# Low-dimensionality and predictability of solar wind and global magnetosphere during magnetic storms

T. Živković, and K. Rypdal,

Abstract. The storm indices  $D_{st}$  and SYM-H and the solar wind velocity v and interplanetary magnetic field  $B_z$  show no signatures of low-dimensional dynamics in quiet periods, but tests for determinism in the time series indicate that SYM-H exhibits a significant low-dimensional component during storm time, suggesting that self-organization takes place during magnetic storms. Even though our analysis yields no discernible change in determinism during magnetic storms for the solar wind parameters, there are significant enhancement of the predictability and exponents measuring persistence. Thus, magnetic storms are typically preceded by an increase in the persistence of the solar wind dynamics, and this increase is reflected in the magnetospheric response to the solar wind.

## 1. Introduction

Under the influence of the solar wind, the magnetosphere resides in a complex, non-equilibrium state. The plasma particles have non-Maxwellian velocity distribution, MHD turbulence is present everywhere, and intermittent energy transport known as bursty-bulk flows occurs as well [Angelopoulos et al., 1999]. The magnetospheric response to particular solar events constitutes an essential aspect of space weather while the response to solar variability in general is often referred to as space climate [Watkins, 2002]. Theoretical approaches to space climate involve concepts and methods from stochastic processes, nonlinear dynamics and chaos, turbulence, self-organized criticality, and phase transitions.

Self-organization can lead to low-dimensional behavior in the magnetosphere [Klimas et al., 1996; Vassiliadis et al., 1990; Sharma et al., 1993]. However, power-law dependence observed in the Fourier spectra of the auroral electrojet (AE) index is a typical signature of high dimensional colored noise indicating multi-scale dynamics of the magnetosphere. In order to reconcile low-dimensional, deterministic behavior with high-dimensionality, Chang [1998] proposed that a high-dimensional system near self-organized criticality (SOC) [Bak et al., 1987] can be characterized by a few parameters whose evolution is governed by a small number of nonlinear equations. Some magnetospheric models, like the one presented in Chapman et al. [1998], are based on the SOC-concept. Here a system tunes itself to criticality and the energy transport across scales is mediated by avalanches which are power-law distributed in size and duration.

On the other hand, it was suggested in Sitnov et al. [2001] that the substorm dynamics can be described as a non-equilibrium phase transition; i.e. as a system tuned externally to criticality. Here, a power-law relation is given, with characteristic exponent close to the input-output critical exponent in a second-order phase transition. In fact, it is claimed in Sharma et al. [2003] that the global features of the magnetosphere correspond to a first order phase transition whereas multi-scale processes correspond to the second-order phase transitions.

The existence of metastable states in the magnetosphere, where intermittent signatures might be due to dynamical

 $^1\mathrm{Department}$  of Physics and Technology, University of Tromsø, Norway.

Copyright 2011 by the American Geophysical Union. 0148-0227/11/\$9.00

phase transitions among these states, was suggested by  $Consolini\ and\ Chang\ [2001]$ , and forced and/or self-organized criticality (FSOC) induced by the solar wind was introduced as a conceptual description of magnetospheric dynamics. The concept of intermittent criticality was suggested by  $Balasis\ et\ al.\ [2006]$  who asserted that during intense magnetic storms the system develops long-range correlations, which further indicates a transition from a less orderly to a more orderly state. Here, substorms might be the agents by which longer correlations are established. This concept implies a time-dependent variation in the activity as the critical point is approached, in contrast to SOC.

In the present paper we investigate determinism and predictability of observables characterizing the state of the magnetosphere during geomagnetic storms as well as during its quiet condition, but the emphasis is on the evolution of these properties over the course of major magnetic storms. The measure of determinism employed here increases if the system dynamics is dominated by modes governed by low-dimensional dynamics. Hence, the determinism in most cases is a measure of low-dimensionality. For a low-dimensional, chaotic system the predictability measure increases when the largest Lyapunov exponent increases, and hence is a really a measure of un-predictability. For a high-dimensional or stochastic system it is related to the degree of persistence in time series representing the dynamics. High persistence means high predictability.

One of the most useful data tools for probing the magnetosphere during substorm conditions is the Kyoto AE minute index which is defined as the difference between the AU index, which measures the eastward electrojet current in the auroral zone, and the AL index, which measures the westward electrojet current, and is usually derived from 12 magnetometers positioned under the auroral oval [Davies and Sugiura, 1966]. The auroral electrojet, however, does not respond strongly to the specific modifications of the magnetosphere that occur during magnetic storms. A typical storm characteristic, however, is a change in the intensity of the symmetric part of the ring current that encircles Earth at altitudes ranging from about 3 to 8 Earth radii, and is proportional to the total energy in the drifting particles that form this current system [Gonzalez et al., 1994]. For the study of storm dynamics the  $D_{st}$  and SYM-H indices are both designed to measure the intensity of the storm-time ring current. They are derived from similar data sources, but SYM-H has the distinct advantage of having 1-min time resolution compared to the 1-hour time resolution of  $D_{st}$ . Wanliss [2006] has recommended that the SYM-H index be used as a de facto high-resolution  $D_{st}$  index. The analysis of these indices are central to this study, where SYM-H is used whenever hourly time resolution proves inadequate.

We also use minute data for the interplanetary magnetic field (IMF) component  $B_z$ , as well as minute data for the solar wind bulk velocity v along the Sun-Earth axis. These data are both retrieved from the OMNI satellite database and are given in the GSE coordinate system. Gaps of missing data in  $B_z$  and v are excluded from the analysis while SYM-H data are analyzed for the entire period.

The typical magnetic storm consists of the initial phase, when the horizontal magnetic field suddenly increases and stays elevated for several hours, the main phase where this component is depressed for one to several hours, and the recovery phase which also lasts several hours. The initial phase has been associated with northward directed IMF (little energy enters the magnetosphere), but it has been discovered that this phase is not essential for the storm to occur Akasofu [1965]. In order to define a storm, we follow the approach of Loewe and  $Pr\ddot{o}lss$  [1997], where the  $D_{st}$  minimum is a common reference epoch, the main-phase decrease is sufficiently steep, and the recovery phase is also defined.

The remainder of the paper is organized as follows: section 2 describes the data analysis methods employed. Section 3 presents analysis results discerning general statistical scaling properties of global magnetospheric dynamics using minute data over several years and data generated by a numerical model which produces realizations of a fractional Ornstein-Uhlenbebck (fO-U) process. In particular we study how determinism and predictability of the geomagnetic and solar wind observables change over the course of magnetic storms. Section 4 is reserved for discussion of results and section 5 for conclusions.

## 2. Methods

#### 2.1. Recurrence plot analysis

The recurrence plot is a powerful tool for the visualization of recurrences of phase-space trajectories. It is very useful since it can be applied to non-stationary as well as short time series [ $Eckmann\ et\ al.$ , 1987], and this is the nature of data we use to explore magnetic storms. Prior to constructing a recurrence plot the common procedure is to reconstruct phase space from the time-series x(t), which has the length N by time-delay embedding [Takens, 1981].

Suppose the physical system at hand is a deterministic dynamical system describing the evolution of a state vector  $\mathbf{z}(t)$  in a phase space of dimension p, i.e.  $\mathbf{z}$  evolves according to an autonomous system of 1st order ordinary differential equations;

$$\frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z}), \ \mathbf{f}: \mathcal{R}^p \to \mathcal{R}^p$$
 (1)

and that an observed time series x(t) is generated by the measurement function  $g: \mathcal{R}^p \to \mathcal{R}$ ,

$$x(t) = g(\mathbf{z}(t)). \tag{2}$$

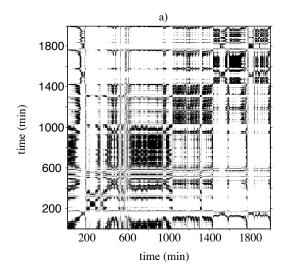
Assume that the dynamics takes place on an invariant set (an attractor)  $\mathcal{A} \subseteq \mathcal{R}^p$  in phase space, and that this set has box-counting fractal dimension d. Since the dynamical system uniquely defines the entire phase-space trajectory once the state  $\mathbf{z}(t)$  at a particular time t is given, we can define uniquely an m-dimensional measurement function,

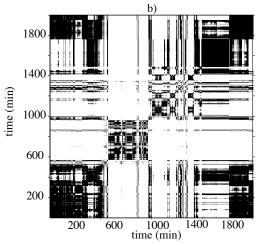
$$\mathbf{g}: \mathcal{A} \to \mathcal{R}^m, \ \mathbf{g}(\mathbf{z}) = (x(t), x(t+\tau), \dots, x(t+(m-1)\tau)).$$
(3)

where the vector components are given by equation (2), and  $\tau$  is a time delay of our choice. If the invariant set  $\mathcal{A}$  is compact (closed and bounded), g is a smooth function and

m>2d, the map given by equation (3) is a topological embedding (a one-to-one continuous map) between  $\mathcal{A}$  and  $\mathcal{R}^m$ . The condition m>2d can be thought of as a condition for the image  $\mathbf{g}(\mathcal{A})$  not to intersect itself, i.e. to avoid that two different states on the attractor  $\mathcal{A}$  are mapped to the same point in the m-dimensional embedding space  $\mathcal{R}^p$ . If such an embedding is achieved, the trajectory  $\mathbf{x}(t) = \mathbf{g}(\mathbf{z})$  (where  $\mathbf{g}(\mathbf{z})$  is given by equation (3)) in the embedding space is a complete mathematical representation of the dynamics on the attractor. Note that the dimension p of the original phase space is irrelevant for the reconstruction of the embedding space. The important thing is the dimension d of the invariant set  $\mathcal{A}$  on which the dynamics unfolds.

There are practical constraints on useful choices of the time delay  $\tau$ . If  $\tau$  is much smaller than the autocorrelation time the image of  $\mathcal{A}$  becomes essentially one-dimensional. If  $\tau$  is much larger than the autocorrelation time, noise may destroy the deterministic connection between the components of  $\mathbf{x}(t)$ , such that our assumption that  $\mathbf{z}(t)$  determines  $\mathbf{x}(t)$  will fail in practice. A common choice of  $\tau$  has been the first minimum of the autocorrelation function, but it has been shown that better results are achieved by selecting the time delay as the first minimum in the average mutual in-





**Figure 1.** Recurrence plot for  $B_z$ . a) during quiet conditions, b) during a magnetic storm.

formation function, which can be percieved as a nonlinear autocorrelation function [Abarbanel, 1996].

The recurrence plot analysis deals with the trajectories in the embedding space. If the original time series x(t) has N elements separated by the time delay  $\tau$ , we have a time series of N-m+1 vectors  $\mathbf{x}(t)$  for  $t=\tau,\ldots,(N-m+1)\tau$ . This time series constitutes the trajectory in the reconstructed embedding space.

The next step is to construct a  $(N-m+1) \times (N-m+1)$  matrix  $R_{i,j}$  consisting of elements 0 and 1. The matrix element (i,j) is 1 if the distance is  $\|\mathbf{x}_i - \mathbf{x}_j\| \leq r$  in the reconstructed space, and otherwise it is 0. The recurrence plot is simply a plot where the points (i,j) for which the corresponding matrix element is 1 is marked by a dot. For a deterministic system the radius r is typically chosen as 10% of the diameter of the reconstructed attractor, but varies for different sets of data. For a non-stationary stochastic process like a Brownian motion there is no bounded attractor for the dynamics, and the diameter is limited by the length of the data record. Examples of recurrence plots are shown in Figure 1, obtained from  $B_z$ , when no storm is present in a), and during a magnetic storm in b).

## 2.2. Empirical mode decomposition

The empirical mode decomposition (EMD) method, developed in *Huang et al.* [1998] is very useful on non-stationary and nonlinear time series. EMD method can give a change of frequency in any moment of time (instantaneous frequency) and a change of amplitude in the system. However, in order to properly define instantaneous frequency, a time series should have the same number of zero crossings and extrema (or they can differ at most by one), and a local mean should be close to zero. The original time series

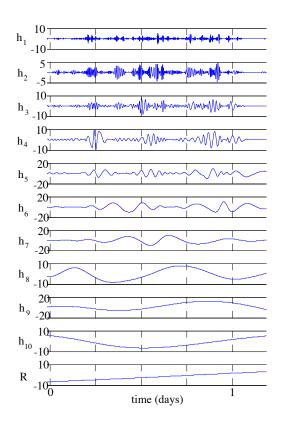
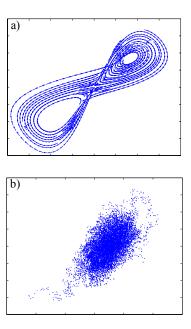


Figure 2. IFs obtained by EMD for  $B_z$  during a magnetic storm

usually does not have these characteristics and should be decomposed into intrinsic mode functions (IF) for which instantaneous frequency can be defined. Decomposition can be obtained through the so-called sifting process. This is an adaptive process derived from the data and can be briefly described as follows: All local maxima and minima in the time series s(t) are found, and all local maxima and minima are fitted by cubic spline and these fits define the upper (lower) envelope of the time series. Then the mean of the upper and lower envelope m(t) is defined, and the difference between the time series and this mean represents the first IF, h(t) = s(t) - m(t), if instantaneous frequency can be obtained, defined by some stopping criterion. If not, the procedure is repeated (now starting from h(t) instead of s(t)) until the first IF is produced. Higher IFs are obtained by subtracting the first IF from the time series s(t) and the entire previously mentioned procedure is repeated until a residual, usually a monotonic function, is left. We use a stopping criterion defined by Rilling et al. [2003], where  $\eta(t) < \theta_1$  on  $1-\gamma$  fraction of the IF, and  $\eta(t) < \theta_2$ , on the remaining fraction of the IF. Here  $\eta = m(t)/a(t)$ , a(t) is the IF amplitude, and  $\gamma = 0.05$ ,  $\theta_1 = 0.05$ , and  $\theta_2 = 0.5$ . By the above definitions, IFs are complete in the sense that their summation gives the original time series:  $s(t) = \sum_1^M h(t) + R(t)$  where M is the number of IFs and R is a residual. In Figure 2 we show the IFs from EMD performed on the IMF  $B_z$  during a magnetic storm.

In order to study stochastic behavior of a time series by means of EMD analysis, we refer to Wu and Huang [2004] who studied characteristics of white noise using the EMD method. They derived for white noise the relationship  $\log E_m = -\log T_m$ , where  $E_m$  and  $T_m$  represents empirical variance and mean period for the m'th IF. Here,  $E_m = (1/N) \sum_{t=1}^N h(t)^2$ , where h(t) is the m'th IF, and  $T_m$  is the ratio of the m'th IF length to the number of its zero crossings. Franzke [2009] analyzed telecommunication indices and noticed a resemblance to autoregressive processes of the first order AR(1), which are stochastic and linear processes. For such processes log  $E_m = \zeta \log T_m$ . For fractional



**Figure 3.** a) Average displacement vectors  $\mathbf{V}_j$  in each box j for a 2-dimensional projection of m-dimensional embedding space. a) Lorenz system for m=3. b) fO-U for m=8.

Gaussian noise processes (H < 1) and fractional Brownian motions (H > 1) we have the connection  $\zeta = 2H - 2$ , where H is the Hurst exponent, as shown by Flandrin and Gonçalvès [2004]. A useful feature of the EMD analysis is the possibility of extraction of trends in the time series [Wu et al., 2007], because the slowest IF components should often be interpreted as trends. This is an advantage compared to the standard variogram or rescaled-range techniques [Beran, 1994], whose estimation of the scaling exponents is biased by the trend.

#### 2.3. A test for determinism

In this paper we employ a simple test for determinism, developed by Kaplan and Glass [1992], where the following hypothesis is tested: When a system is deterministic, the orientation of the trajectory (its tangent) is a function of the position in the phase space. Further, this means that a recurrent trajectories in the same box in phase space, will have parallel directions since these are all functions of the same position in the phase space. On the other hand, trajectories in a stochastic system have directions which do not depend on the position and are equally probable in any direction. This hypothesis can only be tested on continuous flows, and is not applicable to maps since consecutive points on the trajectory may be very separated in the phase space. In this context, a trajectory orientation is defined through a vector of a unit length, whose direction is given by the displacement between the point where trajectory enters the box j to the point where the trajectory exits the same box. The displacement in m-dimensional embedding space is given from the time-delay embedding reconstruction:

$$\Delta \mathbf{x}(t) = [x(t+b) - x(t), x(t+\tau+b) - x(t+\tau), \dots, x(t+(m-1)\tau+b) - x(t+(m-1)\tau)],$$
(4)

where b is the characteristic time the trajectory spends inside a box. The orientation vector for the kth pass through box j is the unit vector  $\mathbf{u}_{k,j} = \Delta \mathbf{x}_{k,j}(t)/|\Delta \mathbf{x}_{k,j}(t)|$ . The estimated averaged displacement vector in the box is

$$\mathbf{V}_{j} = \frac{1}{n_{j}} \sum_{k=1}^{n_{j}} \mathbf{u}_{k,j}, \tag{5}$$

where  $n_j$  is the number of passes of the trajectory through box j. If the embedding dimension is sufficiently high and in the limit of vanishingly small box size, the trajectory directions should be aligned and  $V_j \equiv |\mathbf{V}_j| = 1$ . In the case of deterministic dynamics and finite box size,  $\mathbf{V}_j$  will not depend very much on the number of passes  $n_j$ , and  $V_j$  will converge to 1 as  $n_j \to \infty$ . In contrast, for the trajectory of a random process, where the direction of the next step is completely independent of the past,  $V_j$  will decrease with  $n_i$  as  $V_i \sim n_i^{-1/2}$ .

 $n_j$  as  $V_j \sim n_j^{-1/2}$ . In Figure 3a we show displacement vectors  $\mathbf{V}_j$  averaged over the passes through the box j, for a three-dimensional embedding of the Lorenz attractor, in Figure 3b the same is shown for a random process, in this case a fractional Ornstein-Uhlenbeck (fO-U) process. These model systems will be used throughout this paper as archetypes of low-dimensional and stochastic systems, respectively. The Lorenz system has the form

$$dx/dt = a(y-x)$$

$$dy/dt = -xz + cx - y$$

$$dz/dt = xy - bz,$$
(6)

with standard coefficient values a=10, b=8/3, and c=28, which give rise to a chaotic flow. The fO-U process is described by the stochastic equation:

$$dS_t = \lambda(\mu - S_t) + \sigma dW_t, \tag{7}$$

where  $dW_t$  is a fractional Gaussian noise with Hurst exponent H [Beran, 1994]. The drift ( $\lambda$  and  $\mu$ ) and diffusion ( $\sigma$ ) parameters are fitted by the least square regression to the time series of the SYM-H storm index. This will be explained in more detail in section 3.2.

The degree of determinism of the dynamics can be assessed by exploring the dependence of  $V_j$  on  $n_j$ . In practice, this can be done by computation of the averaged displacement vector

$$L_n \equiv \langle V_j \rangle_{n_j = n},\tag{8}$$

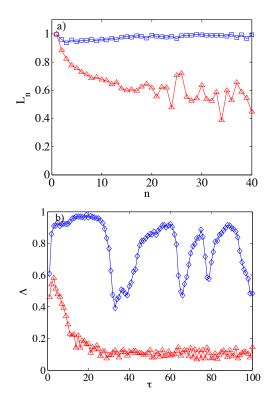
where the average is done over all boxes with same number n of trajectory passes. As shown in Kaplan and Glass [1993], the average displacement of n passes in m-dimensional phase space for the Brownian motion is

$$R_n = \frac{1}{\sqrt{n}} \left(\frac{2}{m}\right)^{1/2} \frac{\Gamma[(m+1)/2]}{\Gamma(m/2)},\tag{9}$$

where  $\Gamma$  is the gamma function. The deviation in  $\langle V_j \rangle$  between a given time series and the Brownian motion can be characterized by a single number given by the weighted average over all boxes of the quantity,

$$\Lambda(\tau) \equiv \frac{1}{\sum_{j} n_{j}} \sum_{j} n_{j} \frac{\langle V_{j} \rangle^{2}(\tau) - R_{n_{j}}^{2}}{1 - R_{n_{j}}^{2}}, \qquad (10)$$

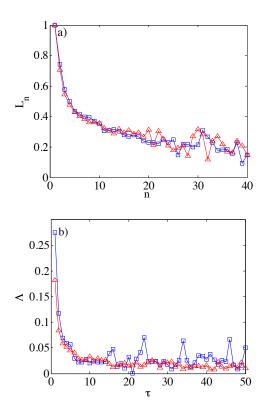
where we have explicitly highlighted that the averaged displacement  $\langle V_j \rangle(\tau)$  of the trajectory in the reconstructed



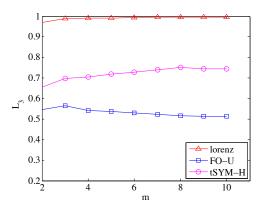
**Figure 4.** a)  $L_n$  vs. n: square symbols are derived from numerical solutions of the Lorenz system, and triangles from these solutions after randomization of phases of Fourier coefficients. b)  $\Lambda$  vs.  $\tau$ : diamonds from Lorenz system, and triangles after randomization of phases.

phase space depends on the time-lag  $\tau$ . For a completely deterministic signal we have  $\Lambda=1$ , and for a completely random signal  $\Lambda=0$ .

All systems described by the laws of classical (nonquantum) physics are deterministic in the sense that they are described by equations that have unique solutions if the initial state is completely specified. In this sense it seems



**Figure 5.** a) $L_n$  vs. n: square symbols are derived from numerical solutions of the fO-U stochastic equation, and triangles from these solutions after randomization of phases of Fourier coefficients. b)  $\Lambda$  vs.  $\tau$ : squares from fO-U equation, and triangles after randomization of phases.



**Figure 6.**  $L_3$  as a function of embedding dimension m for Lorenz equations (triangles), fO-U process (squares) and tSYM-H (circles).

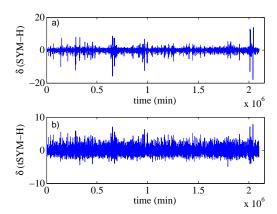


Figure 7. a) Increments for the SYM-H index. b) Increments for the transformed signal tSYM-H.

meaningless to provide tests for determinism. The test described in this section is really a test of low dimensionality. The test is performed by means of a time-delay embedding, for embedding dimension m up to a maximum value M, where M is limited by practical constraints. High M requires longer time series in order to achieve adequate statistics. A test that fails to characterize the system as deterministic for  $m \leq M$  in reality only tells us that the embedding dimension is too small, i.e. the number of degrees of freedom d of the system exceeds M/2. Such systems will in the following be characterized as random, or stochastic.

In Figure 4a, we plot  $L_n$  versus n for a time series generated as a numerical solution of the Lorenz system. In the same plot we also show the same characteristic for the surrogate time series generated by randomizing the phases of the Fourier coefficients of the original time series. This procedure does not change the power spectrum or autocovariance, but destroys correlation between phases due to nonlinear dynamics. For low-dimensional chaotic systems such randomization will change  $L_n$ , as is demonstrated for the Lorenz system in Figure 4a. We also calculate  $\Lambda$  versus  $\tau$  for these time series and plot the results in Figure 4b. Again,  $\Lambda$  for the original and surrogate time series are significantly different.

For the numerically generated fO-U process we observe in Figure 5 that  $L_n$  and  $\Lambda$  for the original and surrogate time series do not differ, demonstrating that these quantities are insensitive to randomization of phases of Fourier coefficients if the process is generated by a linear stochastic equation.

One should pay attention to the nature of the experimental data used in the test of determinism. For lowdimensional data contaminated by low-amplitude noise or Brownian motions, the analysis results will depend on the box size b, but the problem is solved by choice of sufficiently large box size. For a low-dimensional system represented by an attractor of dimension d the results may also depend on the choice of embedding dimension m. The estimated determinism  $L_n$  tends to increase with increasing m until it stabilizes at  $L_n \approx 1$  as m approaches 2d. For a random signal there is no such dependence on embedding dimension, as demonstrate by example in Figure 6. Here we plot the determinism  $L_3$  ( $L_n$  when n=3) versus embedding dimension m for the Lorenz and fO-U time series. For comparison we also plot this for geomagnetic storm index during magnetic storm times. It converges to a value less than 1, and for embedding dimensions higher than the Lorenz time series. This indicates that the geomagnetic index during magnetic storms exhibit both a random and a deterministic

component, and that the dimensionality of this component is higher than for the Lorenz system.

## 2.4. A test for predictability

This test is primarily developed for deterministic systems, which have limited predictability if the attractor is chaotic. Here we perform an analysis which is based on the diagonal line structures of the recurrence plot. In our study we use the average inverse diagonal line length:

$$\Gamma \equiv \langle l^{-1} \rangle = \sum_{l} l^{-1} P(l) / \sum_{l} P(l), \tag{11}$$

where P(l) is a histogram over diagonal lengths;

$$P(l) = \sum_{i,i=1}^{N} (1 - R_{i-1,j-1})(1 - R_{i+l,j+l}) \prod_{k=0}^{l-1} R_{i+k,j+k}.$$

For a low-dimensional, chaotic deterministic system (for which the embedding dimension is sufficient to unfold the attractor)  $\Gamma$  is an analog to the largest Lyapunov exponent, and is a measure of the degree of unpredictability.

For stochastic systems, the recurrence plots do not have identifiable diagonal lines, but rather consists of a pattern of dark rectangles of varying size, as observed in Figure 1. For embedding dimension m=1 such a dark rectangle corresponds to time intervals  $I_1 = (t_1, t_1 + \Delta t_1)$  on the horizontal axis and  $I_2 = (t_2, t_2 + \Delta t_2)$  on the vertical axis, for which the signal x(t) is inside the same  $\epsilon$ -interval whenever t is included in either  $I_1$  or  $I_2$ . In this case the length of unbroken diagonal lines l is a characteristic measure of the linear size of the corresponding rectangle, and the PDF P(l)a measure of the distribution of residence times of the trajectory inside  $\epsilon$ -intervals. For selfsimilar stochastic processes such as fractional Brownian motions P(l) can be computed analytically, and  $\Gamma$  computed as function of the selfsmilarity exponent h. Since the residence time l inside an  $\epsilon$ -box increases as the smoothness of the trajectory increases (increasing h), we should find that  $\Gamma(h)$  is a monotonically decreasing function of h. In section 3.3 we compute  $\Gamma(h)$ numerically for a synthetically generated fO-U process and thus demonstrate this relationship between  $\Gamma$  and h. Hence both  $\Gamma$  and h can serve as measures of predictability, but  $\Gamma$  is more general, because it is not restricted to selfsimilar processes or processes with stationary increments, and applies to low-dimensional chaotic as well as stochastic systems.

## 3. Results

In Rypdal and Rypdal [2010] it is shown that the fluctuation amplitude (or more precisely; the one-timestep increment)  $\Delta y(t)$  of the AE index is on the average proportional to the instantaneous value y(t) of the index. This gives rise to a special kind of intermittency associated with multiplicative noises, and leads to a non-stationary time series of increments. However, the time series  $\Delta y(t)/y(t)$  is stationary, implying that the stochastic process  $x(t) = \log y(t)$ has stationary increments. Thus, a signal with stationary increments, which still can exhibit a multifractal intermittency, can be constructed by considering the logarithm of the AE index. Similar properties pertain to the Dst and SYM-H, although in these cases we have to add a constant  $c_1$  before taking the logarithm, i.e.  $x(t) = \log(c_1 + c_2 y(t))$ has stationary increments. Using the procedure described in Rypdal and Rypdal [2010] the estimated coefficients are  $c_1 = 0.7725$  and  $c_2 = 0.0397$ . In Figure 7a we show the increments for the original SYM-H data, while in Figure 7b we show the increments for the transformed signal x(t), which in the following will be denoted tSYM-H.

## 3.1. Scaling of storm- and solar wind parameters

In this section, we employ EMD and variogram analysis to tSYM-H, IMF  $B_z$  and solar wind flow speed v. The EMD analysis is used to compute intrinsic mode functions (IF) for time intervals of 50000 minutes using data for the entire period from January 2000 till December 2005. The empirical variance estimates E versus mean period T for each IF component in tSYM-H,  $B_z$ , and v are shown as log-log plots in Figure 8a. In section 2.2 we mentioned that Flandrin and Gonçalvès [2004] has demonstrated that for fractional Gaussian noise the slope  $\zeta$  is equal to  $\zeta = 2H - 2$ , where H is the Hurst exponent. This estimate for the slope seems valid for our data as well, as is shown in the figure from comparison with the variogram, even though the time series on scales up to  $10^4$  minutes are non-stationary processes having the character of fractional Brownian motions [Rypdal and Rypdal, 2010. The results from the two different methods shown in Figures 8a and 8b are roughly consistent, using the relations h = H - 1 and  $\zeta = 2H - 2$ , which implies  $2h = \zeta$ . In practice, we have calculated  $\zeta$  from EMD as a function of 2h for fractional Gaussian noises and motions with self-similarity exponent h, and have derived a relation  $\zeta = 0.94 (2h) + 0.1143.$ 

The variogram represent a second order structure function:

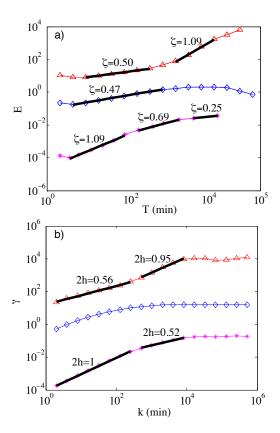
$$\gamma_t = \frac{1}{(N-t)} \sum_{n=1}^{N-t} (s_{n+t} - s_n)^2,$$
 (12)

which scales with a time-lag t as  $\gamma_t = t^{2h}$ , h is denoted as selfsimilarity exponent, and s is a time series. Note that a Hurst exponent H > 1 implies that the process is a nonstationary motion, and if the process is selfsimilar, the selfsimilarity exponent is h = H - 1. In our terminology a white noise process has Hurst exponent H = 0.5 and a Brownian motion has H = 1.5.

From Figure 8a we observe three different scaling regimes for tSYM-H. For time scales less than a few hundred minutes it scales like an uncorrelated motion ( $h \approx 0.5$ ). On time scales from a few hours to a week it scales as an antipersistent motion ( $h \approx 0.25-0.35$  depending on analysis method), and on longer time scales than a week it is close to a stationary pink noise ( $h \approx 0$ ). Similar behavior was observed for log AE in Rypdal and Rypdal [2010], but there the break between non-stationary motion and stationary noise (where h changes from h > 0 to  $h \approx 0$ ) occurs already after about 100 minutes, indicating the different time scales involved in ring current (storm) dynamics and electrojet (substorm) dynamics.

Results for v indicate a regime with antipersistent motion (h=0.25) up to a few hundred minutes, and then an uncorrelated or weakly persistent motion (h=0.5) up to a week. On longer time scales than this the variogram indicates that the process is stationary.

The exponent h for  $B_z$  can not be estimated from the variogram since it is difficult to obtain a linear fit to the concave curve in Figure 8b. The concavity is less pronounced in the curve derived from the EMD method i Figure 8a and  $\zeta = 0.47$ , corresponding to an antipersistent motion (h = 0.23) can be estimated on time scales up to a few hundred minutes.  $B_z$  becomes stationary already after a few hundred minutes, which is similar to the behavior in  $\log AE$ , as pointed out by Rypdal and Rypdal [2010]. In Rypdal and Rypdal [2011] the concavity of the variogram follows from modelling  $B_z$  as a (multifractal) Ornstein-Uhlenbeck process with a strong damping term that confines the motion on time scales longer than 100 minutes. Accounting for



**Figure 8.** a) E vs. T, b)  $\gamma_k$  vs. k: stars are for tSYM-H, diamonds for IMF  $B_z$ , and triangles for v. Note that a generalization of the result  $\zeta = 2H - 2$  in Flandrin and Gonçalvès [2004] yields  $2h = \zeta$ .

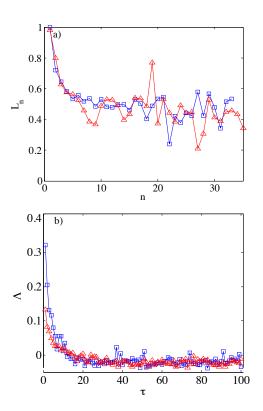
this confinement the "true" selfsimilarity exponent of the stochastic term turns out to be  $h\approx 0.5$ . Thus the antipersistence derived from the EMD analysis may be a spurious effect from this confinement. The conclusion in Rypdal and Rypdal [2011] is that  $B_z$  and  $\log AE$  behave as uncorrelated motions up to the scales of a few hours and become stationary on scales longer than this. Moreover, the stochastic term modelling the two signals share the same multifractal spectrum. In comparison, tSYM-H and v are non-stationary motions on scales up to a week before they reach the stationary regime.

## 3.2. Change of determinism during storm times

In Figure 9 we show  $L_n$  and  $\Lambda$  for tSYM-H and its surrogate time-series with randomized phases of Fourier coefficients. We observe that  $L_n$  and  $\Lambda$  for the surrogate time series does not deviate from those computed from the original tSYM-H, indicating that the dynamics of tSYM-H is not low-dimensional and nonlinear. The same results are obtained for IMF  $B_z$  and flow speed v.

In the following analysis we test for determinism in tSYM-H,  $B_z$  and v for ten intense storms that occurred in the period between 2000-2004. The reference point in our analysis is the storm's main phase, and then we analyze all the data spanning the time interval three days before and after the storm in tSYM-H,  $B_z$  and v.

We compute  $L_n$  for n=4 with a time resolution of 12 hours. The choice of n=4 from the  $L_n$  vs. n curve is a compromise between clear separation between low-dimensional



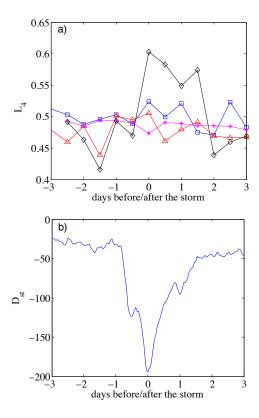
**Figure 9.** a)  $L_n$  vs. n: square symbols are for tSYM-H, and triangles are for this signal after randomization of phases of Fourier coefficients. b)  $\Lambda(\tau)$  vs.  $\tau$ : squares is for tSYM-H, and triangles are after randomization of phases.

and stochastic dynamics and small error bars (which increase with increasing n). In order to improve statistics for  $L_4$ , we compute determinism using data from all ten storms. As a reference, we compute  $L_4$  for the fO-U process as well. In Figure 10a we plot  $L_4$  for tSYM-H  $B_z$ , v, and fO-U, and in Figure 10b the  $D_{st}$  index averaged over all ten storms is plotted, since this index shows precisely when the storm takes place. We can observe that  $L_4$  is essentially the same for  $B_z$ , v and fO-U, and stays approximately constant during the course of a storm. However,  $L_4$  for tSYM-H increases during storm time.

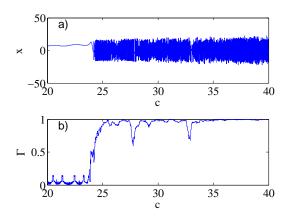
#### 3.3. Change of predictability during magnetic storms

Even though we deal with a predominantly stochastic system, its correlation and the degree of predictability changes in time, and our hypothesis is that such abrupt transitions in the dynamics take place during events like magnetic storms and substorms. We therefore employ recurrence plot quantification analysis as a tool for detection of these transitions.

We compute the average inverse diagonal line length  $\Gamma \equiv \langle l^{-1} \rangle$  as defined in equation (11), but the same results can be drawn from other quantities that can be derived from the recurrence plot (see *Marwan et al.*(2007)).  $\Gamma$  can be used as a proxy for the positive Lyapunov exponent in a system with chaotic dynamics, and is sensitive to the transition from regular to chaotic behavior, as can be shown heuristically for the case of the Lorenz system, where we use  $a=10,\ b=8/3$ , and c is varied from 20 to 40, such that transient behavior is obtained. For c=24.74 a Hopf bifurcation occurs, which corresponds to the onset of



**Figure 10.** a)  $L_4$  for IMF  $B_z$  (squares), v (triangles), tSYM-H (diamonds), fO-U (stars). b) averaged  $D_{st}$  index.



**Figure 11.** a) x component of the Lorenz system as a function of the parameter c. b)  $\Gamma \equiv \langle l^{-1} \rangle$  as a function of the parameter c.

chaotic flows. In Figure 11a we plot a bifurcation diagram for the x component of the Lorenz system as a function of the parameter c, while in Figure 11b we show  $\Gamma$  for the x component again as a function of the parameter c. Similar results have been obtained from the longest diagonal length, when applied to the logistic map  $Trulla\ et\ al.\ [1996]$ . We use embedding dimension m=1, because the results do not seem dependent on m and because, in the case of stochastic or high-dimensional dynamics, a topological embedding

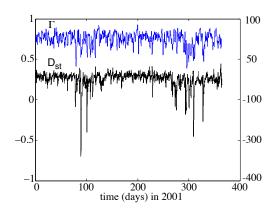


Figure 12. Upper graph:  $\Gamma$  for  $D_{st}$  index for year 2001. Lower graph:  $D_{st}$  index for the same period.

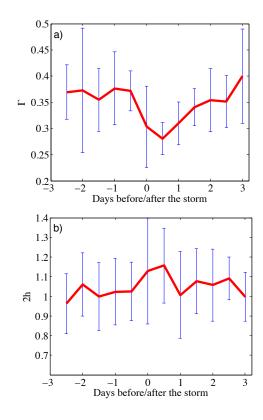
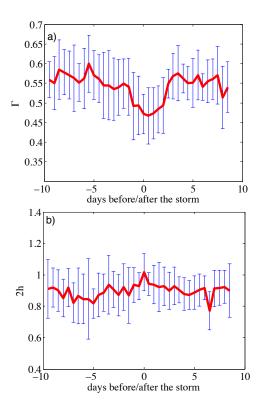


Figure 13. a)  $\Gamma$  for tSYM-H, b) 2h for tSYM-H averaged over ten storms. Error bars represent standard deviation based on data from these ten storms. Time origin is defined by the minimum of the  $D_{st}$  index in each storm.

cannot be achieved for any reasonable embedding dimension. This fact demonstrates the robustness of the recurrence-plot analysis, which responds to changes in the dynamics of the system even if it is a stochastic or high-dimensional system for which no proper phase-space reconstruction is possible. Next, we show how  $\Gamma$  in  $D_{st}$  index varies during the year of 2001. That year two major storms occurred: one on March, 31st, and the other on November, 6th. We compute  $\Gamma$  over



**Figure 14.** a)  $\Gamma$  for  $B_z$ , b) 2h for  $B_z$  averaged over the same storms as in the previous figure.

windows of 24 hours (only 24 points since this index is only given with hourly resolution).

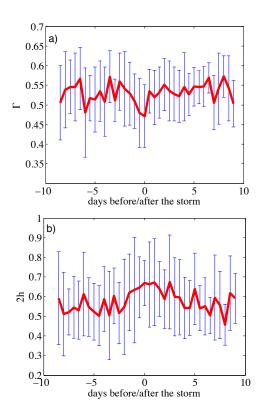
In Figure 12 we display  $\Gamma$  and the  $D_{st}$  index itself as a function of time and observe that predictability increases (Γ decreases) around storm time. Similar conclusions were drawn by Bálasis et al. [2006]. Since reduction in  $\Gamma$  means increase of predictability it may also be a signature of higher persistence in a stochastic signal. This motivates plotting  $\Gamma$ and 2h (computed as a linear fit from the variogram over the time scales up to 12 hours) for both solar wind parameters and tSYM-H. Figure 13a shows Γ for tSYM-H averaged over 10 magnetic storms while Figure 13b shows the same for 2h. Also, in Figure 14 and Figure 15 the same characteristics are plotted for  $B_z$  and v. In each of these parameters we observe a slight increase of 2h, and a decrease of  $\Gamma$ , suggesting increase of persistence and predictability during storms. As a reference,  $\Gamma$  and 2h for hundred realizations of fO-U were simulated with the coefficients in the stochastic equation fitted to model the tSYM-H signal. For these synthetic data we find  $\Gamma = 0.4$  and 2h = 1, in good agreement with the results obtained from the tSYM-H time series. The relationship between  $\Gamma$  and h can be explored through numerical realizations of fO-U processes. Figure 16 shows Γ computed for varying h as a mean value of 100 realizations of such a process for each h. For persistent motions (h > 0.5) there is a linear dependence between  $\Gamma$  and h, and a best fit yields

$$\Gamma \approx 0.72 - 0.57h. \tag{13}$$

This analysis shows the importance of  $\Gamma$  as a universal measure for predictability: in low-dimensional systems it is a measure for the Lyapunov exponent, while for persistent stochastic motions it is a measure of persistence through the relation 13.

# 4. Discussion of results on determinism

The determinism (as measured by  $L_4$ ) of the storm index tSYM-H has been shown to exhibit a pronounced increase at storm time. A rather trivial explanation of this enhancement would be that it is caused by the "trend" incurred by the wedge-shaped drop and recovery of the storm indices associated with a magnetic storm. We test this hypothesis by superposing such a wedge-shaped pulse to an fO-U process and compute  $L_4$ . Next, we take tSYM-H for ten storms and for each set of data subtract the wedge-shaped pulse (computed by a moving-average smoothing). The residual signal represents the "detrended" fluctuations. The result is shown



**Figure 15.** a)  $\Gamma$  for v, b) 2h for v averaged over the same storms as in the previous figures.

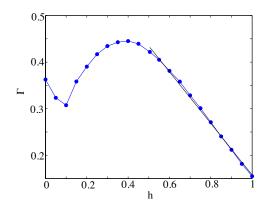


Figure 16.  $\Gamma$  vs. h for fO-U.

in Figure 17 and reveals that the trend has no discernible influence on the determinism while, on the other hand, we observe that the enhancement of  $L_4$  around storm time persists in the "detrended" fluctuations. This result suggests that the increasing determinism during storms is a result of an enhanced low-dimensional component in the storm indices. As mentioned in section 2.3 for low-dimensional dynamics, nonlinearity may be important for the measure of determinism. For a nonlinear, low-dimensional system the destruction of nonlinear coupling by randomizing phases of Fourier coefficients will in general reduce the determinism, while for a linear, stochastic process we will observe no such effect. But what role will nonlinearity play if it is introduced in the deterministic terms of a stochastic equation? The deterministic term in the fractional Langevin equation representing the fO-U is a linear damping term. However, the best representation of the damping/drift term in an fO-U model for tSYM-H is not linear. Following Rypdal and Rypdal [2011], if y = tSYM-H, the drift term is given as the conditional probability density  $\delta y(t, \delta t)$  given that y(t) = y:

$$M(y, \delta t) = E[\delta y(t, \delta t) \mid y(t) = y]. \tag{14}$$

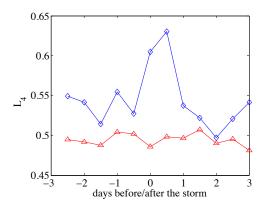
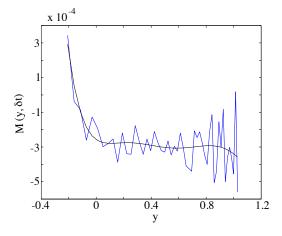


Figure 17.  $L_4$ : triangles are derived from an fO-U process with a "storm trend" imposed, diamonds are derived from the "detrended" tSYM-H.



**Figure 18.** The drift term in the fO-U equation computed from tSYM-H. The smooth solid curve is a six-order polynomial fit.

In fO-U  $M(y, \delta t)$  is a linear function of y, but a polynomial fit to drift term derived from tSYM-H data requires a sixth order polynomial, confirming the nonlinearity of the tSYM-H process. This is shown in Figure 18. Next, we test determinism for the nonlinear fO-U process, whose scaling exponent h is estimated from the variogram of tSYM-H. Figure 19a shows  $L_n$  vs. n for numerical realizations of this process compared with the same analysis after randomization of the phases of the Fourier coefficients. The result reveals that the nonlinear fO-U process is not more deterministic than its randomized version. Next, we add the x-component of the Lorenz attractor  $x = x_L + 1.85x_n$ , where  $x_L$  is the solution of the Lorenz system and  $x_n$  is the nonlinear fO-U process, both signals with zero mean and unit variance. Now  $L_n$  versus n decreases when the phases are randomized, as shown in Figure 19b, which further lead us to the conclusion that determinism is a measure of low-dimensionality.

## 5. Conclusions

The storm indices  $(D_{st}$  and SYM-H) and the solar wind observables (flow velocity v and IMF  $B_z$ ) show no clear signatures of low-dimensional dynamics during quiet periods. However, low-dimensionality increases in SYM-H during storm times, indicating that self-organization of the magnetosphere takes place during magnetic storms. This conclusion is drawn from the study of ten intense, magnetic

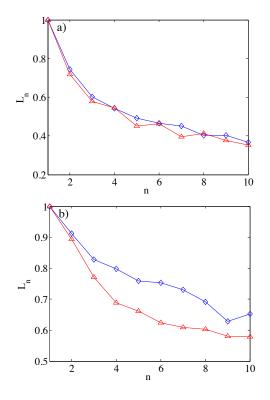


Figure 19.  $L_n$  vs. n. a) diamonds are derived from numerical solutions of the nonlinear fO-U. Triangles are from these solutions after randomization of phases of Fourier coefficients. b) diamonds are derived from numerical solutions of the nonlinear fO-U with a solution of the x component of the Lorenz system superposed. Triangles are from the latter signals after randomization of phases.

storms in the period from 2000-2004. Even though our analysis shows no discernible change in determinism during magnetic storms for solar wind parameters, there is an enhancement of the predictability of the solar wind observables as well as the geomagnetic storm indices during major storms. We interpret this as an increase in the persistence of the stochastic components of the signals.

There are, however, significant differences between  $B_z$  on one hand, and v and  $D_{st}/{\rm SYM-H}$  on the other. While the former is a non-stationary, slightly anti-persistent, motion up to time scales of approximately 100 minutes, and a pink noise on longer time scales, the latter are slightly persistent motions on scales up to several days and noises on longer time scales. These differences represent the different physics and time constants of the storm ring current systems, but also the different role the solar wind  $B_z$  and the velocity v play in driving this current system. Magnetic storms are typically preceded by an increase in the persistence of the solar wind dynamics, and this increase is reflected in the magnetospheric response to the solar wind driver.

Acknowledgments. Recurrence plot and its quantities are computed by means of the Matlab package downloaded from http://www.agnld.uni-potsdam.de/~marwan/toolbox/.

The authors acknowledge illuminating discussions with M. Rypdal and B. Kozelov.

## References

- Abarbanel, H. (1996), Analysis of observed chaotic data, Institute for nonlinear science, Springer, New York.
- Akasofu, S. I. (1965), The development of geomagnetic storms without a preceding enhacement of the solar plasma pressure, Planet. Space Sci., 13, 297.
- Angelopoulos, V., T. Mukai, S. Kokobun (1999), Evidence for intermittency in Earth's plasma sheet and implications for self-organized criticality, *Phys. Plasmas*, 6, 4161.
- Bak, P., C. Tang, and K. Wiesenfeld (1987), Self-organized criticality: An explanation of 1/f noise, Phys. Rev. Lett., 59, 381.
- Balasis, G., I. A. Daglis, P. Kapiris, M, Mandea, D. Vassiliadis, K. Eftaxias (2006), From pre-storm activity to magnetic storms: a transition described in terms of fractal dynamics, Ann. Geophys, 24, 3557.
- Beran, J. (1994), Statistics for long-memory processes, Monographs on statistics and applied probability, Chapman& Hall/CRC, Boca Raton.
- Chang, T. (1998), Self-organized criticality, multi-fractal spectra, and intermittent merging of coherent structures in the magnetotail, Astrophysics and space science, 264, 303.
- Chapman, S.C., N. Watkins, R. O. Dendy, P. Helander, and G. Rowlands (1998), A simple avalanche model as an analogue for magnetospheric activity, *Geophys. Res. Lett.*, 25, 2397.
- Consolini, G., and T. S. Chang (2001), Magnetic field topology and criticality in geotail dynamics: relevance to substorm phenomena, Space Sci. Rev., 95, 309.
- Davies, T. N, and M. Sugiura (1966), Auroral electrojet activity index AE and its universal time variations, J. Geophys. Res., 71, 785.
- Eckmann, J. P, S. O. Kamphorst and D. Ruelle (1987), Europhys. Lett. 5, 973.
- Flandrin, P., and P. Gonçalvès (2004), Empirical mode decomposition as data-driven wavelet- like expansion, International Journal of Wavelets, Multiresolution and Information Processing, 2, 1.
- Franzke, C. (2009), Multi-scale analysis of teleconnection indices: climate noise and nonlinear trend analysis, *Nonlin. Processes Geophys.*, 16, 65.

- Gonzalez, W. D., J. A. Joselyn, Y. Kamide, H. W. Kroehl, G. Rostoker, B. T. Tsurutani, and V. M. Vasyliunas (1994), What is a geomagnetic storm?, J. Geophys. Res., 99, 5771.
- Huang, N.E, Z. Seng, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. Yen, C. C. Thung, H. H. Liu (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. Lond. A 454, 903.
- Kaplan, D. T., and L. Glass (1992), Direct test for determinism in a time series, Phys. Rev. Lett., 68, 427.
- Kaplan, D. T., and L. Glass (1993), Coarse-grained embedding of time series: random walks, Gaussian random processes, and deterministic chaos, *Physica D*, 64, 431.
- Klimas, A., D. Vassiliadis, D. N. Baker, and D. A. Roberts (1996), The organized nonlinear dynamics of the magnetosphere, J. Geophys. Res., 101, 13, 089-13,113.
- Loewe, C. A., and G. W. Prölss (1997), Clasiffication and mean behavior of magnetic storms, J. Geophys. Res., 102, 14,209.
- Marwan, N., M. C. Romano, M. Thiel and J. Kürths (2007), Recurrence plots for the analysis of complex systems *Physics Reports* 438, 237.
- Rilling, G., P. Flandrin, and P. Goncalves (2003), On empirical mode decomposition and its algorithms, IEEE-EURASIP Workshop on Nonlinear signal and image processing NSIP-03, GRADO(I).
- Rypdal, M., and K. Rypdal (2010), Stochastic modeling of the AE index and its relation to fluctuations in  $B_z$  of the IMF on time scales shorter than substorm duration, *J. Geophys. Res.*, 115. A11216.
- Rypdal, M., and K. Rypdal (2011), Discerning a linkage between solar wind turbulence and ionospheric dissipation by a method of confined multifractal motions, J. Geophys. Res., 116, A02202.
- Sharma, A. S., D. Vassiliadis, and K. Papadopoulos (1993), Reconstruction of low-dimensional magnetospheric dynamics by singular spectrum analysis, Geophys. Res. Lett., 20, 335.
- Sharma, A. S., A. Y. Ukhorskiy, and M. Sitnov (2003), Modeling the magnetosphere using time series data, Geophysical Monograph, 142, 231.
- Sitnov, M. I., A. S. Sharma, and K. Papadopoulos (2001), Modeling substorm dynamics of the magnetosphere: From self-organization and self-organized criticality to nonequlibrium phase transitions, *Phys. Rev. E*, 65, 016116.
- Takens, F. (1981), Detecting strange attractors in fluid turbulence, in: Dynamical Systems and Turbulence, edited by D. Rand, and L. S. Young, Springer, Berlin.
- Trulla, L. L, A. Guliani, J. P. Zbilut, and C. L. Webber, Jr. (1996), Recurrence quantification analysis of the logisite equation with transients, *Phys. Lett. A* 223, 255.
- Vassiliadis, D. V, A. S. Sharma, T. E. Eastman, and K. Papadopoulos (1990), Low-dimensional chaos in magnetospheric activity from AE time series, Geophys. Res. Lett., 17, 1841.
- Wanliss, J. A., and K. M. Showalter (2006), High-resolution global storm index: Dst versus SYM-H, J. Geophys. Res., 111, A02202.
- Watkins, N. W., (2002), Scaling in the space climatology of the auroral indices: is SOC the only possible description, Nonlin. Processes in Geophys., 9, 389.
- Wu, B. Z, and N. E. Huang (2004), A study of the characteristics of white noise using the empirical mode decomposition method, Proc. R. Soc. Lond. A, 460, 1597.
- Wu, B. Z, N. E. Huang, S. R. Long, C.K. Peng (2007), On the trend, detrending, and variability of nonlinear and nonstationary time series, PNAS, 104, 38.
- T. Živković, Department of physics and Technology, University of Tromsø, Prestvannveien 40, 9037 Tromsø, Norway (tatjana.zivkovic@uit.no)
- K. Rypdal, Department of physics and Technology, University of Tromsø, Prestvannveien 40, 9037 Tromsø, Norway