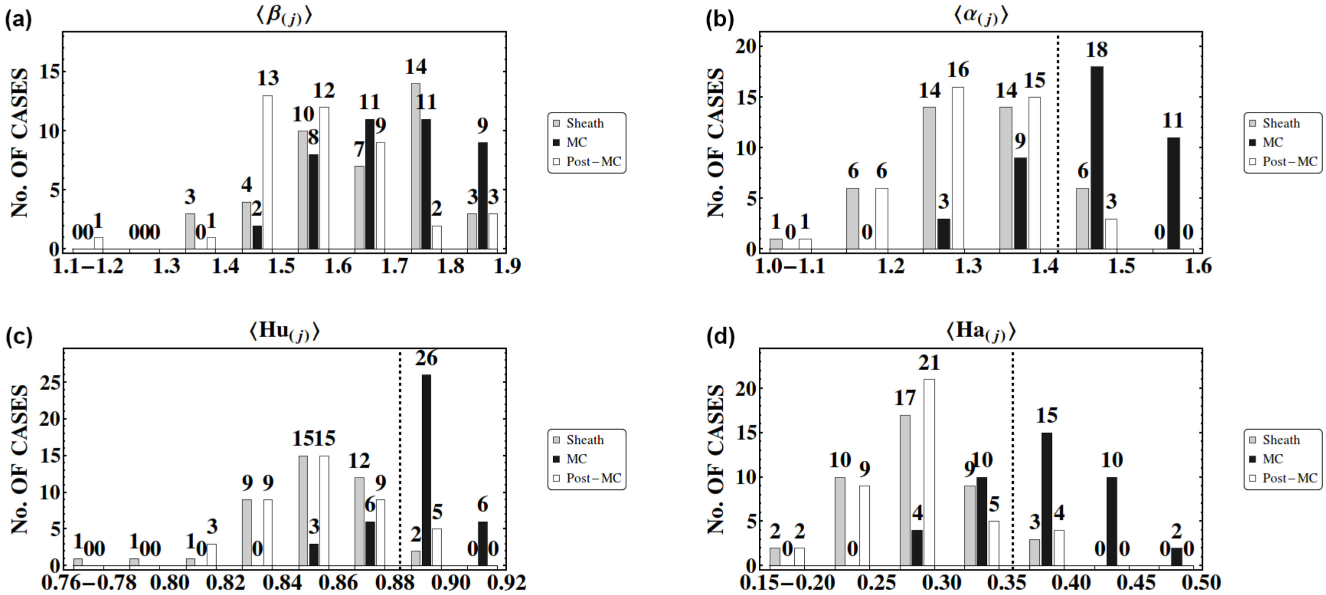

Figure 1. In **(a)**, a histogram is construct from a frequency table of $\langle\beta_{(j)}\rangle$ values, plotted in Fig. 2a. We want to have a better view of the distribution of $\langle\beta_{(j)}\rangle$ values between the three regions. **(b)**, **(c)**, and **(d)** are similar to **(a)**, except for the $\langle\alpha_{(j)}\rangle$, $\langle Hu_{(j)}\rangle$, and $\langle Ha_{(j)}\rangle$ exponents, respectively.

In conclusion, the PSD scaling exponent is not a suitable tool to study persistence in IMF components in the SW. The three exponents report the largest persistence in 33 of 41 MC regions. In 80.5 % of 41 cases, these tools are able to separate the region of the cloud from neighboring regions.

In Fig. 3, the histogram shows the number of cases vs. temporal extension (in hours) of MCs and plasma sheaths, respectively. Temporal extension is largest in the MCs. However, in the previous figures, there is a pattern in the persistence values between all MC events. We believe that these results are valid, because we know that MCs are organized structures in the plasma (Ojeda et al., 2005, 2013, 2014) that have an increase of "memory" in the time series.

We considered a better way to view these results. Thus, the average values for each exponent from 41 events and for each of the three regions are calculated. The equations for calculating the average values are

$$\langle\beta_{(j)}\rangle_T = \frac{1}{N}\sum_{i=1}^{N}\langle\beta_{(j)}\rangle^{(i)}, \tag{11}$$

$$\langle\alpha_{(j)}\rangle_T = \frac{1}{N}\sum_{i=1}^{N}\langle\alpha_{(j)}\rangle^{(i)}, \tag{12}$$

$$\langle Hu_{(j)}\rangle_T = \frac{1}{N}\sum_{i=1}^{N}\langle Hu_{(j)}\rangle^{(i)}, \tag{13}$$

$$\langle Ha_{(j)}\rangle_T = \frac{1}{N}\sum_{i=1}^{N}\langle Ha_{(j)}\rangle^{(i)}, \tag{14}$$

with $N=41$ and $j=1\equiv$ sheath, $j=2\equiv$ MC, and $j=3\equiv$ post-MC.

The calculation of the standard deviation shows how much variation or dispersion exists from the average. If a rectangular area is built using the mean and standard deviation, then there is a validity region in which all exponents join up. Following the above idea, the panels of Fig. 4 are built. In Fig. 4a, the black points $(\langle\alpha_{(j)}\rangle_T, \langle\beta_{(j)}\rangle_T)$ are in each one of three regions, from 41 events plotted in the Fig. 2a and b. For 2-D graphics, filling is done in the x and y directions between the standard deviation of the mean, and the shade rectangular regions are the set of validations of the persistence for each region. Thus, the graph allows a conjugate analysis of persistence. Figure 4a shows in the $\langle\beta_{(j)}\rangle_T$ axis that the MC is the region with the largest average value. However, shade-rectangular regions overlap. It is not possible to separate the MC region. Nevertheless, the result is important, because we can see that persistence is large in the MCs. On the other hand, if we see the $\langle\alpha_{(j)}\rangle_T$ axis then 75 % of the shade-rectangular regions do not overlap. The MCs have $\langle\alpha_{(j)}\rangle$ values from 1.39 to 1.54. A vertical dashed line is drawn in the point 1.392. We propose the use of this value as a threshold when the alpha exponent is calculated in MC regions.

The alpha value characterizes a multiscale phenomenon that can be observed from the fluctuations of the amplitude of the IMF. The coherent structures associated with magnetic clouds are related to scales of hours. However, there are several components of fine structures, which we call noise components (on the order of seconds). These disturbances may be caused by different processes (e.g., Alfvén waves

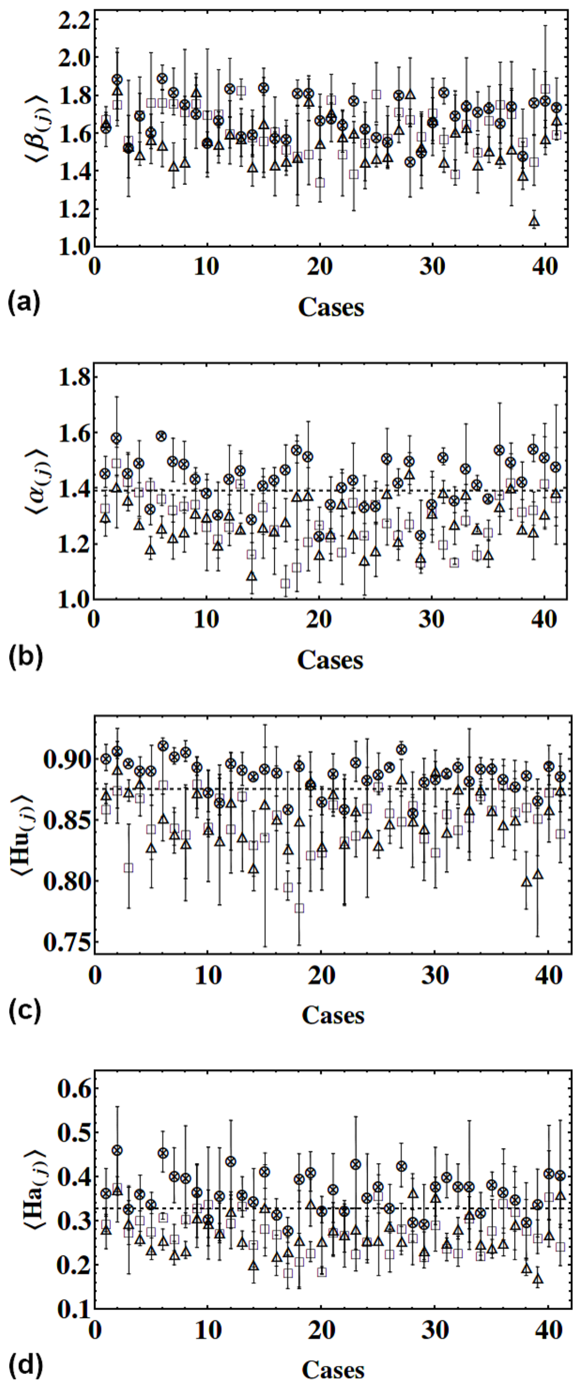

Figure 2. In **(a)**, the PSD scaling exponent $\langle\beta_{(j)}\rangle$ values vs. number of events (see Table 1) plotted, where ("□"), ("⊗"), and ("△") symbols corresponds to the sheath, MCs, and post-MC regions, respectively. The other **(b)**, **(c)**, and **(d)** are similar to **(a)** but for $\langle\alpha_{(j)}\rangle$, $\langle Hu_{(j)}\rangle$, and $\langle Ha_{(j)}\rangle$ exponents, respectively. The results in the four panels show long-range persistence in IMF time series $(1<\langle\beta_{(j)}\rangle<2,\ 1<\langle\alpha_{(j)}\rangle<1.6,\ 0.75<\langle Hu_{(j)}\rangle<0.95$, and $0.1<\langle Ha\rangle<0.5)$. The horizontal dashed line is a threshold derived from Fig. 1.

interacting with the cloud). Another possible nonlinear component at small scales is the fact that there are disturbances outside the coherence B_x and B_z plane (see e.g., Fig. 3, Bothmer and Schwenn, 1998). Here, the calculation of the exponent alpha is taken as the average of the alpha values of each component (B_x, B_y, and B_z). Therefore, the threshold values represent the average complexity signature of the maximum fluctuation of the system. The fluctuation is not self-similar; it is a self-affine phenomenon. This means that there are similar patterns of fluctuation, but only in some scales, not all. An analysis of multi-resolution (for example, by using wavelets) may be important for future work to investigate this process. In the classification of persistence processes, the value of alpha, in the range 1.39 to 1.54, only indicates that, in the transition region, the variability pattern is typically a nonstationary process very close to a Brownian-like fluctuation (≈ 1.5). However, more important than characterizing the process in this context, the detection of the transition should be addressed as the most important issue.

Figure 4b has the same format as Fig. 4a, but, in the y axis, $\langle Ha_{(j)}\rangle_T$ was plotted. Along the $\langle Ha_{(j)}\rangle_T$ axis, the shaded-rectangular region corresponding to the MC has less overlap with other regions than those that are seen in Fig. 4a. Only the MCs have $\langle Ha_{(j)}\rangle$ values between 0.320 and 0.420. A horizontal dashed line is drawn in the point 0.327.

Figure 4c has the same format as Fig. 4a and b, but, in the y axis, $\langle Hu_{(j)}\rangle_T$ was plotted. In addition, the MCs were separated from the other two regions, and the horizontal dashed line is drawn in the point 0.875. The regions with least overlap correspond to the Hurst and Hausdorff exponents, respectively. In Fig. 4d, $(\langle Hu_{(j)}\rangle_T \pm \sigma)$ vs. $(\langle Ha_{(j)}\rangle_T \pm \sigma)$ is plotted. The Hurst and Hausdorff exponents provide good results, and the clouds are separated from the other two regions. This graphic could be used to evaluate the quality when a new MC is identified using other methods; i.e., they are useful to categorize the ranges previously identified by another method.

With these results, we conclude that the persistence values increase in the IMF components inside of MCs. The persistence analysis is not able to distinguish physical differences between sheath and post-MC regions, but the same average values suggest that both regions may be influenced by the noise component (nonlinear processes at finer scales involved in the dynamics of the IMF).

In this study, the investigated period covers the rising phase of solar activity (1998–1999), solar maximum (2000), and the early declining phase (2001–2003) when defined by the yearly sunspot number. We had a variety of MCs in 5 years (1998–2003), and the rotation of the magnetic field direction can occur in any direction relative to the ecliptic. However, there are some MCs for which identification is not completely secure. For example, WIND MC table[2] and Lepping's list show a quality factor ($1 \equiv$ excellent, $2 \equiv$ good, $3 \equiv$ poor) when MC intervals are identified. This

[2]http://wind.nasa.gov/mfi/mag_cloud_pub1.html

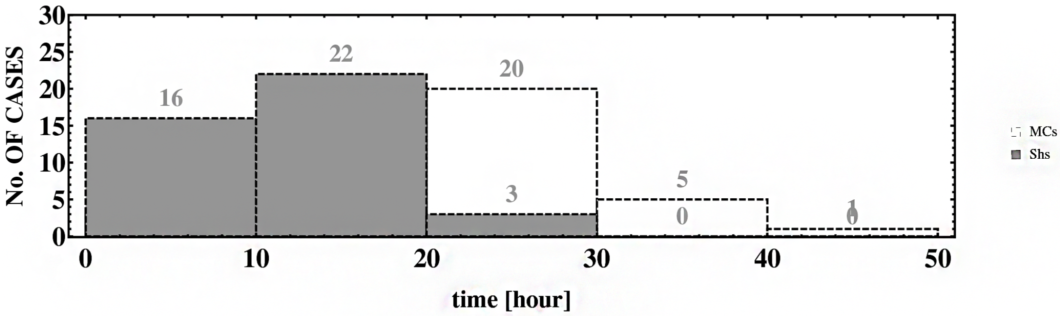

Figure 3. Histogram of 41 MCs and their respective plasma sheaths that are studied in this paper. The histogram shows the number of cases vs. temporal extension (in hours) of MCs and plasma sheaths, respectively.

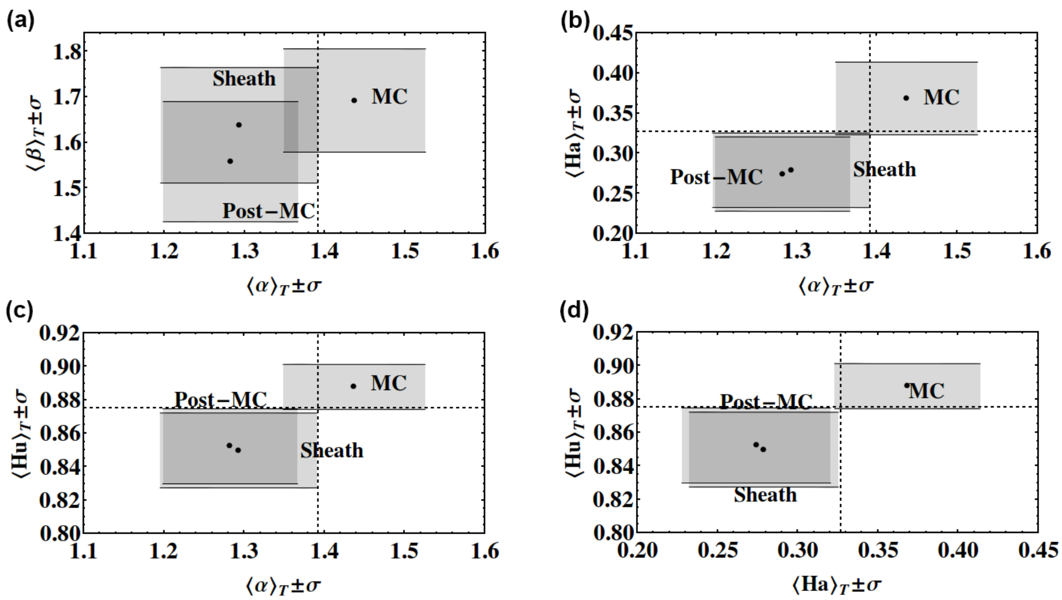

Figure 4. In **(a)**, the black points ($\langle \alpha_{(j)} \rangle_T$, $\langle \beta_{(j)} \rangle_T$) are in each of the three regions of the 41 events plotted in Fig. 2a and b. We calculated the standard deviation of the mean for each persistence exponent that is shown in Eq. (11). For 2-D graphics, filling is done in the x and y directions between the standard deviation of the mean. The filling rectangular regions are the set of validations of the persistence for each regions: ($\langle \beta_{(j)} \rangle_T \pm \sigma$) vs. ($\langle \alpha_{(j)} \rangle_T \pm \sigma$). **(b)**, **(c)**, and **(d)** are similar to **(a)**, except for other exponents combinations; i.e., **(b)** ($\langle Ha_{(j)} \rangle_T \pm \sigma$) vs. ($\langle \alpha_{(j)} \rangle_T \pm \sigma$); **(c)** ($\langle Hu_{(j)} \rangle_T \pm \sigma$) vs. ($\langle \alpha_{(j)} \rangle_T \pm \sigma$); **(d)** ($\langle Hu_{(j)} \rangle_T \pm \sigma$) vs. ($\langle Ha_{(j)} \rangle_T \pm \sigma$).

methodology can help to evaluate the quality of the identification. After identifying an MC, if their persistence exponents occupy non-overlapping regions in Fig. 4b, c, and d, then the cloud was identified with good quality. An advantage of the proposed methodology is that plasma data are not required. The plasma data sometimes have large gaps and poor time resolution when compared with the magnetic field data.

In Table 3, we check whether the 41 events are all in Lepping's list. The first two columns are the same as those published in Lepping's list (MC code and quality factor). Seven events are not published in Lepping's list. These events are shown with the symbol "–". They are the events 5, 10, 16, 17, 20, 27, and 28, as shown in the third column. Lepping's quality factor informs us how well their model identifies each MC. The quality factor is published in a range of 1 to 3. We used the previous idea to create a quality factor that can help to evaluate the quality of the identification; i.e., $Q_p = 1 \equiv$ excellent (three exponents are larger than threshold values), $Q_p = 2 \equiv$ good (two exponents are larger than threshold values), $Q_p = 3 \equiv$ poor (only one exponent is larger than the threshold value), and $Q_p = 0 \equiv$ ill-defined (three exponents are lower than the threshold values, and the field shows little evidence of MCs). The numbers that are greater than the threshold ($\langle \alpha \rangle > 1.392$; $\langle Ha \rangle > 0.327$; $\langle Hu \rangle > 0.875$) are shown in bold font. We found 83 % (34 of 41×100 %) of MCs with quality factor $Q_p = 1$ or $Q_p = 2$. The previous result is better than the 70.6 % (24 of

34 × 100 %) reported in Lepping's list. Of 24 cases reported by Lepping with $Q = 1$ or $Q = 2$, only one disagrees with our results. However, some conflicting results could be expected, because Lepping used a different data set to identify MCs. Seven cases were not reported by Lepping, and $Q_p = 0$ was found in two of them. In Table 3 in the last four columns, there is a summary of the results derived from Fig. 4b, c, and d.

6 Conclusions

The physical bases for the use of the techniques are the plasma features related to the MC processes. Physical–mathematical techniques have been selected for their ability to allow the investigation of MC occurrences. Those techniques have been developed in an original approach to characterize MC events in the SW. They consist of techniques of persistence exponents: Hurst, Hausdorff, the beta exponent from power-spectral density (Fourier), and the alpha exponent from detrended fluctuation analysis, respectively. Those numerical tools have a great advantage, because they are easy to implement with low computational cost and could be the creation of an automatic operation detection. In addition, they characterize MC regions using (as input data) only the three components of the IMF measured by satellites at convenient space location, e.g., the Lagrangian point L1.

We mainly worked with data of B_x, B_y, and B_z with a temporal resolution of 16 s measured by the ACE. We worked with a total of 41 MCs between the years 1998 and 2003, published in the paper of Huttunen et al. (2005). The criteria used to select these 41 cases were the existence of a plasma sheath in front of the MC, and, in these cases, clouds were well identified. We have studied persistence in the 41 ICMEs divided into three regions: plasma sheath, MCs, and post-MCs, respectively. The persistence exponent values increased inside cloud regions, and it was possible to select the following threshold values: $\langle\alpha_{(j)}\rangle = 1.392$; $\langle Ha_{(j)}\rangle = 0.327$; $\langle Hu_{(j)}\rangle = 0.875$. These values prove useful as another test to evaluate the quality of the identification. After identifying a cloud, persistence analysis can be performed in the full extent of temporal series of the three IMF components. If the cloud is well structured, then the persistence exponent values exceed thresholds.

The PSD scaling exponent is not a suitable tool to study persistence in IMF components in the SW. Nevertheless, the other three exponents are suitable for studying persistence, and the exponent values have an increase in the cloud region. It means that the three exponents report the largest persistence in 33 of a total of 41 cloud regions. In 80.5 % of the cases studied, these tools were able to separate the region of the cloud from neighboring regions. The Hausdorff exponent (*Ha*) provides the best results.

One difficulty in studying the persistence in time series is the dimension of it. However, we can see a pattern in

Table 3. The first two columns are the same as those that were published in Lepping's list. MCs that were not identified in Lepping's list are shown with "–". The 41 events in Table 1 are shown in the third column. The last four columns from the left to the right give the Hurst exponent, the Hausdorff exponent, the alpha exponent, and the quality of the MCs, respectively. In columns 4, 5 and 6, the numbers that are greater than the threshold $\langle\alpha\rangle > 1.392$; $\langle H_a\rangle > 0.327$; $\langle H_u\rangle > 0.875$) are shown in bold font.

Code	$Q.$[a]	Table 1	$\langle Hu(j)\rangle$	$\langle Ha(j)\rangle$	$\langle\alpha_{(j)}\rangle$	$Q._p$[b]
28	1	01	**0.901**	**0.365**	**1.460**	1
30	2	02	**0.907**	**0.463**	**1.587**	1
31	1	03	**0.897**	**0.329**	**1.457**	1
32	3	04	**0.891**	**0.363**	**1.496**	1
–	–	05	**0.891**	**0.341**	1.330	2
35	1	06	**0.912**	**0.457**	**1.593**	1
36	2	07	**0.903**	**0.404**	**1.503**	1
37	3	08	**0.907**	**0.400**	**1.493**	1
38	1	09	**0.894**	**0.369**	**1.437**	1
–	–	10	0.874	0.307	1.388	0
39	3	11	0.866	**0.358**	1.310	3
40	3	12	**0.898**	**0.439**	**1.440**	1
41	1	13	**0.892**	**0.362**	**1.470**	1
43	3	14	**0.886**	**0.347**	1.293	2
44.1	3	15	**0.893**	**0.414**	**1.413**	1
–	–	16	**0.890**	0.316	**1.435**	2
–	–	17	0.860	0.280	**1.474**	3
46	2	18	**0.895**	**0.398**	**1.542**	1
47	2	19	**0.879**	**0.412**	**1.521**	1
–	–	20	0.866	0.326	1.234	0
49	2	21	**0.889**	**0.375**	1.349	2
50	3	22	0.860	0.326	**1.408**	3
51	1	23	**0.898**	**0.432**	**1.437**	1
52	2	24	**0.884**	**0.355**	1.337	2
53	3	25	**0.888**	**0.380**	1.340	2
54	2	26	**0.894**	**0.332**	**1.514**	1
–	–	27	**0.909**	**0.427**	**1.423**	1
–	–	28	0.857	0.299	**1.502**	3
57	2	29	**0.882**	0.296	1.235	3
58	2	30	**0.884**	**0.380**	1.348	2
59	2	31	**0.889**	**0.402**	**1.516**	1
60	1	32	**0.895**	**0.381**	1.360	2
62	3	33	**0.883**	**0.382**	**1.477**	1
65	2	34	**0.892**	0.321	**1.419**	2
66	1	35	**0.893**	**0.384**	1.366	2
68	1	36	**0.885**	**0.368**	**1.542**	1
71	2	37	**0.878**	**0.350**	**1.498**	1
72.2	3	38	**0.887**	0.299	**1.428**	2
73	1	39	0.867	**0.341**	**1.545**	2
76	2	40	**0.895**	**0.411**	**1.517**	1
77	2	41	**0.887**	**0.407**	**1.483**	1

[a] From Lepping's list – quality: 1 ≡ excellent, 2 ≡ good, 3 ≡ poor. [b] Our results – quality: 1 ≡ excellent (three exponents are larger than threshold values), 2 ≡ good (two exponents are larger than threshold values), 3 ≡ poor (only one exponent is larger than the threshold value), 0 ≡ ill-defined, the field shows little evidence of MCs (three exponents are lower than the threshold values).

the persistence values between all MC events. An additional analysis by other techniques that consider processes with non-Gaussian features and multifractality is underway and will be presented later (Campos-Velho et al., 2001; Bolzan, M. J. A. et al., 2002).

Fluctuations in time series can also be studied from techniques based on bilateral asymmetries that can be found in the gradient domain of the data. The technique known as gradient pattern analysis (GPA), originally formulated to analyze spatiotemporal data (Rosa et al., 1999), was adapted to analyze patterns of asymmetries that appear exclusively in the time domain (Assireu et al., 2002). The GPA for time series (known as "GPA-1D") compares amplitude values considering different scales of time fluctuation mapped in its gradient field (Rosa et al., 2008). Within the scope of the GPA-1D, the value of the gradient asymmetry coefficient can also present relations with the values obtained from DFA, power spectra and fractal measures. Therefore, the use of gradient pattern analysis (GPA-1D) (Assireu et al., 2002; Rosa et al., 2008) will be explored further in an complementary work.

Appendix A: Autocorrelations and semivariograms

A summary taken from Malamud and Turcotte (1999) is presented here. The correlation of a time series with itself, i.e., $y(t+s)$ compared with $y(t)$ at lag s, is called autocorrelation function ($r(s)$). The autocorrelation function can be used to quantify the persistence or anti-persistence of a time series. This is given by

$$r(s) = \frac{c(s)}{c(0)}, \tag{A1}$$

with the autocovariance function, $c(s)$, given by

$$c(s) = \frac{1}{(T'-s)} \int_0^{T'-s} \left[y(t+s) - \overline{y}\right]\left[y(t) - \overline{y}\right] \mathrm{d}t,$$

and the autocovariance function at 0 lag, $c(0)$, given by

$$c(0) = \frac{1}{T'} \int_0^{T'} \left[y(t) - \overline{y}\right]^2 \mathrm{d}t = V_a.$$

The time series, $y(t)$, is prescribed over the interval $0 \le t \le T'$. The average and variance of $y(t)$ over the interval T' are $\overline{y}$ and V_a. The autocorrelation function, $r(s)$, is dimensionless and does not depend on the units of $y(t)$ or t. The plot of $r(s)$ vs. s is known as correlogram (Malamud and Turcotte, 1999).

For a discrete time series, the autocorrelation function, r_k, is given by

$$r_k = \frac{c_k}{c_0}, \tag{A2}$$

with the autocovariance, c_k, given by

$$c_k = \frac{1}{(N-k)} \sum_{n=1}^{N-k} (y_{n+k} - \overline{y})(y_n - \overline{y}), \tag{A3}$$

and the autocovariance at 0 lag (the variance) given by

$$c_0 = \frac{1}{N} \sum_{n=1}^{N} (y_n - \overline{y})^2 = V_a. \tag{A4}$$

If the mean or variance vary with the length of the interval considered, then the time series is nonstationary. The correlograms is inappropriate to study nonstationary time series, because $r(s)$ has $\overline{y}$ in its definition. However, the method to measure long-range correlation, which is valid for both stationary and nonstationary time series, is the semivariogram γ. Like the autocorrelation function, the semivariogram measures the dependence of values in time series that are separated by lag, s.

For a discrete time series, the semivariogram, $\gamma(s)$, is given by

$$\gamma_k = \frac{1}{2}(N-k) \sum_{n=1}^{N-k} (y_{n+k} - y_n)^2. \tag{A5}$$

For a stationary time series, the semivariogram, γ_k, and the autocorrelation function, r_k, are related. The mean of the time series, $\overline{y}$, can be added and subtracted within the summation in Eq. (A5) to give

$$\gamma_k = \frac{1}{2(N-k)} \sum_{n=1}^{N-k} \left[(y_{n+k} - \overline{y}) - (y_n - \overline{y})\right]^2.$$

When expanded, this gives

$$\gamma_k = \frac{1}{2(N-k)} \left[\sum_{n=1}^{N-k} (y_{n+k} - \overline{y})^2 + \sum_{n=1}^{N-k} (y_n - \overline{y})^2 - \sum_{n=1}^{N-k} 2(y_{n+k} - \overline{y})(y_n - \overline{y})\right]. \tag{A6}$$

Provided the time series is stationary, two of the terms in Eq. (A6) are equivalent to the variance in Eq. (A4), giving

$$\gamma_k = V_a - \frac{1}{(N-k)} \sum_{n=1}^{N-k} (y_{n+k} - \overline{y})(y_n - \overline{y}). \tag{A7}$$

Substituting the definition for c_k from Eq. (A3) into Eq. (A7) and using the definitions of c_0 from Eq. (A4) and r_k from Eq. (A2), the new equation is

$$\gamma_k = (V_a - c_k) = \left(V - V\frac{c_k}{c_0}\right) V(1 - r_k). \tag{A8}$$

For an uncorrelated time series, we have $r_k = 0$ and $\gamma_k = V_a$. Several authors have applied both the autocorrelation function and semivariograms to both real and synthetic time series that exhibit long-range persistence (e.g., Ramos et al., 2004; Rosa et al., 2008).

Using the definition for the semivariogram, γ_k, given in Eq. (A5), a computational code was implemented.

```
function [Ha, R1] = Semivariogram(y)
N1 = size(y, 1);
potencia2 = floor(log2(N1));
gammaT_k = 1 : potencia2;
xi = 1 : potencia2;
for i = 1 : potencia2
k = 2^i;
contador = 0;
for n = 1 : (N1 − k)
contador = contador + (y(n + k) − y(n))^2;
end
gam_k = (1/(N1 − k))*contador;
gammaT_k(i) = gam_k;
```

```
xi(i) = k;
end
yi = gammaT_k;
[a, R] = RegresionLinear(log10(xi), log10(yi));
Ha = a/2;
R1 = R;
end
```

```
function [a, R] = RegresionLinear(xi, yi)
n1 = size(xi, 2);
a = (n1*sum(xi.*yi)-sum(xi)*sum(yi))/(n1*sum(xi.^2)-
sum(xi)^2);
b = (sum(yi) - a*sum(xi))/n1;
R = ((sum(xi.*yi)-(sum(xi)*sum(yi))/n1)^2)/((sum(xi.^2)-
(sum(xi)^2)/n1)*(sum(yi.^2) - (sum(yi)^2)/n1));
end
```

Acknowledgements. This work was supported by grants from CNPq (grant 483226/2011-4), FAPESP (grant 2012/072812-2) and CAPES (grant 1236-83/2012). A. Ojeda González thanks CAPES and CNPq (grant 141549/2010-6) for his PhD scholarship and CNPq (grants 150595/2013-1, 503790/2012-5) for his postdoctoral research support. We also wish to thank the anonymous referees.

Edited by: A. S. Sharma
Reviewed by: two anonymous referees

References

Assireu, A. T., Rosa, R. R., Vijaykumar, N. L., Lorenzzetti, J. A., Rempel, E. L., Ramos, F. M., Sá, L. D. A., Bolzan, M. J. A., and Zanandrea, A.: Gradient patter analysis of short nonstationary time series: an application to Lagrangian data from satellite tracked drifters, Physica D, 168–169, 397–403, doi:10.1016/S0167-2789(02)00527-4, 2002.

Baroni, M. P. M. A., Wit, A. D., and Rosa, R. R.: Detrended fluctuation analysis of numerical density and viscous fingering patterns, Europhys. Lett., 92, 64002, doi:10.1209/0295-5075/92/64002, 2010.

Bashan, A., Bartsch, R., Kantelhardt, J. W., and Havlin, S.: Comparison of detrending methods for fluctuation analysis, Physica A, 387, 5080–5090, doi:10.1016/j.physa.2008.04.023, 2008.

Bolzan, M. J. A., Ramos, F. M., Sá, L. D. A., Neto, C. R., and Rosa, R. R.: Analysis of fine-scale canopy turbulence within and above an Amazon forest using Tsallis' generalized thermostatistics, J. Geophys. Res., 107, 8063, doi:10.1029/2001JD000378, 2002.

Bothmer, V. and Schwenn, R.: The structure and origin of magnetic clouds in the solar wind, Ann. Geophys., 16, 1–24, doi:10.1007/s00585-997-0001-x, 1998.

Bryce R. M. and Sprague K. B.: Revisiting detrended fluctuation analysis, Sci. Rep., 2, 1–6, 2012.

Buldyrev, S. V., Goldberger, A. L., Havlin, S., Mantegna, R. N., Matsa, M. E., Peng, C.-K., Simons, M., and Stanley, H. E.: Long-range correlation properties of coding and noncoding DNA sequences: genbank analysis, Phys. Rev. E, 51, 5084–5091, 1995.

Burlaga, L. F.: Magnetic clouds, in: Physics of the Inner Heliosphere, Vol. 2, edited by: Schwenn, R. and Marsch, E., Springer-Verlag, New York, 1–2, 1991.

Burlaga, L. F., Sittler, E., Mariani, F., and Schwenn, R.: Magnetic loop behind an interplanetary shock: Voyager, Helios and IMP 8 observations, J. Geophys. Res., 86, 6673–6684, 1981.

Calzadilla, M. A. and Lazo, B.: Sheffield Space Plasma Meeting (2001) multipoint measurements vs. theory, in: Inproceedings, ESA: Non-lineal time series analysis of Dst geomagnetic index, ESA Publications Division, SP-492, 24–26 April 2001, edited by: Warmbein, B., Sheffield, UK, 121–125, 2001.

Campos-Velho, H. F., Rosa, R. R., Ramos, F. M., Pielke, R. A., Degrazia, G. A., Neto, C. R., and Zanandrea, A.: Multifractal model for eddy diffusivity and counter-gradient term in atmospheric turbulence, Physica A, 295, 219–223, 2001.

Dasso, S., Mandrini, C., Démoulin, P., Luoni, M., and Gulisano, A.: Large scale MHD properties of interplanetary magnetic clouds, Adv. Space Res., 35, 711–724, 2005.

Heneghan, C. and McDarby, G.: Establishing the relation between detrended fluctuation analysis and power spectral density analysis for stochastic processes, Phys. Rev. E, 62, 6103–6110, 2000.

Hu, K., Ivanov, P. C., Chen, Z., Carpena, P., and Stanley, H. E.: Effect of trends on detrended fluctuation analysis, Phys. Rev. E, 64, 011114, doi:10.1103/PhysRevE.64.011114, 2001.

Hurst, H. E.: Long-term storage capacity of reservoirs, T. Am. Soc. Civ. Eng., 116, 770–808, 1951.

Hurst, H. E., Black, R. P., and Simaika, Y. M.: Long-Term Storage: an Experimental Study, Constable, London, 145 pp., 1965.

Huttunen, K. E. J., Schwenn, R., Bothmer, V., and Koskinen, H. E. J.: Properties and geoeffectiveness of magnetic clouds in the rising, maximum and early declining phases of solar cycle 23, Ann. Geophys., 23, 625–641, doi:10.5194/angeo-23-625-2005, 2005.

Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E.: Multifractal detrended fluctuation analysis of nonstationary time series, Physica A, 316, 87–114, 2002.

Kay, S. and Marple, S. J.: Spectrum analysis – a modern perspective, Proc. IEEE, 69, 1380–1419, doi:10.1109/PROC.1981.12184, 1981.

Klausner, V., Ojeda, G. A., Domingues, M. O., Mendes, O., and Papa, A. R. R.: Study of local regularities in solar wind data and ground magnetograms, J. Atmos. Sol.-Terr. Phy., 112, 10–19, doi:10.1016/j.jastp.2014.01.013, 2014.

Klein, L. W. and Burlaga, L. F.: Interplanetary magnetic clouds at 1 AU, J. Geophys. Res., 87, 613–624, 1982.

Lepping, R. P., Acuña, M. H., Burlaga, L. F., Farrell, W. M., Slavin, J. A., Schatten, K. H., Mariani, F., Ness, N. F., Neubauer, F. M., Whang, Y. C., Byrnes, J. B., Kennon, R. S., Panetta, P. V., Scheifele, J., and Worley, E. M.: The WIND magnetic field investigation, Space Sci. Rev., 71, 207–229, 1995.

Little, M., McSharry, P., Moroz, I., and Roberts, S.: Nonlinear, biophysically-informed speech pathology detection, in: ICASSP 2006 Proceedings, 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 14–19 May 2006, Toulouse, France, 1080–1083, doi:10.1109/ICASSP.2006.1660534, 2006.

Malamud, B. D. and Turcotte, D. L.: Self-affine times series: I. Generation and analyses, Adv. Geophys., 40, 1–90, doi:10.1016/S0065-2687(08)60293-9, 1999.

Mandelbrot, B. B.: The Fractal Geometry of Nature, W. H. Freeman and Company, New York, 480 pp., 1983.

Mandelbrot, B. B. and Ness, J. W. V.: Fractional brownian motions, fractional noises and applications, SIAM Rev., 10, 422–437, 1968.

Mandelbrot, B. B. and Wallis, J. R.: Some long-run properties of geophysical records, Water Resour. Res., 5, 321–340, 1969.

Mcsharry, P. E. and Malamud, B. D.: Quantifying self-similarity in cardiac inter-beat interval time series, in: Computers in Cardiology, IEEE, Lyon, 459–462, doi:10.1109/CIC.2005.1588136, 2005.

Narock, T. W. and Lepping, R. P.: Anisotropy of magnetic field fluctuations in an average interplanetary magnetic cloud at 1 AU, J. Geophys. Res., 112, A06108, doi:10.1029/2006JA011987, 2007.

Ojeda, G. A., Calzadilla, M. A., Lazo, B., Alazo, K., and Savio, S.: Analysis of behavior of solar wind parameters under different IMF conditions using two nonlinear dynamics techniques, J. Atmos. Sol.-Terr. Phy., 67, 1859–1864, doi:10.1016/j.jastp.2004.12.014, 2005.

Ojeda, G. A., Mendes, O., Calzadilla, M. A., and Domingues, M. O.: Spatio-temporal entropy analysis of the magnetic field to help magnetic cloud characterization, J. Geophys. Res., 118, 5403–5414, 2013.

Ojeda, G., A., Mendes, O., Domingues, M. O., and Menconi, V. E.: Daubechies wavelet coefficients: a tool to study interplanetary magnetic fluctuations, Geofis. Int., 53, 101–115, 2014.

Pelletier, J. D. and Turcotte, D. L.: Self-affine times series: 2. Applications and models, Adv. Geophys., 40, 91–166, 1999.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L.: Mosaic organization of DNA nucleotides, Phys. Rev. E, 49, 1685–1689, 1994.

Percival, D. B. and Walden, A. T.: Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques, Cambridge University Press, Cambridge, New York, NY, USA, p. 583, 1993.

Price, C. P. and Newman, D. E.: Using the R/S statistic to analyze AE data, J. Atmos. Sol.-Terr. Phy., 63, 1387–1397, 2001.

Priestley, M. B.: Spectral Analysis and Time Series, Academic Press, London, New York, 890 pp., 1981.

Ramos, F. M., Bolzan, M. J. A., Sá, L. D. A., and Rosa, R. R.: Atmospheric turbulence within and above an Amazon forest, Physica D, 193, 278–291, doi:10.1016/j.physd.2004.01.026, 2004.

Rosa, R. R., Sharma, A. S., Valdivia, J. A.: Characterization of asymmetric fragmentation patterns in spatially extended systems, Int. J. Mod. Phys. C., 10, 509–514, 1999.

Rosa, R. R., Karlický, M., Veronese, T. B., Vijaykumar, N. L., Sawant, H. S., Borgazzi, A. I., Dantas, M. S., Barbosa, E. B. M., Sych, R. A., and Mendes, O.: Gradient pattern analysis of short solar radio bursts, Adv. Space Res., 42, 844–851, 2008.

Ruzmaikin, A., Feynman, J., and Robinson, P.: Long-term persistence of solar activity, Sol. Phys., 149, 395–403, 1994.

Sharma, A. S. and Veeramani, T.: Extreme events and long-range correlations in space weather, Nonlin. Processes Geophys., 18, 719–725, doi:10.5194/npg-18-719-2011, 2011.

Tapiero, C. and Vallois, P.: Run length statistics and the Hurst exponent in random and birth-death random walks, Chaos Soliton. Fract., 7, 1333–1341, doi:10.1016/0960-0779(96)00032-X, 1996.

Tsurutani, B. T., Sugiura, M., Iyemori, T., Goldstein, B. E., Gonzalez, W. D., Akasofu, S. I., and Smith, E. J.: The nonlinear response of AE to the IMF Bs driver: A spectral break at 5 hours, Geophys. Res. Lett., 17, 279–282, 1990.

Turcotte, D. L.: Fractals and chaos in geology and geophysics, 2nd revised Edn., Cambridge University Press, New York, USA, 1997.

Vanouplines, P.: Rescaled range analysis and the fractal dimension of pi, University Library, Free University Brussels, Brussels, Belgium, 1995.

Veronese, T., Rosa, R., Bolzan, M., Fernandes, F. R., Sawant, H., and Karlicky, M.: Fluctuation analysis of solar radio bursts associated with geoeffective X-class flares, J. Atmos. Sol.-Terr. Phy., 73, 1311–1316, doi:10.1016/j.jastp.2010.09.030, 2011.

Voss, R. F.: Random fractals: characterization and measurement, in: Scaling Phenomena in Disordered System, edited by: Pynn, R. and Skjeltorp, A., Springer, USA, 1–11, doi:10.1007/978-1-4757-1402-9_1, 1985a.

Voss, R. F.: Random fractal forgeries, in: Fundamental Algorithms for Computer Graphics, NATO ASI, F17, edited by: Earnshaw, R. A., Springer-Verlag, Berlin, Heidelberg, 805–835, 1985b.

Voss, R. F.: Characterization and measurement of random fractals, Phys. Scripta, 13, 27–32, 1986.

Wei, H. L., Billings, S. A., and Balikhin, M.: Analysis of the geomagnetic activity of the D_{st} index and self–affine fractals using wavelet transforms, Nonlin. Processes Geophys., 11, 303–312, doi:10.5194/npg-11-303-2004, 2004.

Zheng, H., Song, W., and Wang, J.: Detrended fluctuation analysis of forest fires and related weather parameters, Physica A, 387, 2091–2099, 2008.