Extreme events and long-range correlations in space weather

A. S. Sharma¹ and T. Veeramani¹,*

¹Department of Astronomy, University of Maryland, College Park, Maryland, 20742, USA
*now at: Department of Electrical Engineering, Indian Institute of Technology Madras, Chennai, India

Received: 29 November 2010 – Revised: 30 September 2011 – Accepted: 7 October 2011 – Published: 20 October 2011

Abstract. Space weather is driven by the solar wind and many geospace storms and substorms are natural hazards with considerable societal impact. The dynamical and statistical features of these events are complicated because of the turbulent nature of their driver, the solar wind. Large-scale data sets of geospace storms and substorms are analysed for this study of the inherent statistical characteristics of extreme events in geospace. The detrended fluctuation analysis, based on the autocorrelation functions, is used and yields scaling behavior representing long-term correlations. The scaling function is represented by two exponents, arising due mainly to the presence of the largely coherent internal dynamics of the magnetosphere and the turbulent nature of the solar wind driver.

1 Introduction

The inherent dynamical and statistical properties of complex phenomena in geosciences are critical to the understanding of extreme events, in particular those leading to natural hazards. Many complex driven systems such as the coupled solar wind – magnetosphere system, are far from equilibrium and the commonly used techniques of statistical analysis can not be applied readily, and the nonlinear dynamics and complexity science provide a natural framework for the study of such systems, in particular in the study of the magnetosphere and space weather (Sharma, 1995; Klimas et al., 1996; Consolini and Chang, 2001; Zelenyi and Milovanov, 2004). The importance of this approach arises from the recognition that dynamical behaviour, including extreme events, is not isolable phenomena but must be understood in terms of interactions among different components, within and without the specific system.

It should be noted that extreme events are of both natural or anthropogenic origin, and are ubiquitous mainly because of their damaging consequences. However, there is no single definition, at least in the scientific sense, of extreme events (Jentsch et al., 2006). The interpretation of the degree of extremeness often involves the attributes of infrequent occurrence, low-probability or unexpected nature, strong impact, etc. In general, it is not clear that extreme events can be characterized by one or even a few measures. However, it is clear that extreme events are rare and in the distribution of events of all magnitudes they are identified as those outside the bulk, viz. the tail of the distribution. A main objective in the analysis of extreme events thus relates directly to the understanding of the distribution function of the events, in particular the outliers. Another feature of extreme events is that they occur suddenly and a well known characteristic of sudden transitions, such as phase transitions, is the emergence of long-range order, i.e. the value of a physical variable at an arbitrary point is correlated with its value at a point located far away (Dixon et al., 1997). Thus, long-range correlations are important indicators of the development of extreme events. In view of these features the dynamical and statistical approaches of complexity science provide a natural framework for the study of extreme events (Sharma et al., 2010).

The dynamical modeling and prediction based on the reconstruction of dynamics from observational time series data has been used extensively in many natural and laboratory systems (Abarbanel et al., 1993; Kantz and Schrieber, 1997). This approach, based on the embedding theorem, has enabled the reconstruction of dynamical models from observational data, independent of modeling assumptions. In the studies of the dynamics of the geospace environment this approach has provided the first predictive models of geomagnetic activity and space weather, enabled by the extensive data from ground-based and space-borne instruments. These studies, focused on the dynamical behavior, led to the earliest space weather forecasting tools (Sharma, 1995).
The nonequilibrium nature of large scale open systems limits the predictive capability of the dynamical models. In particular, for extreme events the statistical properties are therefore essential for deriving important properties such as the probabilities of recurrence. Recent developments in the studies using the data of many natural phenomena such as floods, climate, earthquakes, etc. have shown long-range interactions to be an inherent feature (Bunde et al., 2004; Altman and Kantz, 2005). The long-range correlations in climate data is identified as leading to many features such as the clustering of extreme events (Bunde et al., 2005) and studies of the data sets of other phenomena are needed to understand the nature of extreme events in general.

In this paper we use the detrended fluctuation analysis to study the nature of long-range correlations in the coupled solar wind – magnetosphere system. In the next section the essential features of space weather and the relevant geospace data are described. The detrended fluctuation analysis with autocorrelation functions computed from large scale data sets of geospace are described in Sect. 3. The main results of the paper are summarized in Sect. 4.

2 Extreme events in space weather

The extreme events in space weather occur during the periods when the magnetosphere is strongly driven by the solar wind, which brings the energetic plasma and fields from the solar eruptive events such as coronal mass ejections to geospace. Many extreme space weather events in the recent past have caused serious damages to technological systems such as satellites, power transmission systems, etc. Some well known examples are: the collapse of Hydro Quebec power grid during the great geomagnetic storm of March 1989, the Canadian telecommunication satellite outage during a period of enhanced energetic electron fluxes at geosynchronous orbit in January 1994, the electrical breakdowns and satellite malfunctions during the magnetic cloud event of July 2000 (Bastille Day event), the disabling of GPS based aviation system during the severe space weather events of October–November 2003 (Halloween Storms), the disturbances in commercial airline traffic during several days of enhanced geomagnetic activity in January 2005, etc. (NRC, 2008). Although these events may not seem devastating by themselves, a confluence of natural hazards in the different regions of the environment of the Earth can make our society and its technological systems highly vulnerable because of their interconnectedness (Baker and Allen, 2000; NRC, 2008). In this aspect the nonlinear dynamical framework for the study of the extreme events become directly relevant to the extended Earth and space system.

The modeling of space weather events rely strongly on the availability of good geospace data and among the most widely used data are the geomagnetic indices (Mayaud, 1980). The data from ground magnetometer stations around the globe have been monitored for more than one and half centuries and these data have been used to compute the geomagnetic indices (Mayaud, 1980; Love, 2008). Among the many indices the auroral electrojet indices (AE, AL and AU) characterize the substorms, and the ring current index Dst represents the geomagnetic or space storms. The substorms, with a characteristic time of ~ 1 h, are episodic in nature and are the essential elements of magnetospheric dynamics. The auroral electrojet indices provide the detailed dynamical features of the global aspects of substorms. On the other hand, the geomagnetic storms, with a typical time scale of ~10 h, are the more global space weather disturbances during which intense substorms occur.

The auroral electrojet indices are computed from the horizontal component of the magnetic field disturbances at a dozen or so ground magnetometer stations distributed around the globe and are readily available with 1 min or longer resolution. These indices reflect the strengths of the large scale ionospheric currents driven by the reconfiguration of the magnetosphere during substorms. They are highly variable during strongly disturbed periods, with peak values of 1000–2000 nT during extreme events cited earlier. The substorms with AL index values less than −1000 nT are considered strong disturbances and these will be considered as extreme events for this study. The geomagnetic storms with Dst values less than −100 nT are referred to as intense storms (Gonzalez et al., 1994). The substorms, with typical time scales of an hour, occur during the storms, with time scales of 10 h or longer. Although the big substorms are accompanied by intense storms the relationships between storms and substorms are not fully resolved (Kamide et al., 1998; Daglis et al., 2003; Sharma et al., 2003). In this study the AE/AL data will be used for the analysis of extreme events in space weather.

The AE index at 1 min resolution for a highly disturbed period, viz. January 1983, is shown in Fig. 1. It should be noted that AE (= AU−AL) has positive values and tracks the AL values closely since AU values are usually not large. The episodic and high variability of the substorms are evident in the sharp peaks of AE whose distribution reflect the nonequilibrium nature of the phenomenon. In this case there are many substorms with AE values above 1000 nT and the corresponding Dst values were close to −100 nT. As is the case with extreme events in general, there is no single measure of the extreme events in space weather. For example, the Dst for the well known “Carrington” storm of 1–2 September 1859 is estimated to be −1760 nT (Tsurutani et al., 2003), and its effects were felt across the globe. The more recent Bastille Day event of 14–16 July 2000 with a Dst minimum of −300 nT (http://wdc.kugi.kyoto-u.ac.jp/dst_final/200007/index.html) was an extreme space weather event which led to significant damages to satellites and other technological infrastructure. It should be emphasized here that the main objective in the studies of extreme events is the nature of their distribution.
In general a system has a time characteristic is shown in Fig. 1. For the AE/AL data (Chen et al., 2011; Abarbanel et al., 1993) is shown in Fig. 2. The characteristic time associated with the average mutual information function is usually taken as the delay time corresponding to half the peak value and this yields ∼10 min. This value can be used as the time delay parameter in many studies such as the reconstruction of magnetospheric dynamics (Sharma et al., 2013; Chen et al., 2008). In general a system has a time scale characterizing the inherent correlations and the mutual information function shows that for the magnetosphere this basic time scale is ∼10 min. For the magnetosphere the Lyapunov exponent computed from the the AL time series is also ∼10 min (Vassiliadis et al., 1991). Thus the long-range correlations for the magnetosphere thus would emerge on time scales much longer than ∼10 min.

Both the autocorrelation function $C(\tau)$ and mutual information function $I(\tau)$ reflect the inherent correlations and can be used to derive other physical quantities. In the studies of long-range correlations the auto-correlation function is the most widely used function, e.g. Bunde et al. (2005), and the following analysis of the long-range correlations is based on this function.

### 3 Detrended fluctuation analysis of AL index

The long-range correlations in a system are analyzed using the scaling behavior of correlation functions. However this requires careful analysis as trends in the data need to be eliminated first so that the long-range correlations as genuine features can be determined. The trends in data are usually caused by external effects, viz. they can be due to the

**Fig. 1.** The auroral electrojet index AE for January 1983 at 1 min resolution. The sharp spikes represent substorms and the strong substorms usually occur during space storms.

**Fig. 2.** The mutual information function of the auroral electrojet index AE for January 1983 at 1 min resolution (Fig. 1). The characteristic time, corresponding to half the peak value, is ∼10 min.
Fig. 3. Autocorrelation function of long-range correlated data (Makse et al., 1996). The correlations are weaker for higher values of the exponent (0.2, 0.4, 0.6 and 0.8).

Fig. 4. Average mutual information function $I(\tau)$ of long-range correlated data for different values of the exponent (Makse et al., 1996). The higher the values of the exponent, the weaker are the correlations.

driver of the system. Thus, in the case of the solar wind-magnetosphere system, the long term trends in the solar wind can potentially show features in the magnetosphere resembling intrinsic long-range correlations. Among the techniques for removing trends in the data, the detrended fluctuation analysis (Peng et al., 1994; Kantelhardt et al., 2001; Gao et al., 2006) is widely used.

Recent advances in the studies of extreme events using the detrended fluctuation analysis have shown the role of long-term memory in the development of extreme events. For example, when the memory function is represented by the autocorrelation function that decays algebraically with an exponent, the probability density function of the return intervals between events become a stretched exponential characterized by the same exponent as the autocorrelation function (Bunde et al., 2005). In the case of uncorrelated data the distribution vanishes for $\tau > 0$. Usually it decays exponentially with a characteristic time $\tau_c$ as $C(\tau) \sim \exp(-\tau/\tau_c)$. In this case a plot of $\ln C(\tau)$ vs. $\tau$ will show a linear dependence. In the presence of long-range correlations $C(\tau)$ decays as a power law, viz. $C(\tau) \sim \tau^{-\gamma}$, with a linear dependence in a $\ln C(\tau)$ vs. $\ln(\tau)$ plot.

The nature of auto-correlation functions in a long-term correlated data can be examined by using the data generated by modified Fourier filtering of white noise (Makse et al., 1996). To obtain such a data-set a sequence of random numbers is generated and then its Fourier components filtered through power law filters. The auto-correlation functions for the data generated with different values of the power law exponent exponent $\gamma$ are shown in Fig. 3. The power law behavior expected in the $\ln C(\tau)$ vs. $\ln(\tau)$ plots for the long-range correlation are clearly seen for $\gamma = 0.2, 0.4, 0.6$ and 0.8. The dependence of the values of the correlation function for different exponents is evident in this figure.

The mutual information function $I(\tau)$, defined by Eq. (2) are computed for the same data set and are shown in Fig. 4. In this case again the long-term correlation structure is clearly depicted by $I(\tau)$ for all values of the exponent $\gamma$. This function however, unlike the autocorrelation function, encompasses all nonlinearities, and provides smoother variations over a wider range of scales.
4 Long-range correlations in AL index data

The large database of the auroral electrojet index AL provides a suitable data set for the studies of long-range correlations in the magnetosphere and space weather. In order to analyze the details of the correlations in AL data different data sets with different resolutions are used. The first case is the hourly averaged AL for the period 1978–1988, which covers a typical solar cycle of 11 yr.

The detrended fluctuation analysis of the AL time series data is accomplished in four steps (Kantelhardt et al., 2001). The first step computes the profile of the data set as:

\[ Y(i) = \sum_{k=1}^{i} x_k - \langle x \rangle \]  

(3)

The subtraction of the global mean \( \langle x \rangle \) of the dataset however is not essential as the third step, described below, removes this and other trends. In the second step the profile \( Y(i) \) is divided into \( N_L = N/L \) non-overlapping segments of length \( L \). In order to avoid a loss of data in the case \( N \) is not a multiple of \( L \), the same process is repeated starting from the other end of the data set, yielding \( 2N_L \) segments. The third step is where the trends in the data are removed by defining a local trend \( q_j(i) \) for each segment \( j \) by a fitting procedure, e.g. a least-squares fit. The detrended time series for the segment duration \( L \) is then defined as:

\[ Y_L(i) = Y(i) - q_j(i) \]  

(4)

The local trend is usually represented by a polynomial and in this study a quadratic function is used, and thus corresponds to DFA2 of Kantelhardt et al. (2001). In the fourth step the variance of each segment \( Y_L(i) \) is calculated:

\[ F^2_L(j) = \langle Y^2_L(i) \rangle = \frac{1}{L} \sum_{i=1}^{L} Y^2_L[(j-1)L + i] \]  

(5)

The detrended fluctuation function \( F(L) \) is then obtained as

\[ F^2(L) = \frac{1}{N_L} \sum_{j=1}^{2N_L} F^2_L \]  

(6)

If the original data are long–range correlated the fluctuation function is expected to have a scaling as

\[ F(L) \propto L^\alpha \]  

(7)

For uncorrelated or short-range correlated data, the exponent is 0.5 and larger values show the presence of long-range correlations (Kantelhardt et al., 2001).

The detrended fluctuation analysis of the 1 h averaged AL data for 1978–1988 yields a scaling function \( F(L) \) shown in Fig. 5. Also shown in this figure is the function using the fluctuation analysis (FA) following Peng et al. (1994) and Karche and Brendel (1993). The DFA function \( F(L) \) yields an exponent \( \approx 0.87 \), thus showing long-range correlations. It should be noted that this data is hourly averaged and as noted earlier, the substorms last typically an hour and higher resolution data are required to confirm this result. The detrended fluctuation analysis using 5 min averaged data yield a similar picture, as shown in Fig. 6. The exponent in this case is 0.90, very close to the case of 1-h averaged data. Thus the scaling of the \( F(L) \) in both the 1 h and 5 min averaged datasets show the presence of long-range correlations in the AL data,
and consequently in the magnetospheric dynamics and space weather. In both the cases there is a break in the function \( F(L) \) at 200–300 min and similar results are obtained in the case of other data sets of AL for different periods. More detailed studies are needed to reach a clear result on the nature of this break in the slope of the fluctuation function \( F(L) \).

## 5 Conclusions

Among the different components of space weather, the magnetosphere plays a critical role. The strong driving by the solar wind is the origin of many extreme geomagnetic events such as geospace storms and substorms, which are potential natural hazards. A database of substorms consisting of more than 5 million events was compiled for this study of the inherent statistical characteristics of extreme events in geospace. The detrended fluctuation analysis with autocorrelation functions is used to obtain the scaling exponents and they show the presence of long-range correlations, but with a break in the scaling. The existence of two exponents due to the break in the scaling of \( F(L) \) is likely to be due to the turbulent nature of the solar wind.

The autocorrelation function, used in the detrended fluctuation analysis, are widely used as the estimate of the correlations in the data. However for most complex systems the linear and nonlinear correlations are essential for the determination of their inherent statistical characteristics. The mutual information function, discussed in Sect. 2, is an information theoretic measure and has the important feature that it encompasses all orders of correlations. This function is thus better suited for the study of long-range correlations and related properties (Sharma and Veeramani, 2011).

**Acknowledgements.** The research was supported by NSF DMS grant 0417800.

Edited by: J. Büchner
Reviewed by: two anonymous referees

## References


www.nonlin-processes-geophys.net/18/719/2011/