



POTSDAM-INSTITUT FÜR  
KLIMAFOLGENFORSCHUNG

**Originally published as:**

**Donner, R. V., Balasis, G. (2013):** Correlation-based characterisation of time-varying dynamical complexity in the Earth's magnetosphere. - *Nonlinear Processes in Geophysics*, 20, 6, 965-975

**DOI:** [10.5194/npg-20-965-2013](https://doi.org/10.5194/npg-20-965-2013)



# Correlation-based characterisation of time-varying dynamical complexity in the Earth's magnetosphere

R. V. Donner<sup>1,2</sup> and G. Balasis<sup>3</sup>

<sup>1</sup>Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Hans-Knöll-Straße 10, 07745 Jena, Germany

<sup>2</sup>Research Domain IV – Transdisciplinary Concepts & Methods, Potsdam Institute for Climate Impact Research, Telegrafenberg A31, 14473 Potsdam, Germany

<sup>3</sup>Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, I. Metaxa & Vas. Pavlou St., 15236, Penteli, Greece

Correspondence to: R. V. Donner (reik.donner@pik-potsdam.de)

Received: 3 July 2013 – Revised: 23 September 2013 – Accepted: 26 September 2013 – Published: 12 November 2013

**Abstract.** The dynamical behaviour of the magnetosphere is known to be a sensitive indicator for the response of the system to solar wind coupling. Since the solar activity commonly displays very interesting non-stationary and multi-scale dynamics, the magnetospheric response also exhibits a high degree of dynamical complexity associated with fundamentally different characteristics during periods of quiescence and magnetic storms. The resulting temporal complexity profile has been explored using several approaches from applied statistics, dynamical systems theory and statistical mechanics. Here, we propose an alternative way of looking at time-varying dynamical complexity of nonlinear geophysical time series utilising subtle but significant changes in the linear autocorrelation structure of the recorded data. Our approach is demonstrated to sensitively trace the dynamic signatures associated with intense magnetic storms, and to display reasonable skills in distinguishing between quiescence and storm periods. The potentials and methodological limitations of this new viewpoint are discussed in some detail.

## 1 Introduction

Accumulated evidence points to the complex nonlinear character of magnetosphere dynamics. The seminal paper by Tsurutani et al. (1990) provided one of the first indications of nonlinear processes in the magnetosphere, claiming that there is a nonlinear response of the AE index to the southward component of the interplanetary magnetic field (IMF). Subsequently, several authors (Baker et al., 1990; Vassiliadis

et al., 1990; Sharma et al., 1993; Pavlos et al., 2003; Vörös et al., 2003) studied the occurrence of low-dimensional chaos in magnetospheric activity, while Klimas et al. (1996) discussed the nonlinear characteristics of magnetosphere dynamics and Angelopoulos et al. (1999) witnessed several aspects of the dynamical complexity of the plasma sheet. Thus, recent advances in the study of complexity and complex systems open new research perspectives to the investigation of the magnetospheric dynamics (Wanliss, 2005; Chang et al., 2010; De Michelis et al., 2012).

Dynamical complexity detection for output time series of complex systems is one of the foremost problems in physics, biology, engineering and economic sciences. Especially in geomagnetism and magnetospheric physics, accurate detection of the dissimilarity between normal and abnormal states (e.g. pre-storm activity and magnetic storms) can vastly improve geomagnetic field modelling as well as space weather forecasting, respectively.

Entropy measures (e.g. non-extensive Tsallis entropy, Shannon entropy, block entropy, Kolmogorov entropy, T-complexity and approximate entropy) have been proven effectively applicable for the investigation of dynamical complexity in time series of the disturbance storm time (Dst) geomagnetic activity index (Balasis et al., 2008, 2009). It has been demonstrated that during transitions between non-storm and stormy periods of the magnetosphere, there is clear evidence of significantly lower complexity in the magnetosphere. The observed higher degree of organisation of the system agrees with results previously inferred from fractal analysis via estimates of the Hurst exponent based on wavelet

transforms (Balasis et al., 2006). This convergence between entropies and linear analyses provides a more reliable detection of the transition from the quiet-time to the storm-time magnetosphere, thus showing evidence that the occurrence of an intense magnetic storm is imminent.

Moreover, based on the general behaviour of complex system dynamics, it has been recently found that Dst time series exhibit discrete scale invariance, which in turn leads to log-periodic corrections to scaling that decorate the pure power law (Balasis et al., 2011c). The latter can be used for the determination of the time of occurrence of an approaching magnetic storm.

In this work, we propose a complementary view on the time-varying dynamical complexity of the Earth's magnetosphere by making use of higher-order characteristics based on the linear autocorrelation structure of Dst time series. Specifically, we modify a recent approach to quantifying the fraction of dynamically relevant variables in multivariate data sets, the linear variance decay (LVD) dimension density method (Donner and Witt, 2006), in such a way that it becomes applicable to univariate time series. Subsequently, the method is applied to a well-studied part of the Dst time series covering the entire year of 2001, which exhibited two particularly marked periods with several intense magnetic storms. Necessary details on both the method and the data are described in Sects. 2 and 3, respectively. The results of our analysis are summarised in Sect. 4, and indicate that the proposed measure has a good capability of tracing temporal variations in the dynamical complexity of nonlinear geophysical time series as reflected in their linear autocorrelation structure. The potentials and methodological limitations of this new approach are studied and discussed in detail. Finally, a methodological framework is proposed for comparing the performance of different dynamical characteristics displaying signatures of time-dependent complexity in geophysical systems.

## 2 Method

In order to characterise the time-varying dynamical complexity of geophysical time series, we exploit some characteristic scaling properties associated with the linear autocovariance structure of the data. Specifically, we utilise a univariate variant of the recently proposed LVD dimension density, which was originally introduced for studying temporal changes in the effective number of linearly independent components in multivariate geophysical recordings (Donner and Witt, 2006).

In the classical multivariate framework, the LVD dimension density characterises the exponential decay of normalised residual variances when performing a dimensionality reduction based on the matrix of pairwise linear (Pearson) correlation coefficients between the individual components (i.e. a principal component analysis, PCA). Specifically, with

$\sigma_i^2$  being the eigenvalues of the correlation matrix given in descending order and normalised to unit sum, it has been shown (Donner and Witt, 2006) that one can find an approximate scaling law:

$$\rho_p = \frac{\sum_{i=p+1}^N \sigma_i^2}{\sum_{i=1}^N \sigma_i^2} \simeq \exp\left(-\frac{p}{N} / \delta(p)\right). \quad (1)$$

The scaling parameter  $\delta(p)$  characterises how fast the residual variances  $\rho_p$  decay towards zero, i.e. how much dynamically relevant information is contained in the leading principal components in comparison with the “background noise”. A continuous estimate of this parameter as a function of the fraction  $f \in (0, 1)$  of explained variance, which has been referred to as the (non-normalised) LVD dimension density  $\delta(f)$ , has been introduced by Donner and Witt (2006); details about the corresponding estimator have been given in Donner et al. (2008). Notably,  $f$  acts as a free parameter (though with typically only minor relevance for the estimator), which is typically chosen as  $f = 0.95$  or  $f = 0.99$  in many real-world situations.

Considering the theoretical minimum and maximum values of  $\delta(f)$  (for perfectly linearly correlated or pairwise linearly independent components, respectively), the normalised LVD dimension density is defined as (Xie et al., 2011)

$$\delta_{\text{LVD}}(f) = \frac{\delta(f) - \delta_{\min}(f)}{\delta_{\max}(f) - \delta_{\min}(f)}. \quad (2)$$

Notably, we have  $\delta_{\text{LVD}} = 0$  if all pairwise linear correlation coefficients between the components of the considered record approach  $\pm 1$ , and  $\delta_{\text{LVD}} = 1$  if they are linearly independent.

One main goal of the present work is the straightforward generalisation and subsequent application of the approach described above to the study of univariate geophysical time series. We note that this idea has been originally mentioned by Donner (2007) and has already been used for characterising complexity patterns among sets of time series from climate as well as engineering contexts (Donner, 2012; Toonen et al., 2012). Specifically, in order to make the LVD dimension density method applicable to univariate data, we replace the formerly considered (cross-) correlation matrix with the Toeplitz matrix of linear autocorrelation values (Donner, 2012). In this case, the formerly fixed number of components  $N$  becomes an additional parameter of the method, which is conceptually related with the embedding dimension in classical time-delay embedding. However, in contrast to time-delay embedding for attractor reconstruction calling for a minimum number of sufficiently decorrelated time-shifted replications of a time series, our method makes use of much higher values of  $N$ , i.e. an intentional “over-embedding”, for allowing access to the scaling properties discussed above. Numerical studies for model systems as well as real-world geoscientific time series reveal that

the estimates of  $\delta_{LVD}$  obtained using this framework commonly approach stationary values as  $N$  is increased (Donner, 2012)<sup>1</sup>.

### 3 Description of the data

The hourly Dst index values have been widely used as a proxy of the magnetospheric ring current strength and, consequently, the intensity of geospace magnetic storms, which are considered the most complex phenomenon of magnetospheric dynamics (Gonzalez et al., 1994; Daglis, 2006). The Dst index is computed as the average change of the horizontal component of the Earth's magnetic field recorded at four mid-latitude magnetic observatories (<http://wdc.kugi.kyoto-u.ac.jp/index.html>). Here, we consider about one year of Dst data (1 January 2001–31 January 2002) including two particularly intense magnetic storms, which occurred on 31 March 2001 and 6 November 2001 with minimum Dst values of  $-387$  and  $-292$  nT, respectively, as well as a number of weaker events (e.g. in May and August 2001 with  $Dst \lesssim -100$  nT in both cases).

## 4 Results

### 4.1 Time-varying complexity

We obtain a quantitative characterisation of the time-varying dynamical complexity of the Earth magnetosphere in terms of the univariate normalised LVD dimension density (see Sect. 2), which is utilised here for the first time in a sliding window analysis. Specifically, we study the Dst time series for windows of width  $w$  and a mutual offset of  $\Delta w$ , using an “embedding” dimension  $N < w - \Delta w$ .

The results of our analysis for some reasonable choice of the method's intrinsic parameters (see below for a more detailed discussion) are shown in Fig. 1. In general, we find the qualitative behaviour being relatively robust under modifications of these parameters over a reasonable range. Specifically, given a sufficiently large choice of  $N$  which takes all relevant scales of magnetospheric dynamics (short-term fluctuations as well as relaxation processes to equilibrium conditions after magnetic storms) into account, we observe several distinct minima of  $\delta_{LVD}$  coinciding with the timing of intense magnetic storms. This result is understood as indicating a very high degree of temporal organisation of the system (i.e. the system exhibits a continuous variability mode for a significant amount of time, representing storm onset and subsequent relaxation as persistent trends without remarkable stochastic fluctuations superimposed to this mode).

<sup>1</sup>Note that in a similar way, one could additionally consider multiples of a fixed time shift  $\tau$  (also known as embedding delay) for resolving the scaling properties with respect to a given basic timescale, an approach that we will not further discuss here. In the following, we will always consider  $\tau = 1$  (the basic time unit as defined by the original sampling step of the data under study).

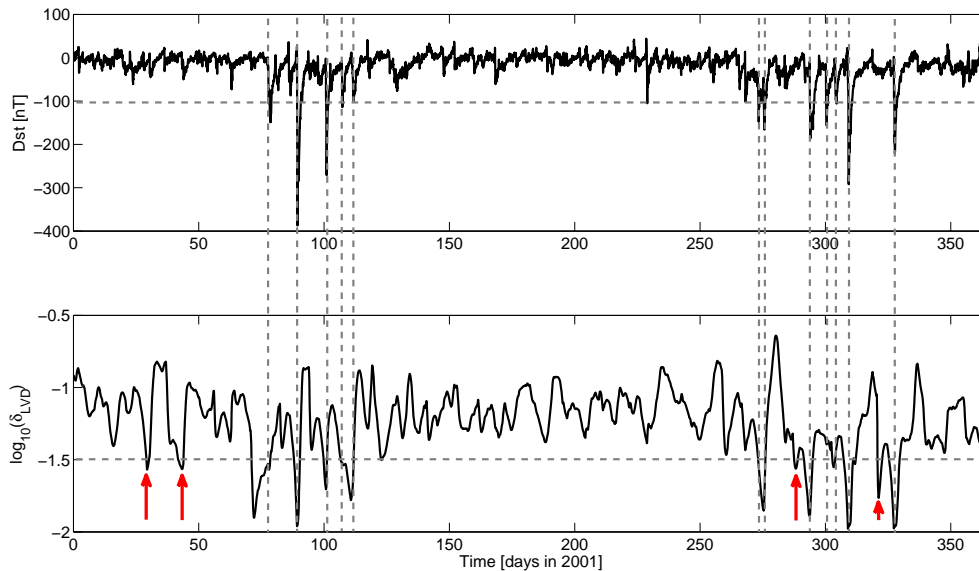
### 4.2 Comparison with other methods

The obtained results regarding qualitative changes in the dynamical properties during intense magnetic storms coincide well with recent findings using frequency-domain characteristics as well as concepts from non-equilibrium statistical mechanics (see Fig. 2).

On the one hand, Balasis et al. (2006, 2011a) applied two independent estimates of the Hurst exponent  $H$  (Hurst, 1951) – a fractal spectral analysis technique based on wavelet transforms and the rescaled range analysis method – to the Dst index variations for 2001. Their analyses allowed for discrimination between normal and abnormal states of the magnetosphere by showing that the abnormal state (associated with intense magnetic storms) complies with the fractional Brownian motion (fBm) model with persistent behaviour ( $H > 0.5$ ), while, conversely, the normal state follows the same model but with anti-persistent behaviour ( $H < 0.5$ ).

On the other hand, Balasis et al. (2008, 2009) analysed the same Dst data set in terms of non-extensive Tsallis entropy (Tsallis, 1988). It is well known that the traditional Shannon entropy (Shannon, 1948) works best in dealing with systems composed of subsystems which can assess all the available phase space and which are either independent or interact via short-range forces. For systems exhibiting long-range correlations, memory or fractal properties, non-extensive Tsallis entropy becomes the most appropriate mathematical tool (Balasis et al., 2011b, d; Tsallis, 2011). A central property of magnetic storm emergence is the occurrence of coherent large-scale collective behaviour with a very rich structure, resulting from repeated nonlinear interactions among the constituents of the global geospace system. In this respect, Tsallis entropy is an appropriate tool for identifying magnetic storm precursors. Specifically, it sensitively shows the dissimilarity of complexity among different “physiological” (normal) and “pathological” states (intense magnetic storms) – similar to the fBm Hurst exponent, Tsallis entropy implies the emergence of two distinct patterns: (i) a pattern associated with intense magnetic storms, which is characterised by a higher degree of organisation, and (ii) a pattern associated with normal (quiet-time) periods, which exhibits a lower degree of organisation. We stress that the anti-persistent time windows correspond to the time windows of high entropy, while the persistent time windows correspond to the time windows of low entropy.

The higher degree of organisation may reflect that “a sufficiently intense and long-lasting interplanetary convection electric field leads, through a substantial energisation in the magnetosphere–ionosphere system, to an intensified ring current sufficiently strong to exceed some key threshold of the quantifying storm time Dst index”, which defines a magnetic storm (Gonzalez et al., 1994).



**Fig. 1.** Dst time series for the year 2001 (top) and temporal profile of the LVD dimension density  $\delta_{\text{LVD}}$  (in logarithmic scale) for  $f = 0.95$ ,  $w = 168$  h,  $\Delta w = 1$  h and  $N = 96$ . For better visibility, intense magnetic storms are indicated by vertical dashed lines; almost all of them are characterised by particularly low values of  $\delta_{\text{LVD}}$  (the corresponding horizontal dashed line displayed in the lower panel is intended to guide the reader's eye and does not represent any kind of optimality criterion for distinguishing between storm and quiescence periods). In turn, we find some additional episodes of low  $\delta_{\text{LVD}}$  without adjoint intense magnetic storms (marked by red arrows), pointing to a different kind of strong temporal organisation in the data over the considered time window.

Figure 1 demonstrates that  $\delta_{\text{LVD}}$  primarily traces the gradual trend associated with the sudden drop in the Dst index followed by the recovery of the magnetosphere, which provides the generic signature of intense magnetic storms. Specifically,  $\delta_{\text{LVD}}$  is plotted as a function of the beginning of the time window of width  $w = 7$  days used for its estimation (this choice of  $w$  is motivated by the typical duration of magnetic storms including post-storm magnetospheric relaxation phases as well as the intention of resolving temporal changes on the associated timescale; see Sect. 4.4 for further discussion). In this representation, most sharp minima of the LVD dimension density coincide precisely with the timing of intense magnetic storms, implying that the decrease in dynamical complexity is actually associated with the recovery phase of the magnetosphere.

The latter behaviour is clearly different from the findings reported for the Hurst exponent and Tsallis entropy, which show consistent trends even considerably *before and after* the storm and thus can be thought of as indicating possible precursory structures. However, note that besides their general trends, both the Hurst exponent and Tsallis entropy often exhibit sharp increases or decreases at the timing of a magnetic storm, respectively (Fig. 2) – whereas these signatures even enhance the trend towards persistent dynamics for the Hurst exponent, they actually decrease the discriminatory power of the Tsallis entropy by leading to values similar to those obtained during periods of magnetospheric quiescence.

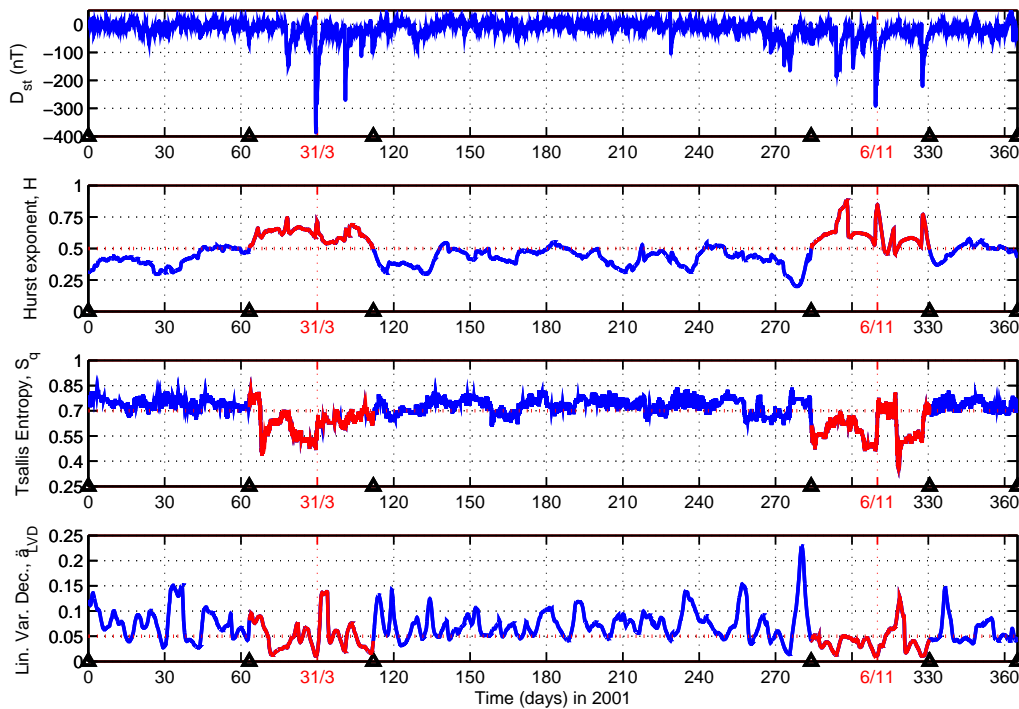
### 4.3 Which properties are captured by $\delta_{\text{LVD}}$ ?

There are multiple possible reasons for the apparently different behaviour of  $\delta_{\text{LVD}}$  in comparison with the other recently studied measures. In the following, we discuss some of these effects in order to develop a better understanding of the specific properties captured by our “linear” complexity measure.

#### 4.3.1 Short-range fluctuations vs. temporary trends

$\delta_{\text{LVD}}$  has been previously found to often display a certain relationship with the decay of correlations as expressed by, for example, the lag-one autocorrelation value (Donner, 2012). In turn, systematic differences between the autocorrelation values of processes with short- and long-range memory can be expected to mainly exist at considerably large time increments  $\tau$ , which in turn would provide only minor contributions to the Toeplitz matrix of autocorrelations, and thus affect the scaling of the residual variances obtained from this matrix to a somewhat lesser extent. With this in mind,  $\delta_{\text{LVD}}$  could be expected to be a less sensitive indicator for the *type* of persistence (which has been studied before in terms of the Hurst exponent, cf. Balasis et al., 2006, 2011a), but rather for the associated de-correlation time.

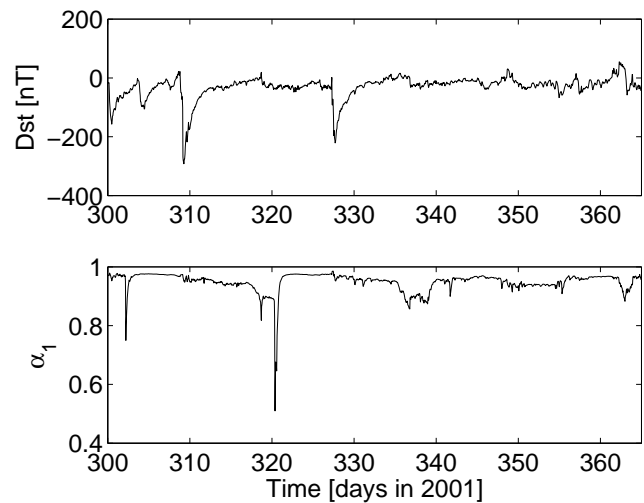
In order to further clarify this point, Fig. 3 shows the behaviour of the lag-one autocorrelation  $\alpha_1$  for some part of the considered Dst record. As can be seen,  $\alpha_1$  generally takes high values (mostly above 0.9), but exhibits sharp drops whenever data associated with a strong magnetic storm are



**Fig. 2.** From top to bottom: Dst time series along with time variations of Hurst exponent ( $H$ , wavelet spectra-based estimate of Balasis et al., 2006), Tsallis entropy ( $S_q$ ) and linear variance decay dimension density ( $\delta_{LVD}$ ). The activity periods around 31 March and 6 November 2001 magnetic storms are marked red. In the subplot of  $H$ , the red dashed line marks the transition between anti-persistent and persistent behaviour ( $H = 0.5$ ). In the subplot of  $S_q$ , the red dashed line gives a possible boundary value for the transition to lower complexity, characterising the different state of the magnetosphere during intense magnetic storms. In the subplot of  $\delta_{LVD}$ , the red dashed line can be related to transitions between physiological and pathological behaviours of the magnetosphere during low and intense magnetic activity, respectively. The triangles denote five time intervals, in which the first, third and fifth intervals correspond to anti-persistent ( $0 < H < 0.5$ ) epochs with high Tsallis entropies, whereas the second and fourth time windows exhibit persistent ( $0.5 < H < 1$ ) behaviour of Dst with lower Tsallis entropies.

taken into account for the first time (i.e. for time windows starting about  $w$  time steps before the storm). In turn, these drops are very well localised in time and persist only for a time interval of at most about one day before returning again to values close to 1 when data corresponding to the post-storm relaxation phase of the magnetosphere contribute. In this spirit, for the performed running window analysis, there is no clear relationship between  $\delta_{LVD}$  and  $\alpha_1$  like recently found within an ensemble of temperature time series (Donner, 2012), where the same methodology has been utilised for studying spatial patterns of *average* dynamical complexity over a time span of more than 50 yr.

The observed behaviour of  $\alpha_1$  underlines that the autocorrelation function is generally sensitive to temporary trends associated with magnetic storms, and as such different measures based on this function display different sensitivity with respect to changes in the short-term fluctuations contained in the data. Notably, this observation also applies to other techniques that do not make explicit use of detrending techniques.



**Fig. 3.** Part of the Dst time series (top) and temporal profile of the lag-one autocorrelation coefficient  $\alpha_1$  ( $w = 168$  h,  $\Delta w = 1$ ).

In addition to  $\delta_{LVD}$  and  $\alpha_1$ , we have also considered other eigenvalue-based measures (not shown), such as the (normalised) largest eigenvalue  $\sigma_1^2$  as well as the associated participation ratio

$$PR(\sigma_1^2) = \left( N \sum_{j=1}^N |v_{1j}|^4 \right)^{-1},$$

which quantifies the degree of homogeneity with which all considered “modes” (i.e. time lags) contribute to the leading eigenvector  $\mathbf{v}_1 = (v_{11}, \dots, v_{1N})$  (Plerou et al., 2002). We find that similar to the signature of  $\delta_{LVD}$ ,  $\sigma_1^2$  exhibits marked *peaks* at the timing of the six strongest magnetic storms within the considered observation period. In turn, the participation ratio behaves rather like  $\alpha_1$  in showing sharp minima at those time intervals ending with intense magnetic storms. Both observations taken together imply that both the largest eigenvalue and  $\delta_{LVD}$  (being clearly influenced by  $\sigma_1^2$ ) trace the presence of gradual trends associated with storms and their associated recovery phases (i.e. take their maxima/minima during time windows which fully cover such a large-scale pattern), whereas the associated eigenvector indicates the non-stationarity associated with the storm (which is most pronounced for time windows ending at the peak phase of the storm).

### 4.3.2 Linear vs. nonlinear statistical associations

As indicated by the interpretation of other dynamical characteristics (Sect. 4.2), magnetic storms are strongly nonlinear phenomena. In this regard, it is possible that some of the more subtle variations in the underlying nonlinear dynamics during magnetic storms are not sufficiently traced by linear autocorrelations yielding the foundation of the LVD dimension density method.

In order to test this hypothesis, one possible way is to study a “nonlinear” version of this measure, where the linear autocorrelation function is replaced by a nonlinear statistical association measure such as the (auto-)mutual information function. Similar approaches using different measures of dynamical similarity have been recently used for reducing the effect of non-Gaussianity in LVD dimension density estimates for small multivariate data sets (Donner, 2011) and studying phase synchronisation processes in networks of coupled oscillators (Donner, 2008). Figure 4 displays the corresponding results, which are qualitatively consistent with those of the “linear” LVD dimension density. Specifically, the main storm periods (note that the emergence of intense magnetic storms is typically clustered in time, which is consistent with the known persistence ( $H > 0.5$ ) of magnetic fluctuations during storm phases) are characterised by a marked reduction in temporal complexity, which appears, however, less marked than in the case of the linear version.

There are several reasons for the less clear signature of magnetic storms in the considered nonlinear dimension density.

First, the typically required amount of data (i.e. window width  $w$ ) for properly estimating nonlinear interdependence measures is much larger than in the case of the linear (Pearson) correlation function. In this spirit, the choice of  $w$  considered in most calculations presented here is most probably too low to guarantee proper mutual information estimates, especially at large delays, i.e. small differences  $w - N$  as studied here. Notably, these large delays most probably become particularly important in storm periods due to the presence of increased long-term memory. We emphasise that there are various competing estimators of mutual information, all of them sharing the problem of large variance (or even potential bias) at short time series lengths (Kraskov et al., 2004; Cellucci et al., 2005).

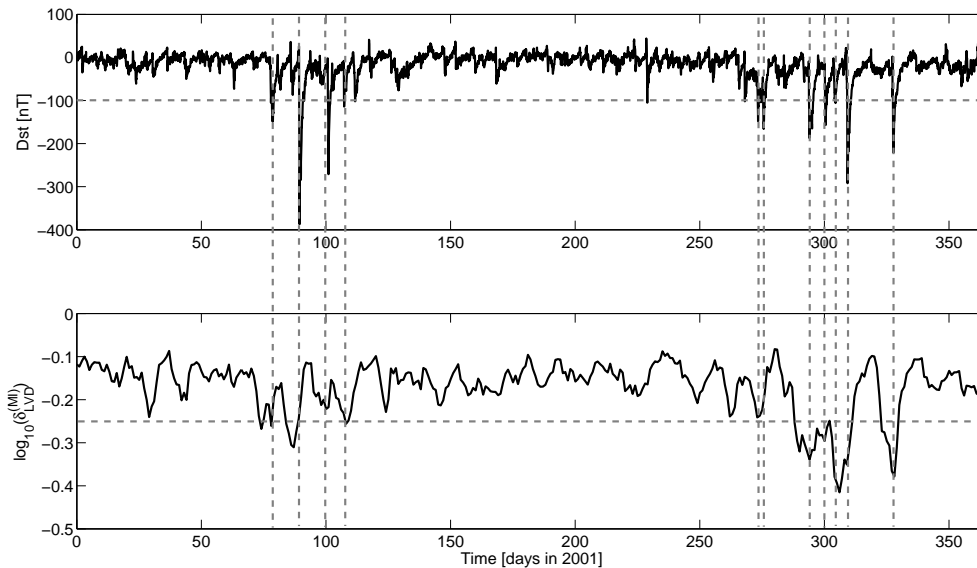
Second, the decay behaviour of the mutual information function is known to be different from that of the autocorrelation function. This is already a consequence of the fact that mutual information is positive semi-definite, whereas linear correlations are not. Consequently, the considered Toeplitz matrix has only non-negative coefficients, which leads to an essentially different spectrum of eigenvalues (i.e. there are no studies on the presence of an approximate exponential decay of the residual sum of eigenvalues, which is the basis of the applied methodology). We do not further explore this point here, since it is beyond the scope of the present work.

Finally, there is the conceptual problem that the eigenvalues of the mutual information matrix do not directly correspond to component variances of some statistical decomposition of the data (such as least-dependent component analysis based on mutual information minimisation, which would be the associated nonlinear counterpart of PCA (Stögbauer et al., 2004)). In this respect, the notion of “residual variances” is not well defined either.

In general, the applied “brute-force” combination of linear and nonlinear approaches has severe methodological drawbacks in comparison with the “classical” LVD dimension density method, requires consideration of larger time windows, and is computationally much more demanding. Given these facts, further research is necessary to provide a fully operational “nonlinear” variant of the LVD dimension density.

### 4.3.3 Beyond the exponential decay model

Finally, the LVD dimension density as discussed so far explicitly assumes the validity of an exponential decay model for the residual variances. In dependence on the specific dynamics, deviations from this idealised assumption may be of different magnitude. In this spirit, quantifying the goodness of fit of the corresponding model might provide additional relevant information taking some complementary aspects related to the system’s dynamical complexity into account.



**Fig. 4.** As in Fig. 1 for a “nonlinear” version of the LVD dimension density (obtained with same values of  $f$ ,  $w$  and  $N$ , but  $\Delta w = 24$  h) based on nearest-neighbour estimates of the (auto-)mutual correlation function introduced by Kraskov et al. (2004) (using  $k = 6$  neighbours).

Instead of performing an explicit goodness-of-fit evaluation, Fig. 5 allows for a general visual assessment by displaying the logarithm of the residual variances  $\rho_p$ . As can be seen, the overall behaviour can be reasonably well approximated by an exponential model in all cases. Notably, during storm periods, we have a higher  $\sigma_1^2$  (see above) and thus generally smaller values of  $\rho_p$  for all  $p$ . As it follows from Fig. 5, in such cases the decay of the residual variances is considerably faster than during “normal” periods, which is expressed by lower values of  $\delta_{LVD}$ . This observation confirms the already discussed linkage between  $\delta_{LVD}$  and  $\sigma_1^2$ , whereas there is no direct correspondence with other correlation-based statistics.

#### 4.4 Performance analysis

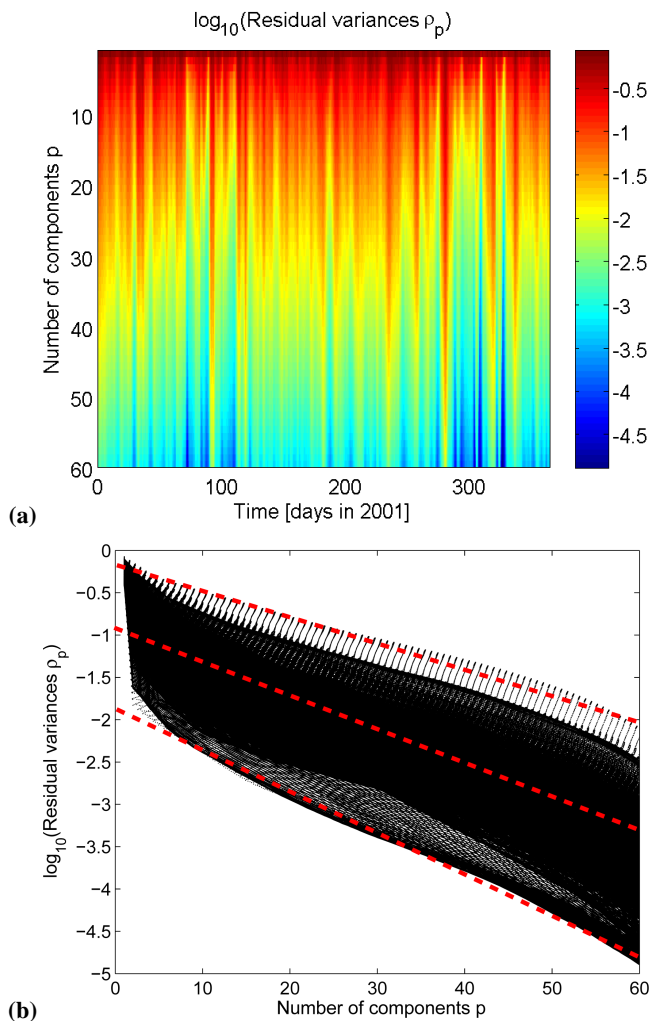
In order to evaluate the performance of  $\delta_{LVD}$  as a means to detecting signatures of intense magnetic storms, a receiver-operating characteristics (ROC) analysis has been performed. For this purpose, we first coarse-grained the Dst time series into periods of width  $\Delta T$  (in the following,  $\Delta T = 1$  day). Regarding the timing of intense magnetic storms, we define a *storm day* as coinciding with one of these periods if at least 50 % of the hourly Dst values within the given time window fall below a threshold Dst value ( $Dst^*$ ). For  $Dst^* = -100$  nT (i.e. the common definition of an intense magnetic storm), we thus find 13 storm days, which are considered as “events” to be traced by extended minima of  $\delta_{LVD}$ . In order to establish if the latter is indeed the case, we apply the same approach to our complexity measure, but with a varying threshold  $\delta_{LVD}^*$ . With increasing  $\delta_{LVD}^*$ , the fraction of correct “detections” of magnetic storms (true positive rate, TPR) necessarily converges towards 1, whereas the fraction of time

intervals without intense magnetic storm erroneously identified as storm periods (false positive rate, FPR) rises as well. Plotting these two fractions against each other yields the ROC curve of  $\delta_{LVD}$  as a means of characterising the performance of this measure.

Figure 6 displays the obtained ROC curves for the same parameters as used before but with different values of the embedding dimension  $N$ . It can be seen that the choice of  $N$  has only a minor influence on the resulting relationship between TPR and FPR, indicating a reasonable degree of robustness of our method. Moreover, since the ROC curves are not only clearly separated from the diagonal TPR = FPR (which would correspond to zero skills of our measure) but are also close to the limiting curve for a perfect discrimination of storm periods (i.e. FPR = 0 for TPR < 1 and TPR = 1 for FPR > 0), we can judge that  $\delta_{LVD}$  displays a close relationship with the underlying Dst index variability.

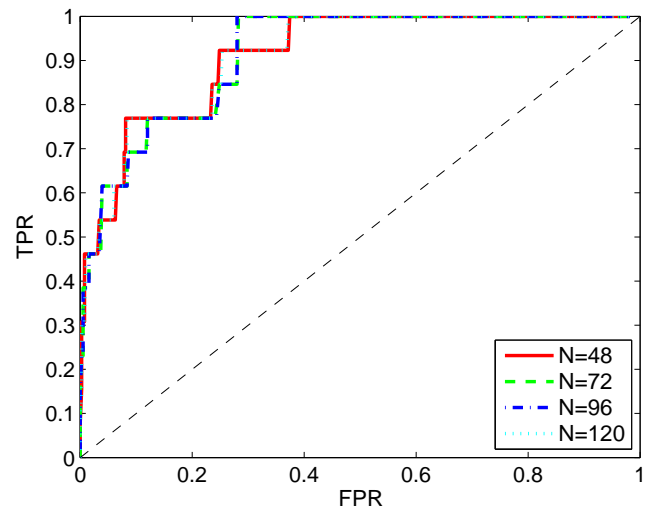
Quantifying the corresponding performance in more detail by computing the corresponding area under the ROC curve (AUC), we mostly find values above 0.9 for the discrimination of storm and non-storm periods by  $\delta_{LVD}$  (Fig. 7). Notably, the corresponding performance is much better than that obtained using the largest eigenvalue  $\sigma_1^2$  of the Toeplitz matrix of autocorrelations, which does not take AUC values of a similar order. Moreover, we find that for embedding dimensions  $N$  being either relatively low or large compared with the window width  $w$ , the performance is the best with about the same values of AUC. Specifically, we find a marked inflection in the symmetry around  $N = w/2$  for variable embedding dimension  $N$  and fixed window width  $w$ .





**Fig. 5.** (a) Residual variances  $\rho_p$  for running windows of the Dst time series (same parameters as before). (b) Detailed inspection of the approximately exponential decay behaviour of  $\rho_p$ . Dashed red lines visualise the typical scale of the decay, which is faster in the case of generally lower residual variances (the typical situation for intense magnetic storms) corresponding to lower values of  $\delta_{\text{LVD}}$ .

In turn, when fixing the embedding dimension  $N = 96$  and increasing the window width  $w$  (thus decreasing the ratio  $N/w$ ), we find the best discrimination for  $N/w \approx 0.7$ , i.e. window sizes of around six days. In this spirit, we conjecture that this value could represent an optimum window width for resolving the signatures of time-varying dynamical complexity in space weather covering all necessary timescales of variability yet allowing for a reasonable temporal resolution. We emphasise that this timescale is in excellent agreement with previous values for the nonlinear magnetospheric dependencies inferred by Johnson and Wing (2005), who explored the nonlinear behaviour of the magnetosphere as characterised by another geomagnetic activity index, i.e. the planetary 3 h range index,  $K_p$ , which is, however, less suited for studying magnetic storms. They demonstrated that strong nonlinear



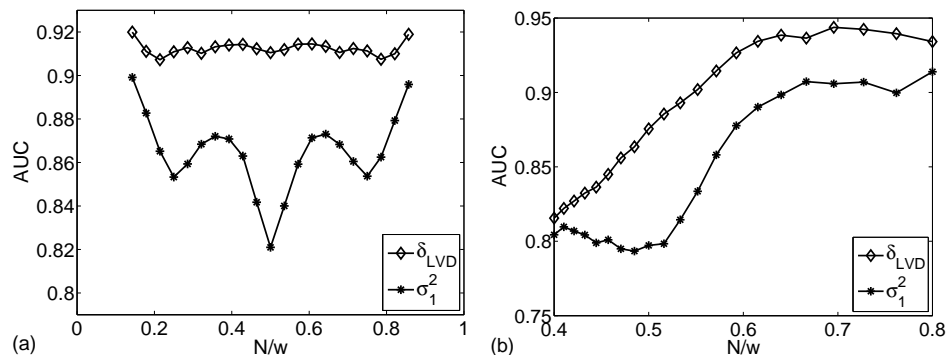
**Fig. 6.** ROC curve for  $\delta_{\text{LVD}}$  obtained using different embedding dimensions  $N$  ( $w = 168$  h,  $\Delta w = 1$  h,  $f = 0.95$ ). The dashed diagonal line displays the relation  $\text{TPR} = \text{FPR}$  corresponding to vanishing discriminatory skills of  $\delta_{\text{LVD}}$ .

magnetospheric dependencies tend to peak on a timescale around 40–50 h, and are statistically significant up to one week. This is in accordance with not only the time windows used earlier for the fractal spectral analysis and to derive Tsallis entropic measures of the Dst data but is also in agreement with the time windows used in this study to estimate the associated LVD dimension densities.

In order to fully explore the optimality of possible parameter choices, it would be necessary to consider AUC with respect to the two-dimensional  $(N, w)$  space, possibly also taking further parameters of the method as additional variables into account, such as  $f$ , the width  $\Delta T$  and phasing of the time intervals used for defining “events” for the ROC analysis, the threshold value  $\text{Dst}^*$  for defining a storm, etc. To this end, we only highlight the possibility of performing such a more detailed analysis, and leave this open for future research activities. In a similar spirit, it is possible to extend the presented analysis to other dynamical characteristics (e.g. entropies, Hurst exponents and even further properties). However, such a quantitative comparison between different methodological approaches is beyond the scope of the present work.

## 5 Conclusions

We have proposed a novel approach for characterising time-varying dynamical complexity in nonlinear geophysical time series based on scaling properties associated with the temporary autocorrelation structure. Our method is a straightforward extension of the established LVD dimension density  $\delta_{\text{LVD}}$  to univariate time series, which allows for characterising subtle changes in the system’s dynamical characteristics



**Fig. 7.** Areas under the ROC curve AUC for  $\delta_{LVD}$  (diamonds) and  $\sigma_1^2$  (asterisks) in dependence on  $N$  (a,  $w = 168$  h) and  $w$  (b,  $N = 96$ ). In all cases,  $\Delta w = 1$  h and  $f = 0.95$  have been used.

as time proceeds and thus appears widely applicable to data from a variety of different fields. Specifically, as we have shown,  $\delta_{LVD}$  is largely influenced by the dominating “mode” of variability, e.g. short-term fluctuations, temporary trends, or even possible oscillatory patterns. In the same vein, if the associated main pattern type is changed,  $\delta_{LVD}$  is expected to exhibit marked changes as well, which is supported by the present study. In order to avoid this behaviour and focus on only one specific timescale of variability, pre-filtering of the time series under study may present a reasonable methodological extension of our proposed approach. We will further explore this possibility in future research.

It is important to note that there is no way to find an “optimum” organisation or complexity measure (Kurths et al., 1995). Moreover, a combination of complexity measures, which refer to different aspects of a system, such as structural versus dynamical properties, is the most promising way to deal with the complexity of a system. Figure 2 provides a synthesis of the results of three conceptually different complexity measures to the same Dst index time series, where each measure represents a different approach to the study of magnetosphere dynamics related to magnetic storms. Specifically, the Hurst exponent, non-extensive Tsallis entropy and univariate normalised LVD dimension density all catch different aspects and features related to the complex character of the Earth’s magnetosphere as described in this study. In this spirit, although our new approach, which is utilised here for the first time in a sliding windows framework, gives results that are consistent with previously known findings, it provides a complementary view on magnetospheric complexity. Hence, we believe that  $\delta_{LVD}$  does not just provide an interesting alternative for characterising temporally varying complexity of magnetospheric fluctuations but could also be of interest in various other fields of application as an easily computable complexity measure.

We have successfully applied  $\delta_{LVD}$  as a means to characterising time-varying dynamical complexity in the Earth’s magnetosphere associated with the emergence of magnetic storm periods. Unlike other measures like the Hurst exponent

or non-extensive Tsallis entropy, our characteristic is able to detect signatures directly associated with a magnetic storm event itself rather than changes in the dynamical properties in the preparatory phase. Along the same line,  $\delta_{LVD}$  as considered in this work appears unable to detect precursory structures associated with strong magnetospheric perturbations. In turn, it allows for classifying the dynamical signatures of such activity phases a posteriori. Specifically, a  $\delta_{LVD}$  value of 0.05 that has been used in Fig. 2 yields a reasonable boundary marking the different activity epochs of the quiet-time and storm-time magnetosphere. For the latter purpose, a detailed ROC analysis has revealed acceptable skills in discriminating between the dynamical characteristics during intense magnetic storms and periods of magnetospheric quiescence. The reduction in complexity with time could be related to the emergence of strong anisotropy in the system (recall that anisotropy is inherent to the system). The action of such strong anisotropy can lead to the appearance of a clearly preferred direction of elementary activities (ion acceleration and earthward transport) before the main phase of the storm, i.e. to the massive and continuous earthward injection of accelerated ions. Theoretical and experimental evidence support the former hypothesis: throughout the entire main phase and during the early recovery phase of intense magnetic storms, the geometry of the energy flow produces a highly asymmetric ring current configuration (Daglis and Kozyra, 2002; Daglis et al., 2003).

We emphasise that the application of the new methodological approach presented in this work to different geomagnetic activity indices like the 1 min SYM-H index could, in the future, provide novel insights into the study of magnetospheric dynamics. We plan to extend the present study accordingly in the near future. In turn, as another possible extension of the work presented here, we note that the used framework for performance analysis of  $\delta_{LVD}$  is very general and could also be applied in future research to quantitatively evaluating and comparing the performance of other recently studied measures associated with non-stationarities in magnetospheric dynamics.

*Acknowledgements.* This work has been financially supported by the joint Greek–German IKYDA 2013 project “Transdisciplinary assessment of dynamical complexity in magnetosphere and climate: A unified description of the nonlinear dynamics across extreme events” funded by IKY and DAAD. G. Balasis additionally acknowledges partial support from the “Hellenic National Space Weather Research Network” co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF) – Research Funding Program “Thales. Investing in knowledge society through the European Social Fund”. The Dst data are provided by the World Data Center for Geomagnetism, Kyoto (<http://swdcwww.kugi.kyoto-u.ac.jp/>). Valuable discussions with K. Eftaxias and C. Papadimitriou are gratefully acknowledged.

The service charges for this open access publication have been covered by the Max Planck Society.

Edited by: G. Lapenta

Reviewed by: two anonymous referees

## References

- Angelopoulos, V., Mukai, T., and Kokubun, S.: Evidence for intermittency in Earth’s plasma sheet and implications for self-organized criticality, *Phys. Plasmas*, 6, 4161–4168, doi:10.1063/1.873681, 1999.
- Baker, D. N., Klimas, A. J., McPherron, R. L., and Büchner, J.: The evolution from weak to strong geomagnetic activity: An interpretation in terms of deterministic chaos, *Geophys. Res. Lett.*, 17, 41–44, doi:10.1029/GL017i001p00041, 1990.
- Balasis, G., Daglis, I. A., Kapiris, P., Manda, M., Vassiliadis, D., and Eftaxias, K.: From pre-storm activity to magnetic storms: a transition described in terms of fractal dynamics, *Ann. Geophys.*, 24, 3557–3567, doi:10.5194/angeo-24-3557-2006, 2006.
- Balasis, G., Daglis, I. A., Papadimitriou, C., Kalimeri, M., Anastasiadis, A., and Eftaxias, K.: Dynamical complexity in Dst time series using non-extensive Tsallis entropy, *Geophys. Res. Lett.*, 35, L14102, doi:10.1029/2008GL034743, 2008.
- Balasis, G., Daglis, I. A., Papadimitriou, C., Kalimeri, M., Anastasiadis, A., and Eftaxias, K.: Investigating dynamical complexity in the magnetosphere using various entropy measures, *J. Geophys. Res.*, 114, A00D06, doi:10.1029/2008JA014035, 2009.
- Balasis, G., Daglis, I. A., Anastasiadis, A., and Eftaxias, K.: Detection of dynamical complexity changes in Dst time series using entropy concepts and rescaled range analysis, in: *The Dynamic Magnetosphere*, edited by: Liu, W. and Fujimoto, M., Springer, Berlin Heidelberg, 211–220, doi:10.1007/978-94-007-0501-2\_12, 2011a.
- Balasis, G., Daglis, I. A., Papadimitriou, C., Anastasiadis, A., Sandberg, I., and Eftaxias, K.: Quantifying dynamical complexity of magnetic storms and solar flares via nonextensive Tsallis entropy, *Entropy*, 13, 1865–1881, doi:10.3390/e13101865, 2011b.
- Balasis, G., Papadimitriou, C., Daglis, I. A., Anastasiadis, A., Athanasopoulou, L., and Eftaxias, K.: Signatures of discrete scale invariance in Dst time series, *Geophys. Res. Lett.*, 38, L13103, doi:10.1029/2011GL048019, 2011c.
- Balasis, G., Papadimitriou, C., Daglis, I. A., Anastasiadis, A., Sandberg, I., and Eftaxias, K.: Similarities between extreme events in the solar-terrestrial system by means of nonextensivity, *Nonlin. Processes Geophys.*, 18, 563–572, doi:10.5194/npg-18-563-2011, 2011d.
- Cellucci, C. J., Albano, A. M., and Rapp, P. E.: Statistical validation of mutual information calculations: Comparison of alternative numerical algorithms, *Phys. Rev. E*, 71, 066208, doi:10.1103/PhysRevE.71.066208, 2005.
- Chang, T., Wu, C. C., Podesta, J., Echim, M., Lamy, H., and Tam, S. W. Y.: ROMA (Rank-Ordered Multifractal Analyses) of intermittency in space plasmas – a brief tutorial review, *Nonlin. Processes Geophys.*, 17, 545–551, doi:10.5194/npg-17-545-2010, 2010.
- Daglis, I. A.: Ring current dynamics, *Space Sci. Rev.*, 124, 183–202, doi:10.1007/s11214-006-9104-z, 2006.
- Daglis, I. A. and Kozyra, J. U.: Outstanding issues of ring current dynamics, *J. Atmos. Solar-Terr. Phys.*, 64, 253–264, 2002.
- Daglis, I. A., Kozyra, J. U., Kamide, Y., Vassiliadis, D., Sharma, A. S., Liemohn, M. W., Gonzalez, W. D., Tsurutani, B. T., and Lu, G.: Intense space storms: Critical issues and open disputes, *J. Geophys. Res.*, 108, 1208, doi:10.1029/2002JA009722, 2003.
- De Michelis, P., Consolini, G., and Tozzi, R.: On the multi-scale nature of large geomagnetic storms: an empirical mode decomposition analysis, *Nonlin. Processes Geophys.*, 19, 667–673, doi:10.5194/npg-19-667-2012, 2012.
- Donner, R.: *Advanced Methods for Analysing and Modelling of Multivariate Palaeoclimatic Time Series*, Ph.D. thesis, University of Potsdam, Germany, 2007.
- Donner, R.: Multivariate analysis of spatially heterogeneous phase synchronisation in complex systems: application to self-organised control of material flows in networks, *Eur. Phys. J. B*, 63, 349–361, doi:10.1140/epjb/e2008-00151-8, 2008.
- Donner, R.: Spatial Correlations of River Runoffs in a Catchment, in: *In Extremis*, edited by: Kropp, J. and Schellnhuber, H.-J., Springer Berlin Heidelberg, 286–313, doi:10.1007/978-3-642-14863-7\_14, 2011.
- Donner, R. and Witt, A.: Characterisation of long-term climate change by dimension estimates of multivariate palaeoclimatic proxy data, *Nonlin. Processes Geophys.*, 13, 485–497, doi:10.5194/npg-13-485-2006, 2006.
- Donner, R. V.: Complexity Concepts and Non-Integer Dimensions in Climate and Paleoclimate Research, in: *Fractal Analysis and Chaos in Geosciences*, edited by: Ouadfeul, S.-A., InTech, Rijeka, 28 pp., doi:10.5772/53559, 2012.
- Donner, R. V., Sakamoto, T., and Tanizuka, N.: Complexity of spatio-temporal correlations in Japanese air temperature records, in: *Nonlinear Time Series Analysis in the Geosciences: Applications in Climatology, Geodynamics and Solar-Terrestrial Physics*, edited by: Donner, R. V. and Barbosa, S. M., Springer, Berlin Heidelberg, 124–154, doi:10.1007/978-3-540-78938-3\_7, 2008.
- Gonzalez, W. D., Joselyn, J. A., Kamide, Y., Kroehl, H. W., Rostoker, G., Tsurutani, B. T., and Vasylunas, V. M.: What is a geomagnetic storm?, *J. Geophys. Res.*, 99, 5771–5792, doi:10.1029/93JA02867, 1994.
- Hurst, H. E.: Long term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, 116, 770–799, 1951.

- Johnson, J. R. and Wing, S.: A solar cycle dependence of nonlinearity in magnetospheric activity, *J. Geophys. Res.*, 110, A04211, doi:10.1029/2004JA010638, 2005.
- Klimas, A. J., Vassiliadis, D., Baker, D. N., and Roberts, D. A.: The organized nonlinear dynamics of the magnetosphere, *J. Geophys. Res.*, 101, 13089–13113, doi:10.1029/96JA00563, 1996.
- Kraskov, A., Stögbauer, H., and Grassberger, P.: Estimating mutual information, *Phys. Rev. E*, 69, 066138, doi:10.1103/PhysRevE.69.066138, 2004.
- Kurths, J., Voss, A., Saparin, P., Witt, A., Kleiner, H., and Wessel, N.: Quantitative analysis of heart rate variability, *Chaos*, 5, 88–94, doi:10.1063/1.166090, 1995.
- Pavlos, G. P., Athanasiu, M. A., Rigas, A. G., Sarafopoulos, D. V., and Sarris, E. T.: Geometrical characteristics of magnetospheric energetic ion time series: evidence for low dimensional chaos, *Ann. Geophys.*, 21, 1975–1993, doi:10.5194/angeo-21-1975-2003, 2003.
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T., and Stanley, H. E.: Random matrix approach to cross correlations in financial data, *Phys. Rev. E*, 65, 066126, doi:10.1103/PhysRevE.65.066126, 2002.
- Shannon, C. E.: A mathematical theory of communication, *Bell Syst. Techn. J.*, 27, 379–423, 1948.
- Sharma, A. S., Vassiliadis, D. V., and Papadopoulos, K.: Reconstruction of low-dimensional magnetospheric dynamics by singular spectrum analysis, *Geophys. Res. Lett.*, 20, 355–358, doi:10.1029/93GL00242, 1993.
- Stögbauer, H., Kraskov, A., Astakhov, S. A., and Grassberger, P.: Least-dependent-component analysis based on mutual information, *Phys. Rev. E*, 70, 066123, doi:10.1103/PhysRevE.70.066123, 2004.
- Toonen, C., Lappe, D., Donner, R., and Scholz-Reiter, B.: Impact of Machine-Driven Capacity Constellations on Performance and Dynamics of Job-Shop Systems, in: *Enabling Manufacturing Competitiveness and Economic Sustainability*, edited by ElMaraghy, H. A., Springer Berlin Heidelberg, 611–616, doi:10.1007/978-3-642-23860-4\_100, 2012.
- Tsallis, C.: Possible generalization of Boltzmann–Gibbs statistics, *J. Stat. Phys.*, 52, 479–487, doi:10.1007/BF01016429, 1988.
- Tsallis, C.: The nonadditive entropy  $S_q$  and its applications in physics and elsewhere: Some remarks, *Entropy*, 13, 1765–1804, doi:10.3390/e13101765, 2011.
- 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  $B_5$  driver: A spectral break at 5 hours, *Geophys. Res. Lett.*, 17, 279–282, doi:10.1029/GL017i003p00279, 1990.
- Vassiliadis, D. V., Sharma, A. S., Eastman, T. E., and Papadopoulos, K.: Low-dimensional chaos in magnetospheric activity from AE time series, *Geophys. Res. Lett.*, 17, 1841–1844, doi:10.1029/GL017i011p01841, 1990.
- Vörös, Z., Baumjohann, W., Nakamura, R., Runov, A., Zhang, T. L., Volwerk, M., Eichelberger, H. U., Balogh, A., Horbury, T. S., Glaßmeier, K.-H., Klecker, B., and Rème, H.: Multi-scale magnetic field intermittence in the plasma sheet, *Ann. Geophys.*, 21, 1955–1964, doi:10.5194/angeo-21-1955-2003, 2003.
- Wanliss, J. A.: Fractal properties of SYM-H during quiet and active times, *J. Geophys. Res.*, 110, A03202, doi:10.1029/2004JA010544, 2005.
- Xie, X., Zhao, X., Fang, Y., Cao, Z., and He, G.: Normalized linear variance decay dimension density and its application of dynamical complexity detection in physiological (fMRI) time series, *Phys. Lett. A*, 375, 1789–1795, doi:10.1016/j.physleta.2011.03.003, 2011.