

Comment on the existence of a long range correlation in the geomagnetic disturbance storm time (Dst) index

Lucas Lacasa

Abstract Very recently (Banerjee et al. in *Astrophys. Space*, doi:1007/s10509-011-0836-1, 2011) the statistics of geomagnetic Disturbance storm (Dst) index have been addressed, and the conclusion from this analysis suggests that the underlying dynamical process can be modeled as a fractional Brownian motion with persistent long-range correlations. In this comment we expose several misconceptions and flaws in the statistical analysis of that work. On the basis of these arguments, the former conclusion should be revisited.

Keywords Solar-terrestrial relations · Solar wind

In Banerjee et al. (2011) the authors make use of some methods for nonlinear time series, including the recently introduced Visibility (Lacasa et al. 2008) and Horizontal Visibility algorithms (Luque et al. 2009) to describe the statistical properties of geomagnetic time series, and conclude accordingly that its structure is compatible with an underlying long-range correlated stochastic process as the dynamical mechanism originating such phenomenon. This is an important statement and therefore it must rely on solid grounds. In this comment we wish to stress some flaws and possible pitfalls that the authors have committed in their statistical analysis, which prevents us from assuming *a priori* the validity of their findings.

According to the Visibility Algorithm (Lacasa et al. 2009), fractional Brownian motion series with Hurst exponent H map into scale free Visibility graphs with a power

law degree distribution of the shape $P(k) \sim k^{-\gamma}$, where

$$\gamma(H) = 3 - 2H \quad (1)$$

The authors initially center their study in the determination of the Hurst exponent of different processes. It must be first stressed that this relation only restricts to the aforementioned kind of non-stationary stochastic processes with self-similar increments. Therefore, it does not apply for instance to random uncorrelated series, such as the ones firstly analyzed by the authors (Fig. 3 and Table 1 in Banerjee et al. 2011). Notice at this point that in Lacasa et al. (2008) it was already pointed out that random uncorrelated series map into visibility graphs with an *exponential* degree distribution, instead of one following a power law relation, and that there is a theorem that precisely proves such a thing within the Horizontal version of the method (Lacasa and Toral 2010) (something that the authors paradoxically seem to be aware of, given the fact that they make use of the Horizontal visibility algorithm further in their study). Note that such exponential relation becomes evident just by visual inspection, in a semi-log plot of $P(k)$. This is presumably the shape behind the results shown in Fig. 3 of Banerjee et al. (2011) (a log-log plot of an exponential function reveals as a curved shape, such as the one found by the authors, and shouldn't be confused with a power law). It is indeed a bit disturbing that the authors claim that the resulting value of $H = 0.2232$ “*coincides with the exact specified value for the perfectly random series*”, a straightforwardly wrong statement.

The second time series addressed by the authors is the Conway series. Again, while this is a self-similar series, it does not fall under the umbrella of an fBm (if only because the Conway series is a deterministic fractal series recursively generated, instead of a stochastic one), and accordingly relation (1) cannot be used here as well (Table 1 in Banerjee et al. 2011).

L. Lacasa (✉)
Dept. Matemática Aplicada y Estadística ETSI Aeronáuticos,
Universidad Politécnica de Madrid, Madrid, Spain
e-mail: lucas.lacasa@upm.es

On a second step, the authors proceed by analyzing, using the same methodology, the empirical geomagnetic datasets, concluding from their analysis that the series have an associated Hurst exponent whose value suggests persistent long-range correlations. While in principle the data under study are a candidate for such an analysis (concretely, since their associated visibility graph have a heavy tailed degree distribution), the results are not conclusive at all, and I will explain the reasons. The data under study range from 2001 to 2006 with six hour sampling rate, i.e. a total set of about 9000 data: this is in principle enough to accurately determine the presence of long-range correlations according to a previous study (Lacasa et al. 2009). However, the authors deliberately only address datasets of one year long (i.e. sizes of about 1500 data each). From a statistical point of view this is problematic, since within the visibility algorithm, the basis for measuring self-similarity (and hence the existence of a well defined Hurst exponent) is a power law degree distribution. Such a distribution is in general very difficult to distinguish from other shapes in the absence of good statistics, and typically one requires at least two decades of statistics prior to performing a likelihood test (Newmann 2005; Clauset et al. 2009). The associated results on the Hurst exponent are therefore ambiguous at this point, and should therefore be validated with larger datasets. Some other misconceptions include plotting non-normalized probability distributions (Figs. 3, 4, 6 and 7).

Finally, the authors calculate the autocorrelation function of the signals and their associated power spectral density (incidentally, these two magnitudes are univocally related via the well known Wiener-Khinchin theorem: one is the Fourier transform of the other, at odds with the authors quote “*the power spectral density of the autocorrelation function*”). First, notice that these magnitudes are well defined only for stationary processes, whereas the main claim of the authors, which reads that the Dst index “*follows the same model that of a stochastic fractional Brownian motion*”, resides on an intrinsically non-stationary process. Second, the problem related to the small size of the datasets is also present here. Third, and more important, what is indicative of long-range correlation is not, as the authors suggest, “*the presence of a high peak in the periodogram*”, but the occurrence of peaks at all scales, something which is usually studied plotting in log-log such periodogram and again performing a maximum likelihood estimation of a power law decaying function. The evidence in this direction is not convincing at all, according to the authors results (Fig. 8).

In summary, the analysis conducted by the authors in Banerjee et al. (2011) present too many flaws and problems to unambiguously conclude whether if the series under study have been generated from a long-range correlated stochastic process.

1 Chatterjee replies

1. Lacasa suggested that a set of 9000 data is ‘in principle enough to accurately determine the presence of long-range correlations’ and a the power law degree distribution associated with the Dst index time series is ‘in general very difficult to distinguish from other shapes in the absence of good statistics.’ Here the statistical significance of 9000 data is not particularly understood. Also the data set are divided into different years for studying the variations of the nature of the Dst index over the rising and falling part of the solar cycle.
2. Regarding the periodogram of power spectral density of the auto-correlation function, the magnitudes are plotted with respect to time, not frequency. So log-log plot is not applicable here. Also the data set here are divided according to the nature of the solar cycle as previously stated. The high peak value in the periodograms represents the period in which the Dst index value repeats in that particular part of the solar cycle.

References

- Banerjee, A., Bej, A., Chatterjee, T.N.: On the existence of a long range correlation in the geomagnetic disturbance storm time (Dst) index. *Astrophys. Space. Sci.* (2011). doi:10.1007/s10509-011-0836-1
- Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C.: From time series to complex networks: the visibility graph. *Proc. Natl. Acad. Sci. USA* **105**(13), 4973 (2008)
- Luque, B., Lacasa, L., Ballesteros, F., Luque, J.: Horizontal visibility graphs: exact results for random time series. *Phys. Rev. E* **80**, 046103 (2009)
- Lacasa, L., Luque, B., Luque, J., Nuno, J.C.: The Visibility Graph: a new method for estimating the Hurst exponent of fractional Brownian motion. *Europhys. Lett.* **86**, 30001 (2009)
- Lacasa, L., Toral, R.: Description of stochastic and chaotic series using visibility graphs. *Phys. Rev. E* **82**, 036120 (2010)
- Newmann, M.E.J.: Power laws, Pareto distributions and Zipf’s law. *Contemp. Phys.* **46**(5), 323–351 (2005)
- Clauset, A., Shalizi, C.R., Newman, M.E.J.: Power-law distributions in empirical data. *SIAM Rev.* **51**(4), 661–703 (2009)