www.nonlin-processes-geophys.net/19/667/2012/
© Author(s) 2012. CC Attribution 3.0 License.

On the multi-scale nature of large geomagnetic storms: an empirical mode decomposition analysis

P. De Michelis1,2, G. Consolini3, and R. Tozzi1
1Istituto Nazionale di Geofisica e Vulcanologia, 00143 Roma, Italy
2Dip. Scienze della Terra, Università degli Studi di Siena, 53100 Siena, Italy
3INAF – Istituto di Astrofisica e Planetologia Spaziali, 00133 Roma, Italy

Correspondence to: P. De Michelis (paola.demichelis@ingv.it)

Received: 30 June 2012 – Revised: 11 November 2012 – Accepted: 12 November 2012 – Published: 29 November 2012

Abstract. Complexity and multi-scale are very common properties of several geomagnetic time series. On the other hand, it is amply demonstrated that scaling properties of geomagnetic time series show significant changes depending on the geomagnetic activity level. Here, we study the multi-scale features of some large geomagnetic storms by applying the empirical mode decomposition technique. This method, which is alternative to traditional data analysis and is designed specifically for analyzing nonlinear and nonstationary data, is applied to long time series of Sym-H index relative to periods including large geomagnetic disturbances. The spectral and scaling features of the intrinsic mode functions (IMFs) into which Sym-H time series can be decomposed, as well as those of the Sym-H time series itself, are studied considering different geomagnetic activity levels. The results suggest an increase of dynamical complexity and multi-scale properties for intermediate geomagnetic activity levels.

1 Introduction

The Earth’s magnetosphere nonequilibrium dynamics in response to the solar wind changes is mainly nonlinear and multi-scale (e.g. Consolini et al., 2008). Evidences of such nonlinear nature are the dynamical complexity, observed in the overall and global response, and the turbulent features of the magnetic field and plasma parameter fluctuations observed in several magnetospheric regions, such as the high latitude polar regions, the magnetospheric equatorial plasma sheet, etc. (e.g. Consolini and Chang, 2001; Chang et al., 2003; Vassiliadis, 2006; Zimbardo et al., 2010). On the other hand, the multi-scale nature of the magnetospheric dynamics, which manifests in the absence of a single characteristic spatial and/or temporal scale in response to solar wind changes, is also widely provided by the scale-invariance of geomagnetic and magnetospheric observations (global and/or in situ time series of magnetic field and plasma parameters measurements). The occurrence of dynamical complexity and turbulent fluctuations has firmly been established in the case of geomagnetic substorms and storms and in situ physical processes related to these phenomena.

In this framework, geomagnetic storms are interesting and intriguing dynamical processes generated from the magnetospheric dynamics and whose effects are clearly visible in the magnetic records of ground-based observatories. The global scale geomagnetic disturbances observed during geomagnetic storms are the result of electric currents encircling the Earth and of the overall increase of magnetospheric plasma convection (e.g. Kamide and Maltsev, 2007). Indeed, during a magnetic storm the number of electrons and ions, which drift azimuthally round the Earth at radial distances of about 4–6 R_E, significantly enhances (e.g. Gonzalez et al., 1994; De Michelis et al., 1999). As a consequence of this increase of the intensity of magnetospheric electric currents, a reduction of the strength of the magnetic field at the Earth’s surface is produced. In first approximation, the strength of this perturbation on the Earth’s surface is given by the so-called Dessler-Parker-Sckopke relationship, which relates the energy of the charged particles trapped in the Earth’s magnetic field to the energy of the magnetic field itself and thus to the Dst geomagnetic index (Dessler and Parker, 1959; Sckopke, 1966). This index was originally designed to describe the variations of the symmetric ring current, but taking into account the specific Dst derivation procedure, it is evident that
Empirical mode decomposition: a brief introduction

In most real systems, either natural or even man-made ones, data describing their dynamics are often characterized by an inherent degree of nonlinearity and nonstationarity. For this reason, the analysis and the description of these systems require the use of different analytical methods based on adaptive bases, directly derived by the data themselves and capable of representing their inherent multi-scale and complex nature. Indeed, an a priori defined function cannot be used to build such a basis, no matter how much sophisticated the basis function might be. A few adaptive methods are available for signal analysis, among them being EMD (Huang et al., 1998). In contrast to almost all of the previous methods, the EMD method is intuitive, direct, and adaptive, with a basis a posteriori defined from the decomposition method, based on and derived from the data.

At the base of EMD is the idea that any time series can be written as the superposition of a number of monocomponent signals, namely IMFs, each characterized by a well-defined frequency. IMFs must satisfy two simple conditions: (1) the number of extrema and of zero crossings must be either equal or differ at most by one; and (2) the mean value of the envelope defined by the local maxima and of the envelope defined by the local minima is zero. To decompose a time series into its IMFs, an iterative procedure must be applied. Basically, the idea is to extract from the signal all its monocomponents, starting from those characterized by the highest frequencies and finishing with those characterized by the lowest frequencies. So, given the time series \( x(t) \), oscillations are locally singled out by locating maxima and minima. The repetition of this operation for the entire time series allows us to determine the highest frequency monocomponent. By subtracting it from the original signal, we obtain a new signal that still contains oscillating modes but characterized by frequencies lower than that of the extracted one. Iterating this process results in the original time series decomposed into a number of oscillating modes and a monotonic residual representing the long term trend.

Numerically, this iterative procedure can be schematized into three steps. In the first step, all local extrema (maxima and minima) of the signal \( x(t) \) are found and fitted with two cubic-splines to find the upper, \( x_{\text{upper}}(t) \), and lower, \( x_{\text{lower}}(t) \), envelopes of the signal. In the second step, the mean \( m(t) = (x_{\text{upper}}(t) + x_{\text{lower}}(t))/2 \) is estimated and then subtracted from the original signal, obtaining, say, \( x_{1,0}(t) \) (the first index representing the IMF number and the second the iteration number). In the third step, it is verified whether \( x_{1,0}(t) \) satisfies requirements (1) and (2) to be an IMF. If it does, \( x_{1,0}(t) \) can be raised to the rank of IMF \( 1 \). But if \( x_{1,0}(t) \) does not fulfill the conditions to be an IMF, the so-called sifting process starts. According to the sifting process, \( x_{1,0}(t) \) is considered as the new initial signal in place of \( x(t) \) and the first two steps are repeated \( n \) times until \( x_{1,n}(t) \) meets conditions (1) and (2). At this point, IMF \( 1 \) will be found and

---


www.nonlin-processes-geophys.net/19/667/2012/
the sifting process stops. However, if sifting is performed too many times, IMF could lose its physical meaning. So to guarantee that the sifting process stops before physical information contained in amplitude and frequency modulation does not get lost, a stopping criterion must be considered. Different criteria have been proposed; here we used a combination of two of them. One evaluates the following quantity:

$$SD = \frac{\sum_{k} |x_{k,(j-1)}(t) - x_{k,0}(t)|^2}{\sum_{k} x_{k,(j-1)}^2(t)},$$

where, as above, the first index is representative of the IMF under evaluation and the second index representative of the number of the iteration. According to Huang et al. (1998), if $SD < 0.2$, sifting can be terminated. Together with this criterion we used also another condition, in particular we required the fulfillment of the following inequality:

$$\text{Max} \left[ \frac{|x_{\text{lower}}(t) + x_{\text{upper}}(t)|}{|x_{\text{lower}}(t)| + |x_{\text{upper}}(t)|} \right] < 0.02. \quad (2)$$

Once IMF$_1$ is found, a new initial signal is built, subtracting IMF$_1$ from the original signal $x(t)$ and applying the entire procedure described above using this new signal in place of $x(t)$. EMD is complete when all IMFs and the residue into which the original signal can be decomposed are found. Subsequent IMFs, i.e. IMF$_1$, IMF$_2$, etc. will be characterized by slower and slower oscillating patterns. The residue will generally consist of a monotonic function, or a function with only one maximum and one minimum. The residue should represent the trend contained in the data, and even for data with zero mean, the final residue still can be different from zero. So, to summarize, considering as the original signal the Sym-H time series, after decomposing it via EMD it will be represented as follows:

$$\text{Sym-H}(t) = \left[ \sum_{k=1}^{N} \text{IMF}_k(t) \right] + \text{res}(t). \quad (3)$$

A more detailed description of EMD and stopping criteria can be found in several references (e.g. Huang et al., 1998, 2003; Huang and Wu, 2008; Flandrin et al., 2004). It is worth underlining here that the choice of one or more stopping criteria is a real critical point of EMD. Actually, different choices of criteria and of associated thresholds could result in different IMFs (both in shape and number). In what follows EMD will be used to decompose six Sym-H time series characterized by high, moderate and low geomagnetic activity levels.

### 3 Data description and analysis

Data used refer to Sym-H index during 6 different periods (reported in Table 1). A portion of each period is characterized by the occurrence of a large geomagnetic storm, while the remaining part of the period is characterized by low or moderate geomagnetic activity.

Figure 1 exhibits Sym-H index during the selected different periods (named from A to F). The basic characteristic of the solar wind at 1 AU, as measured by L1 satellites (ACE/WIND), during each of the six selected time intervals are well described in Zhang et al. (2007). Each period can be divided into 3 subintervals, each characterized by either low, moderate or intense geomagnetic activity, respectively.

We apply the EMD, as described in Sect. 2, to the selected time series to separate them into their corresponding IMF. As example of our analysis, Fig. 2 reports the obtained IMFs relative to the Sym-H index from 1 July 2000 to 31 July 2000 (period B).

As a result of the decomposition, one can see that the characteristic scale of each IMF increases with the mode index $k$ and that there is a general separation of the data into locally non-overlapping time scale components. We recall that the number of IMFs can vary depending on the algorithm used to implement EMD. In particular, as already mentioned in Sect. 2, the critical point is represented by the stopping criterion. However, being the EMD decomposition is a sort of diadic filter bank (Flandrin et al., 2004), it is reasonable to expect a number of IMFs of the order of $\log_2(N)$, where $N$ is the length (number of points) of the considered time series. In the case of the analysis of period B, the original time series is decomposed in 19 IMFs plus a residue. This number is higher than $\log_2(N) \sim 15$, suggesting that the studied time series possesses a more complex structure and a higher information content than one of a purely stochastic noise. A similar number of IMFs is obtained after the application of the EMD method to the other selected periods. From this decomposition of the original data, we comprehend that an IMF represents a simple local oscillatory mode as a counterpart to the simple harmonic function of the Fourier series method, but it is much more general – it is a function with both amplitude and frequency modulations. From Fig. 2 it is possible to verify that the typical temporal scale of fluctuations in each IMF increases with the index $k$ and that each IMF has a different mean frequency. We can estimate it as

<table>
<thead>
<tr>
<th>Period</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3 Apr 2000</td>
<td>2 May 2000</td>
</tr>
<tr>
<td>B</td>
<td>1 Jul 2000</td>
<td>31 Jul 2000</td>
</tr>
<tr>
<td>C</td>
<td>15 Mar 2001</td>
<td>14 Apr 2001</td>
</tr>
<tr>
<td>D</td>
<td>5 Nov 2001</td>
<td>13 Dec 2001</td>
</tr>
<tr>
<td>E</td>
<td>1 Mar 2003</td>
<td>9 Apr 2003</td>
</tr>
<tr>
<td>F</td>
<td>29 Oct 2004</td>
<td>30 Nov 2004</td>
</tr>
</tbody>
</table>

Table 1. Time intervals of 1-min Sym-H index analyzed in the present study.
Sym-H index according to the following expression:

\[ \delta \text{Sym-H}(t) = (\text{Sym-H}|_{\text{max}} - \text{Sym-H}|_{\text{min}})_{T}. \]  

(4)

For each selected subinterval \( T = \text{I, II, III} \), we estimate the mean frequency \( \langle f_k \rangle \) of each mode \( k \) as

\[ \langle f_k \rangle = \frac{\int_{-\infty}^{+\infty} f S_k(f) df}{\int_{-\infty}^{+\infty} S_k(f) df}, \]  

(5)

where \( S_k(f) \) is the Fourier spectrum of IMF\(_k\) (Huang et al., 1998).

In Fig. 3, the Fourier spectra for all the IMFs are plotted on a double-logarithmic scale (left side) together with the estimated mean frequency of each IMF mode (right side) for the three different time intervals.

The relation between the index \( k \) of the IMF and its mean frequency \( \langle f_k \rangle \) suggests the following exponential law:

\[ \langle f_k \rangle = f_0 \rho^{-k}, \]  

(6)

Fig. 1. 1-min Sym-H(t) time series of the six selected time periods.

Fig. 2. From top to bottom: Sym-H time series for the time period \( B \), from 1 July 2000 to 31 July 2000; 19 IMFs and residue resulting from EMD. Vertical lines denote the separation into three subintervals (I, II, III). The characteristic scale increases with the mode number \( k \).
where $f_0$ is a constant and $\rho$ is given by the slope of the linear fit performed in the semi-log representation of $\log\left(\langle f_k \rangle \right)$ vs. $k$. This result implies that the mean frequency $\langle f_k \rangle$ of a given mode is $\rho$ times larger than the mean frequency of the next one. In our case, taking into account the first ten mean frequency values, we obtained $\rho$ values that are significantly different from 2, which would correspond to a dyadic filter bank as obtained for white noise (Wu and Huang, 2004) and fractional Gaussian noise (Flandrin and Gonçalves, 2004; Flandrin et al., 2004). Taking into account that the lower $\rho$ is, the larger the number of scales involved in the description of a phenomenon, we can consider $\rho$ as a measure of the signal multi-scale features and complexity. On the other hand, it has also been shown that departures of $\rho$ from the value expected for stochastic noises (i.e. $\rho = 2$) are observed in the case of intermittent turbulence, where $\rho$ is significantly lower than 2.

For the above reason, it is interesting to analyze the dependence of $\rho$ on $\delta$Sym-H. Thus, we repeat the previous analysis for the other selected periods (see Table 1). First of all, we apply the EMD methodology to the original data of Sym-H in order to understand how the full spectrum process is split into its intrinsic mode functions. Secondly, we divide the original signal and its IMFs into three subintervals of the same length. For each interval, we estimate the mean frequency of each IMF and successively we determine the value of $\rho$. The obtained results are reported in Fig. 4, where it is possible to evaluate the trend of $\rho$ as function of different levels of geomagnetic activity as described by $\delta$Sym-H.

The diagram suggests that the complexity and the multi-scale nature of the magnetospheric response to the solar wind forcing is higher for intermediate values of $\delta$Sym-H (from 150 nT to 300 nT) than for the other cases. In other words, the multi-scale features are a function of the geomagnetic activity level. Being the diagram of Fig. 4 represents the major result of our multi-scale analysis, the detailed discussion of this figure is postponed to the next section.

### 4 Discussion and conclusions

In the present study, we have analyzed the multi-scale features of a set of large geomagnetic storms by applying the EMD technique as introduced by Huang et al. (1998). The EMD analysis has emphasized the multi-scale and complex nature of the geomagnetic field fluctuations during large geomagnetic storms. This point is corroborated by the high number of intrinsic mode functions necessary to decompose the considered Sym-H time series.

We note that, in all the investigated large storms, the number of IMFs obtained from the EMD technique is higher than what is normally expected for purely stochastic noise. This
suggests that the information content of the storm time signal is higher than one of purely stochastic noise, and that the multi-scale features are more relevant. This is closely reflected in the values of $\rho$, which are lower than 2. According to Huang et al. (2008), this higher number of IMFs could be due to the intermittent features of the investigated time series. Indeed, intermittency is the counterpart of a more complex behavior, which involves non-trivial scaling features and perhaps a higher number of scales.

Our major result is reported in Fig. 4 where the dependence of $\rho$ on $\delta\text{Sym-H}$ is shown. Considering the $\rho$ parameter as a sort of measure of the multi-scale character of the observed fluctuations, the plot seems to suggest that the multi-scale nature assumes a maximum value during time intervals characterized by intermediate $\delta\text{Sym-H}$ values. Taking into account that $\rho$ has been obtained from the dependence of $\langle f_k \rangle$ on $k$ at frequencies $\leq 10^{-3}\text{ min}^{-1}$, the multi-scale nature of fluctuations refers to typical time scales below 1000 min (below 1 day). Thus, the increase of the multi-scale nature of $\text{Sym-H}$ is maximum for medium to large geomagnetic activity, while it is reduced in the case of extreme geomagnetic storm events.

In a recent study on storm–substorm relationship (De Michelis et al., 2011), it has been demonstrated that the influence (in terms of information flows) of substorms on storms is maximum in the case of moderate/intermediate geomagnetic activity level. In particular, during intense magnetic storms, which are mainly caused by an enhanced magnetospheric convection following a long-lasting period of southward interplanetary magnetic field conditions (Kamide, 1992; McPherron, 1988), the driving of storms by substorms seems to be less effective. On the other hand, in several papers it has been widely documented how the magnetospheric dynamics in the course of magnetic substorms shows dynamical complexity and an inherent multi-scale nature (e.g. Consolini and Chang, 2001; Chang et al., 2003; Klimas et al., 2000; Lui et al., 2000; Chapman and Watkins, 2001).

The above considerations permit us a physical interpretation of the results plotted in Fig. 4 in terms of the relevance of internal dynamics with respect to externally-driven enhanced convection. To fully understand and correctly interpret the results reported in Fig. 4, we remind that $\text{Sym-H}$ essentially describes the ring current dynamics, and consequently we have to properly examine all those phenomena that are related to substorms and considerably influence this current system. Indeed, we can suppose that during a moderate activity level the internal magnetospheric dynamics are strongly affected by the impulsive and bursty character of plasma transport in the equatorial magnetotail regions (De Michelis et al., 1999). This plasma transport process, which is responsible for the substorm related phenomena, has been shown to be characterized by a strong intermittent coherent character (Consolini and Chang, 2001; Chang et al., 2003; Klimas et al., 2000). This might be why the multiscale character becomes highly relevant during periods of moderate geomagnetic activity. In contrast, during low and high geomagnetic activity levels, the geomagnetic field fluctuations seem to be a consequence of a very stochastic dynamics, similar to the global dynamics that characterize a Markovian nonequilibrium relaxation process (e.g. de Groot and Mazur, 1984). This interpretation is also in agreement with the range of scales where the increment of the multi-scale nature is observed, being these scales are those typical of substorm events. In other words, the increase of the multi-scale nature during moderate activity levels reflects the influence of substorms to storms that is maximum during these periods. We can understand this point by considering the different roles played by two main processes: the overall magnetospheric plasma convection and the impulsive relaxation events observed in the magnetotail regions, which are responsible for the geomagnetic disturbances during different geomagnetic activity levels. We can suppose that during low-to-moderate geomagnetic activity periods the observed geomagnetic fluctuations are dominated by the impulsive relaxation events that dominate the substorm dynamics, while during periods of high geomagnetic levels the fluctuations and the internal magnetospheric dynamics are principally dominated by the enhancement of the overall plasma convection. This interpretation is also corroborated by the results on the nonequilibrium phase transition nature of the magnetospheric dynamics in response to solar wind changes, which displays both first and second order features (Sitnov et al., 2000, 2001).

In conclusion, this preliminary study indicates the inherent multi-scale nature of the magnetospheric dynamics during large storm events and its dependence on the overall geomagnetic disturbance level. Clearly, more work is necessary to confirm these results on a larger survey of data.

Acknowledgements. We acknowledge the Kyoto World Data Center for providing Sym-H index time series used in the present study.

Edited by: B. Tsurutani
Reviewed by: A. C.-L. Chian and A. L. Clua de Gonzalez

References


Consolini, G., De Michelis, P. and Tozzi, R.: On the earth’s magnetospheric dynamics: Nonequilibrium evolution


