Prediction of Geomagnetic Activity: A Novel Approach

G.K. Rangarajan*
B.D. Kadam**
S. Alex**

Resumen

El índice Aa aprobado por IAGA, disponible desde 1868, ha sido extendido hasta 1844, por Nevanlinna, utilizando las observaciones de Declinación del observatorio de Sodankyla. Los valores promedio mensuales del índice Aa extendido pudieron descomponerse en partes dependientes directamente de la actividad solar y una parte residual controlada predominantemente por un torrente de vientos solares de elevada velocidad, como lo sugirió Feynman (1982). Utilizamos la técnica adaptativa de datos, recientemente desarrollada, llamada Análisis de Espectro Singular (SSA) en las dos partes del índice Aa, para identificar los principales componentes, por arriba del nivel del ruido. Los principales componentes individuales son extrapolados utilizando un modelo autorregresivo adaptando el algoritmo de entropía máxima de Burg. Los valores extrapolados se combinan para producir los valores promedio mensuales calculados para 36 meses. La comparación de los valores observados y calculados de Aa para los años recientes mostraron un buen acuerdo, indicando la potencialidad de esta nueva aproximación en un exitoso pronóstico de las variaciones a largo plazo en la actividad geomagnética.

Abstract

IAGA-approved Aa index, available from 1868, has been extended backwards up to 1844 by Nevanlinna using the Declination observations of Sodankyla observatory. The monthly mean values of the extended Aa index could be decomposed into a part directly dependent on solar activity and a residual part controlled dominantly by recurrent high speed solar wind streams, as suggested by Feynman (1982). We

* Observatorio Nacional, Rio de Janeiro, Brazil.
** Indian Institute of Geomagnetism Mumbai India.

use the recently developed data adaptive technique, called Singular Spectrum Analysis (SSA), on the two parts of the Aa index to identify significant principal components, above the noise level. Individual principal components are extrapolated using an autoregressive model adapting Burg's maximum entropy algorithm. The extrapolated values are then combined to yield predicted monthly mean values for 36 months ahead. Comparison of the observed and predicted values of Aa for the recent years show good agreement indicating the potentiality of this novel approach in successful prediction of long-term variations in the geomagnetic activity.

Introduction

Indices of geomagnetic activity provide a condensed summary of the state of the Earth's magnetic environment and have been in use for more than a century. Mayaud (1973) produced a long, homogeneous series of geomagnetic index Aa, derived from two nearly antipodal observatories in England and in Australia. This series, available from 1868, has been extensively studied for analysis of long term trends in geomagnetic activity (Feynman 1982; Borello-Filisetti *et al.* 1992 and others).

Using the hourly values of Declination at Helsinki, Finland, Nevanlinna and Kataja (1993) were able to extend the Aa series backward in time up to 1844. They found that the monthly means of the derived Aa index showed excellent linear correlation with the actual Aa index for the common period of availability between 1868 and 1880.

Forecasting the level of geomagnetic activity over short time scales using Interplanetary Magnetic Field and solar wind conditions has been attempted based on empirical models (Burton *et al.* 1975; Feldstein *et al.* 1984, 1994) and these are useful in taking precautions to protect the delicate instruments on board the orbiting satellites. Long term forecast of geomagnetic activity, on the other hand, will be useful both in the planning stages of space experiments and as a proxy for the interplanetary conditions that may be expected. NOAA Space Environment Center in Boulder USA, for example, carry out such exercises on a regular basis.

In this investigation, we use the homogeneous series of Aa index and a recently developed technique called Singular Spectrum Analysis (Vautard *et al.* 1992) to show the success of a novel approach to forecast the expected level of geomagnetic activity for each month for succeeding several months.

Data and Methodology

The data base consists of the monthly mean values of Aa index covering the period 1844 to 1995. To avoid the undue influence of any single large value (due to sporadic severe geomagnetic disturbance) on the monthly mean values, we take overlapping means covering three years at a time. This smoothed time series is used in the subsequent analyses.

The efficacy of Singular Spectrum Analysis (SSA) in providing quantitative and qualitative details about the deterministic and stochastic parts of a time series, how the method generates data adaptive filters which isolate significant spectral components, how it helps reconstruct the original time series from a few numbers of individual components and how the noisy component is eliminated in the process have been highlighted by Vautard and Ghil (1989). The mathematical steps involved in the processing of the time series to generate the data adaptive filters and the filtered components are given in detail by Dettinger *et al.* (1995), Rangarajan and Araki (1997) and Rangarajan and Iyemori (1997).

The method, in brief, is described next. From the zero-mean time series, autocorrelations are computed for successive lags up to an appropriately chosen maximum lag M (termed the embedding space or the viewing window). These are used in generating a Toeplitz matrix with the first row given by the series of autocorrelations. Eigen values and eigen vectors of this matrix are determined. All eigen values are expected to be positive as the matrix is positive and symmetric. Number of non-zero eigen values correspond to the number of independent variables in the system. Other eigen values, close to zero, define the 'noise floor' (Sharma *et al.* 1993). The elements of the eigen vectors corresponding to significant eigen values serve as data adaptive filters and when used on the time series, extract the individual component represented by it. As the eigen vectors are orthogonal, the extracted components are relatively independent of each other. When cumulatively added, these individual components reproduce the original time series, less the noisy part.

Penland *et al.* (1991) demonstrated, using a synthetic signal and a meteorological time series, that the spectrum of the individual components obtained through SSA are much smoother, can be computed with very low order Auto Regressive (AR) process and they can be combined to arrive close to the true spectrum of the time series, uncontaminated by noise. Keppenne and Ghil (1992) utilized this feature to extrapolate individual component series using the low order AR coefficients and combine them subsequently to successfully forecast the time series of the Southern Oscillation Index.

To improve the reliability of the forecast values, we modified the procedure of Keppenne and Ghil (1992) as described below:

For an AR process, we have

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \ldots\ldots + a_m X_{t-m} + n$$

where a1 to am are the Prediction Error Filter coefficients (PEF) derived using the algorithm of Burg (1968) and 'n' is the noise component that cannot be modeled.

Using the PEFs and the observed values, we can extrapolate one step ahead at a time. Using next the extrapolated value as input we can estimate the second next etc. up to any desired limit. It should, however, be kept in mind that as we extend further into the future more and more of extrapolated rather than observed values are utilized and there will be progressive loss of reliability.

To improve the stability of the forecast values, we proceed as follows:

As an example, let us consider 100 values of a time series to be extrapolated to 110 values. We first use the 100 observed values, isolate the individual components, extrapolate the next 10 values for each of the components using the equation above. Next, we use the first 99 values, leaving out the last value to estimate the next 9 values (one value being an observed value) and so on till we arrive at a time series of first 91 values with one extrapolated value for 101. Thus the first extrapolated value is estimated 10 times, the next 9 times etc. which are averaged with weights 1.0, 0.9,........,0.2 and 0.1 for the first to the tenth trials respectively. The averaged extrapolated values of individual components are then combined and the mean value of the time series is added to provide a forecast of the time series.

Feynman (1982) showed that the geomagnetic activity, and by proxy the solar wind, has two distinct components: R component with the phase and relative amplitude of the sunspot cycle and the I component related to long-lived solar structures such as the coronal holes. She also showed that the observed annual mean Aa index can be considered a sum of AaR and AaI where AaR is given by the relation:

$$\mathrm{AaR} = 0.12\mathrm{Rz} + 5.38$$

and Rz is the annual mean sunspot number for the corresponding year.

Nevanlinna and Kataja (1993) found that this relationship holds good even when the Aa data was extrapolated backwards in time to 1844. Figure 1 depicts clearly that except for a couple of stray points, all the others lie above the line corresponding to the relation given by Feynman (1982). Using this linear relationship between the sunspot number and the Aa index, we generate two time series AaR and AaI representing two distinct components of geomagnetic activity, analyze them separately, extrapolate them individually and then combine the two to get an estimate of the geomagnetic activity for the succeeding months.

Results and Discussion

In Figure 2a, we show the first six eigen vectors corresponding to the first six significant eigen values. The percentage variance accounted for by the component corresponding to each vector as a fraction of the total variance of the time series are indicated. These are derived from the relation:

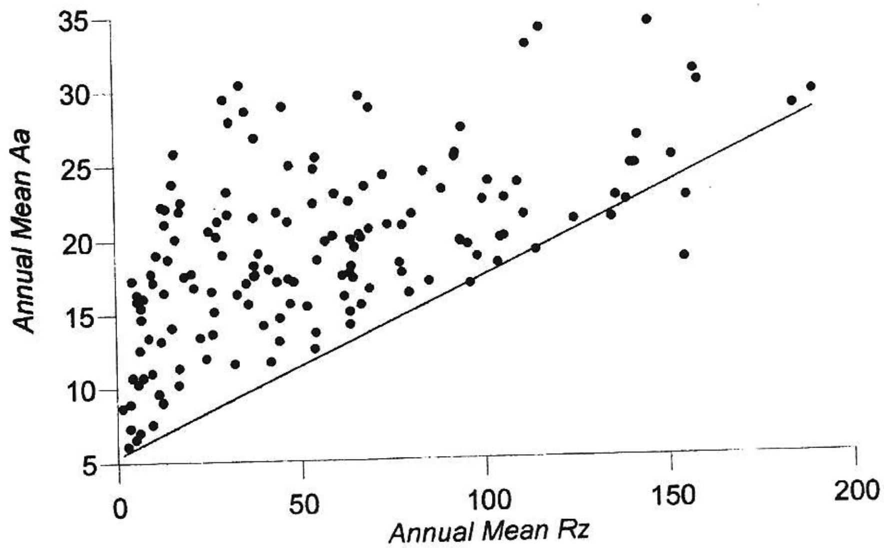

Figure 1. Annual mean sunspot number and corresponding annual mean Aa index for the Period 1844-1995.

$$\text{Percent Variance} = [\lambda_k/\Sigma\lambda_k] \times 100$$

We find that just six distinct components together account for ~99% of the total variance.

We use a window size of 132 months (to ensure that the 11-year solar cycle component is properly incorporated). The 132 elements of each eigen vector with values between + 1 and –1 constitute the weights of the digital filter operating on the time series, one at a time. The resulting filtered series of the six components for AaR are shown on Figure 2b. They, in effect, correspond to two solar cycle components, a long term trend in the AaR series and other quasiperiodic oscillations which have been discussed in great detail by Rangarajan (1998).

When AaI series are analyzed in a similar fashion, we find that one needs at least 12 components to account for 98 percent of the total variance, indicating that in comparison to the AaR series, this series is more complex. This is understandable because the AaR series results from variations in Rz which changes smoothly and systematically over approximately 11 years whereas the AaI component results from long-lived solar features which have much less regularity. The 12 reconstructed components for AaI are shown in Figure 3.

The success of the methodology of SSA is clearly demonstrated in Figure 4 which compares the original AaR and AaI with the reconstructed AaR and AaI and

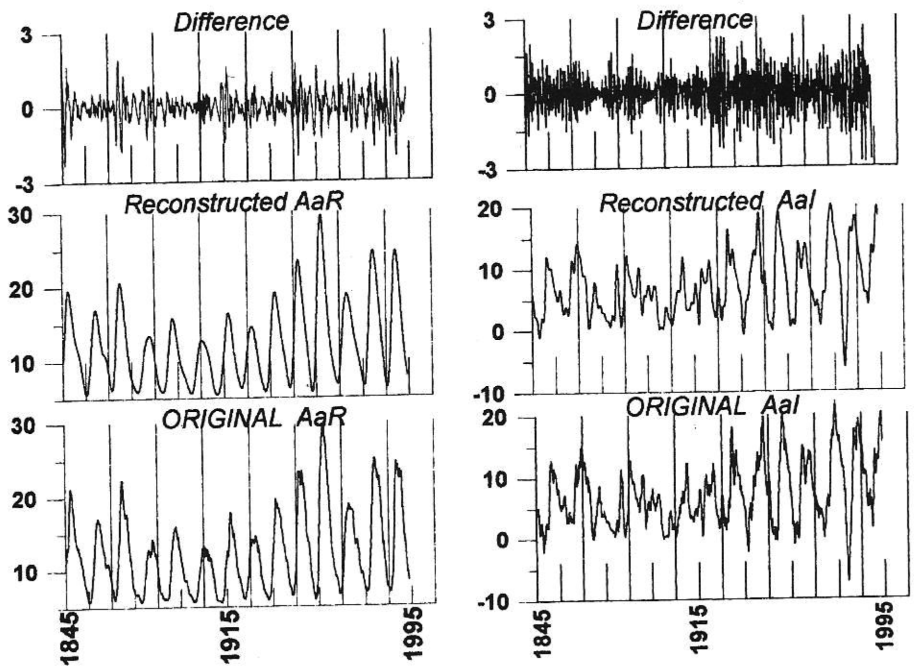

Figure 4. Comparison of the original and reconstructed AaR and AaI time series. Also shown at the top are the noisy part, not accounted for by the reconstructed components.

plot the residuals. The influence of noise reduction by the technique can also be seen when one compares the short spiky oscillations in the original data being smoothed out in the reconstructed series.

Next, we extrapolate each of the 18 time series —six for AaR and 12 for AaI— over the next 36 months to cover the years 1995 to 1997 using appropriate low order AR coefficients, as outlined in the Methodology. Six extrapolated series for AaR and 12 extrapolated series for AaI are then cumulatively added and the predicted values for the three years are shown in Figure 5a.

We notice that while the sunspot- dependent component- AaR- has its minimum values during 1995 and 1996 and start rising from early 1997, the solar wind-associated component —AaI— decreases from 1995 towards 1997. This is consistent with the expectation that geomagnetic activity caused by solar coronal features tend to maximize during the declining and minimum phase of the solar cycle. The shape of the predicted curves are also in conformity with the suggestion of Feynman (1982) that the I component is nearly 180 deg. out of phase with the R component and has nearly equal amplitude. The top curve of Figure 5a indicates the time profile of the anticipated monthly mean values of Aa indices for the three years and is obtained as the sum of the two lower curves.

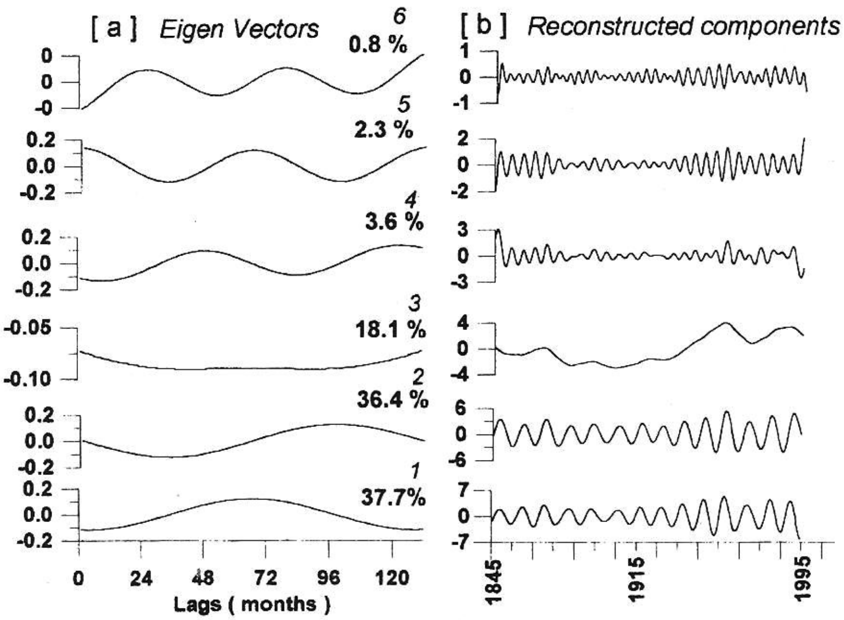

Figure 2. (a) The first six eigen vectors derived from the time series corresponding to AaR. The percentage variance accounted for by individual components are given. (b) The reconstructed components of the AaR time series using the elements of the eigen vectors as data adaptive filter weights.

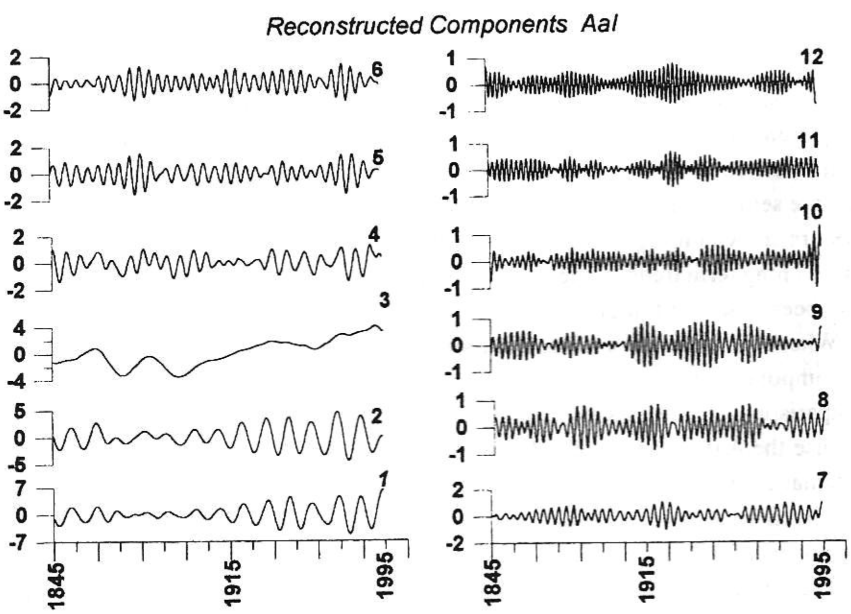

Figure 3. The first twelve reconstructed components of the AaI time series.

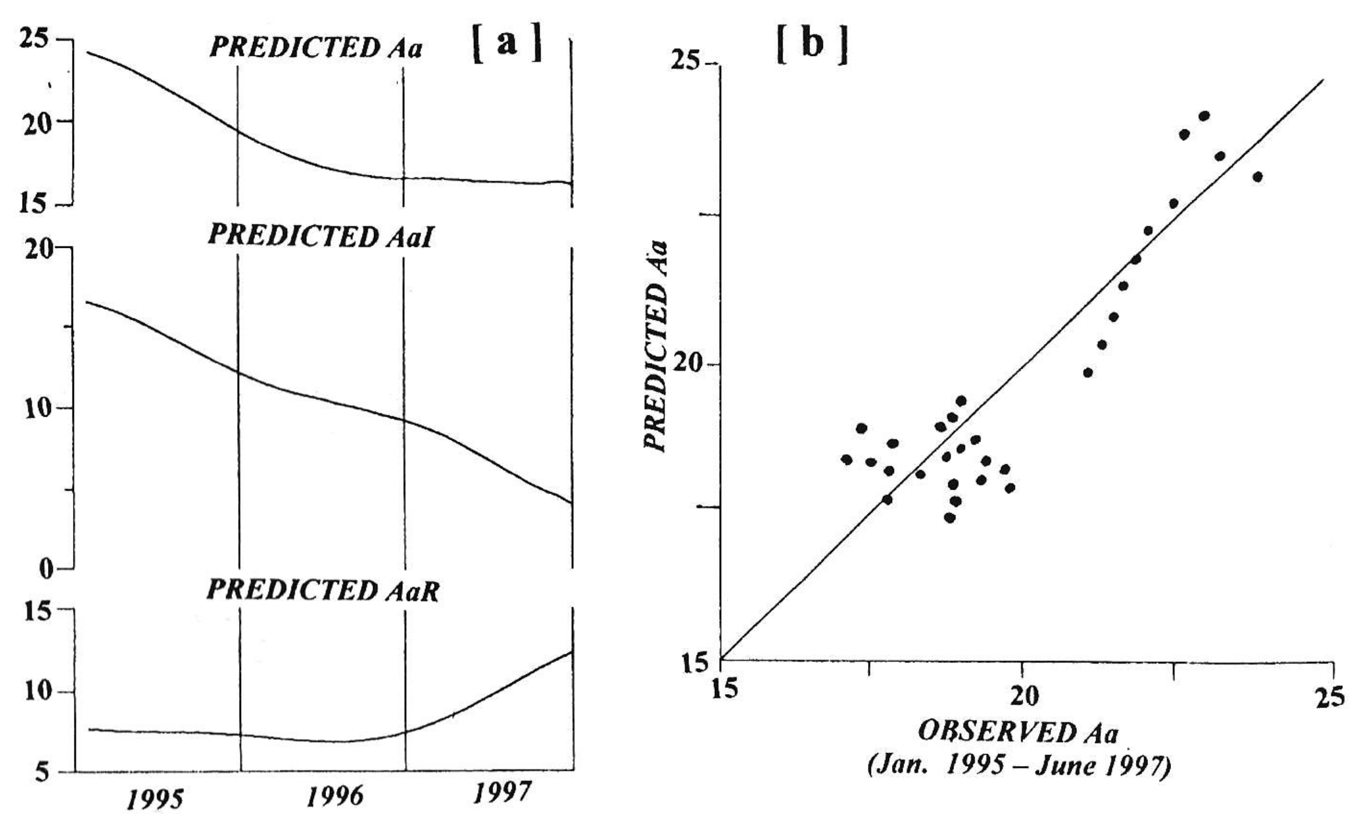

Figure 5. (a) Predicted monthly values of AaR, AaI and their sum (Aa) for the period 1995 to 1997. (b) Scatter plot of the observed and predicted monthly values of Aa index for the period Jan. 1995 to June 1997.

To show that the prediction approach outlined in this investigation has been quite successful, we compare the actual observed mean monthly values (37-month overlapped means) with the predicted values. The scatter plot is shown on Figure 5b. For an ideal prediction, the points should lie on the straight line of equal values, shown in the figure. We notice that though they do deviate from the straight line, at no point the departure is of magnitude 2 or more.

In conclusion, we emphasize that the novel approach of using Singular Spectrum Analysis to isolate significant spectral components in the time series of geomagnetic activity —split into two distinct components with different causative mechanisms— and combining their individual extrapolated values using a low order AR process appears to be very efficient in the forecast of geomagnetic activity several months ahead. The method has the potentiality to be applied to other similar time series which are fairly orderly in their time progression.

References

Borello-Filisetti, O.; Mussino, V.; Parisi, M. and Storini, M., "Long-term variations in the geomagnetic activity level. I A connection with solar activity," *Ann. Geophys.,* 10:668-675, 1992.

Burg, J.P. "Maximum entropy spectral analysis" Presented at 37th annual meeting of SEG-Oklahoma, 1968 (also in Modern Spectrum Analysis. Ed. D.G. Childers, pp. 34-88. IEEE Press, New York, 1978).

Burton, R.K.; McPherron, R.L. and Russell, C.T., "An emperical relationship between interplanetary conditions and Dst," *J. Geophys. Res.,* 80:4204-4214, 1975.

Dettinger, M.D.; Ghil, M.; Strong, C.M.; Weibel, W. and Grafe, A., "Software expedites singular spectrum analysis of noisy time series," *EOS Trans. AGU.,* 76(2):12, 14, 21, 1995.

Feldstein, Ya.I.; Pisaraki, V.Yu.; Rudneva, N.M. and Grafe, A., "Ring current simulation in connection with interplanetary conditions," *Planet. Space Sci.,* 32:975-984, 1984.

Feldstein, Ya.I.; Veshchezerova, U.B. and Golyshev, S.A., "The ring current and auroral Electrojet and their relationships to the interplanetary medium parameters," *Geomag. Aeron,* 33:755-762, 1994.

Feynman, J., "Geomagnetic and solar wind cycles 1900-1975," *J. Geophys. Res.,* 87:6153-6163, 1982.

Keppenne, C.L. and Ghil, M., "Adaptive filtering and prediction of the southern oscillation index," *J. Geophys. Res.,* 97:20449-20454, 1992.

Mayaud, P.N., "A hundred year series of geomagnetic data, 1868-1967, indices aa, storm sudden commencements," *IAGA Bulletin,* 33, IUGG Publication Office, Paris, 1973.

Nevanlinna, H. and Kataja, E., "An extension of the geomagnetic activity index series aa for two solar cycles (1844-1868)," *Geophys. Res. Lett.*, 20:2703-2706, 1993.

Penland, C.M.; Ghil, M. and Weickmann, K.M., "Adaptive filtering and maximum entropy spectra with application to changes in atmospheric angular momentum," *J. Geophys. Res.*, 96:22659-22671, 1991.

Rangarajan, G.K., "Sunspot variability and an attempt to predict solar cycle 23 by adaptive filtering," *Earth Planets Space,* 50:91-100, 1998.

Rangarajan, G.K. and Araki, T., "Multiple time scales in the fluctuations of the equatorial Dst index through singular spectrum analysis," *J. Geomagn. Geoelec,* 49:3-20, 1997.

Rangarajan, G.K. and Iyemori, T., "Time variations of geomagnetic activity indices Kp and Ap: an update,". *Ann. Geophysicae,* 15:1271-1290, 1997.

Sharma, A.S.; Vassilliadis, D. and Papadapoulous, K., "Reconstruction of low dimensional magnetospheric dynamics by singular spectral analysis," *Geophys. Res. Lett.*, 20:335-338, 1993.

Vautard, R. and Ghil, M., "Singular spectrum analysis in non linear dynamics with application to palaeoclimatic time series," *Physica D,* 35:395-424, 1989.

Vautard, R.; Yiou, P. and Ghil, M., "Singular spectrum analysis: a toolkit for short noisy chaotic signal," *Physica D,* 58:95-126, 1992.