Long-term prediction of solar and geomagnetic activity daily time series using singular spectrum analysis and fuzzy descriptor models

Masoud Mirmomeni[1], E. Kamaliha[2], Masoud Shafiee[3], and Caro Lucas[1,4]

[1]*Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering,*
University College of Engineering, University of Tehran, Tehran, Iran
[2]*Computer Engineering Department, Sharif University of Technology, Tehran, Iran*
[3]*Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran*
[4]*School of Cognitive Sciences, Institute for studies in theoretical Physics and Mathematics, Tehran, Iran*

Of the various conditions that affect space weather, Sun-driven phenomena are the most dominant. Cyclic solar activity has a significant effect on the Earth, its climate, satellites, and space missions. In recent years, space weather hazards have become a major area of investigation, especially due to the advent of satellite technology. As such, the design of reliable alerting and warning systems is of utmost importance, and international collaboration is needed to develop accurate short-term and long-term prediction methodologies. Several methods have been proposed and implemented for the prediction of solar and geomagnetic activity indices, but problems in predicting the exact time and magnitude of such catastrophic events still remain. There are, however, descriptor systems that describe a wider class of systems, including physical models and non-dynamic constraints. It is well known that the descriptor system is much tighter than the state-space expression for representing real independent parametric perturbations. In addition, the fuzzy descriptor models as a generalization of the locally linear neurofuzzy models are general forms that can be trained by constructive intuitive learning algorithms. Here, we propose a combined model based on fuzzy descriptor models and singular spectrum analysis (SSA) (FD/SSA) to forecast a number of geomagnetic activity indices in a manner that optimizes a fuzzy descriptor model for each of the principal components obtained from singular spectrum analysis and recombines the predicted values so as to transform the geomagnetic activity time series into natural chaotic phenomena. The method has been applied to predict two solar and geomagnetic activity indices: geomagnetic *aa* and solar wind speed (SWS) of the solar wind index. The results demonstrate the higher power of the proposed method—compared to other methods—for predicting solar activity.

Key words: Space weather, geomagnetic disturbance, solar activity indices, prediction, fuzzy descriptor models, singular systems, singular spectrum analysis, GLoLiMoT.

1. Introduction

Most space weather phenomena are influenced by variations in solar activity (Vassiliadis *et al.*, 2000). As such, many of the changes that occur in space weather originate from solar activity, which varies in an 11-year period, called the solar cycle. The solar cycle consists of a period of activity, the solar maximum, and a period of quiet, the solar minimum. There are an increased number of flares during the solar maximum, and these cause a significant increase in solar cosmic ray intensity. The high-energy particles disturb communication systems and affect the lifetime of satellites. Coronal mass ejections and solar flares cause shocks in the solar wind and geomagnetic disturbances in Earth's magnetosphere. A high rate of geomagnetic storms and substorms result in atmospheric heating and increased drag of low earth orbit (LEO) satellites (Mirmomeni *et al.*, 2006). Accurate and reliable long-term solar activity forecasting is especially useful to space mission centers because the orbital trajectory parameters of satellites are greatly affected by changes in solar activity.

Several methods have recently been proposed to predict solar and geomagnetic activity (Haykin, 1994; Lillekjendlie *et al.*, 1994; Nelles, 2001; Leung *et al.*, 2001). Given the achievements in the field of chaotic systems, several such methods can be used in the prediction of solar activity indices, namely polynomial function approximation, reconstructions using Lyapunov exponents, and the detection of periodicity in chaotic time series (Vautard and Ghil, 1989). Of these methods, the physical precursor (Brown, 1992; Thompson, 1993; Joselyn *et al.*, 1997) and solar dynamo techniques (Schatten *et al.*, 1978, 1996; Schatten and Sofia, 1987; Schatten and Pesnell, 1993; Sofia *et al.*, 1998), which are based on simple linear and nonlinear empirical studies, seem to have superseded the numerical and black box techniques in terms of user preference (Tong and Lim, 1980; Weigend *et al.*, 1992; Tong, 1996; Uluyol *et al.*, 1998; Lucas *et al.*, 2003; Gholipour *et al.*, 2003). In the black box modeling technique, in comparison with the white box modeling technique (both of which are fully derived by applying principles of physics, chemistry, biology, economy,

etc.), both model structure and parameters are determined from experimental modeling. In other words, to build a black box model, no or very little prior knowledge is exploited (Nelles, 2001).

In comparison, descriptor systems have attracted increased interest from researchers during the past two decades due to their many practical applications (Raouf and Boukas, 2004; Yonchev *et al.*, 2004; Mirmomeni and Shafiee, 2005a, b). Such systems describe a wider class of systems, including physical models and non-dynamic constraints. The descriptor system is much tighter than the state-space expression for representing real independent parametric perturbations (Taniguchi *et al.*, 2000; Lu and Ho, 2003; Meng and Zhang, 2006). In addition, fuzzy descriptor (FD) models as generalizations of Takagi-Sugeno fuzzy models and locally linear neurofuzzy models are general forms that can be trained by constructive intuitive learning algorithms. They fulfill the principle of network parsimony, which results in high generalization property. These models were first reported by Tadanari *et al.* (Taniguchi *et al.*, 1999a, b). The Takagi-Sugeno fuzzy model has been proven to be a universal approximator of any smooth nonlinear systems that are first order differentiable (Taniguchi *et al.*, 2000). This property provides a strong basis for using the FD models in the prediction of some solar activity indices, which have complex dynamics. In addition, the application of spectral analysis for eliciting the main pattern of a time series is useful for determining a number of long-term trends in solar activity indices. In turn, these indices can be used for making long-term predictions by a number of data-driven approaches. One of the well-known methods in spectral analysis is singular spectrum analysis (SSA). This method was originally designed to extract information from short noisy chaotic time series, provide an insight into the unknown dynamics, and enhance the signal-to-noise ratio (SNR) (Vautard and Ghil, 1989; Vautard *et al.*, 1992). In addition, SSA performs a data adaptive filtering in the lag coordinate space of data and yields the principal components of time series which have narrow band frequency spectra and obvious temporal patterns. The principal components include linear or nonlinear trends, periodic and quasi-periodic patterns, and some lower amplitude signals that can be considered as colored noise. Most of the narrow band periodic components can be estimated via simple and optimal linear models, while there are always a number of more complex patterns that present nonlinear characteristics. Thus, when reconstructing the original time series from the principal components, one should use both linear and nonlinear techniques and also the linearity tests.

Here, we propose a new method that combines FD models with SSA, denoted here as FD/SSA, for the long-term prediction of solar activity indices. First, SSA is used to elicit the main patterns of the solar and geomagnetic activity indices. After these main patterns or principal components of the original time series have been elicited, a FD model is trained by an incremental learning algorithm, which is called Generalized Locally Linear Model Tree (GLoLiMoT) algorithm, for each principal component to make an intuitive nonlinear black box modeling technique (which is based solely on measurement data) applicable to

the long-term prediction of solar and geomagnetic activity indices. To demonstrate the advantage of this method, we compare the performance of the FD model with singular spectrum analysis to several neural and neurofuzzy models for predicting two geomagnetic activity indices: solar wind speed (SWS) of the solar wind index and the geomagnetic *aa* index as a proxy indicator of the "new" magnetic field. Our results demonstrate the excellent performance of this combined model in predicting solar activity as compared to other neural and neurofuzzy models.

This paper is structured as follows. Section 2 briefly introduces the nonlinear descriptor system, describing its characteristics compared to regular systems and addressing a number of complex phenomena that can be used for some good test beds to show the performance of such systems in describing the characteristics of such phenomena. Section 3 introduces SSA and FD models and discusses their characteristics. Section 4 is devoted to describing the learning methodology that is used for FD models to predict solar activity indices. In Section 5, the FD model and SSA are used to predict SWS and geomagnetic *aa* indices to show the performance of this method in comparison with other methods. The last section contains the concluding remarks.

2. Characteristics of Nonlinear Descriptor Systems

A singular implicit differential equation is an implicit ordinary differential equation which takes the form of

$$F(\dot{x}(t), x(t), u(t), t) = 0, \quad x(t_0) = x_0 \quad (1)$$

where x is an n-dimensional state vector, u is an m-dimensional control vector, t is time, and the Jacobian matrix $\frac{\partial F}{\partial \dot{x}}$ is singular (Xiaoping, 1995). A system which is described by a singular implicit differential equation is called a singular system (Campbell, 1980; Newcomb, 1981). Singular systems are often referred to as differential algebraic equations because they frequently are a mixture of differential and algebraic equations; that is, they take the form of

$$\begin{aligned} \dot{x}(t) &= f(x, u, t) \\ 0 &= g(x, u, t) \end{aligned} \quad (2)$$

One can define a matrix E such that

$$\begin{aligned} E\dot{x}(t) &= F(x, u, t), \\ E &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \\ F(x, u, t) &= \begin{bmatrix} f(x, u, t) \\ g(x, u, t) \end{bmatrix} \end{aligned} \quad (3)$$

This general form of the differential equation is used as a canonical form of nonlinear singular systems in many papers. These systems are also called descriptor systems (Luenberger, 1977) because they are the way in which the system is initially described. Other names for descriptor systems are constrained systems (Xiaoping, 1995), degenerate systems (Pandolfi, 1981), generalized state-space systems (Verghese *et al.*, 1981), semi-state systems (Dziurla and Newcomb, 1979), non-canonic systems (Xiaoping,
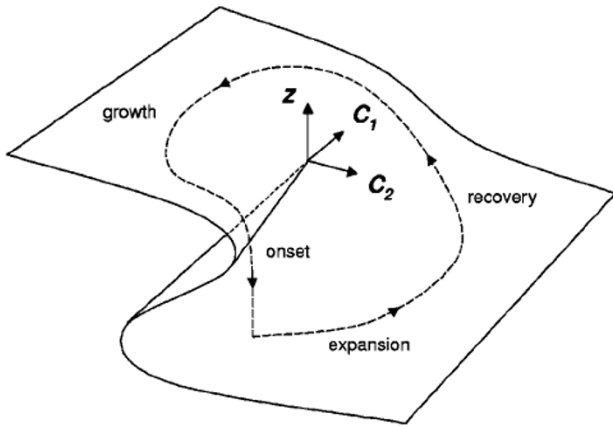
Fig. 1. Hypothetical cusp catastrophe manifold that was expected to approximate the sub-storm dynamics of the magnetosphere according to the model (Lewis, 1991). The evolution of an isolated sub-storm is shown by dashed arrows.

1995), and differential equation on a manifold (Xiaoping, 1995). These systems appear during the study of robotics, optimal control, economics, large-scale interconnected systems, among others (Dziurla and Newcomb, 1979; Kang and Tang, 2005; Shafiee and Amani, 2005).

If E is regular, the implicit ordinary differential equation (ODE) (Eq. (3)) is equivalent to the explicit ODE: $\dot{x}(t) = E^{-1}F(x(t), u(t))$. This case has been comprehensively studied and is now rather well understood. When E is singular in (Eq. (1)), resulting in what we shall term a *generalized state-space system* or *descriptor system*, this behavior is considerably modified. In contrast to the *regular state-space system*, we find that the number of degrees of freedom of the system, i.e., the number of independent initial values that $x(0_-)$ (here the $x(0_-)$ means the states of the system before the beginning of the process; in other words, 0_- means before the time zero) can take is now evidently reduced to

$$f = \operatorname{rank} E < n \qquad (4)$$

The term *generalized order* has been proposed for f (Verghese *et al.*, 1981). Therefore, in such systems, state spaces have to satisfy some constraints, and state variables have to be on a manifold in state space. In addition, it has been proven that the output of such systems may include some impulsive motions even if there is no impulse input to the system (Verghese *et al.*, 1981). Such characteristics show the power of descriptor systems in describing complex phenomena. For example, in modeling the sub-storm dynamics of the magnetosphere, the long-term prediction of which is the aim of this paper, both the surface and the corresponding circulation flows turn out to be surprisingly close to a very simple low-dimensional scheme of the magnetospheric sub-storm as a cusp catastrophe, which was first proposed by Lewis (1991) and illustrated in Fig. 1 (where the z Parameter is the state parameter and c_1 and c_2 are the control parameters) (Sitnov *et al.*, 2001). Cusp catastrophe is one of the seven basic catastrophic forms which belong to nonlinear dynamics, with exotic characteristics introduced by René Thom in his fascinating "Catastrophe theory" in

1972. This theory, as a qualitative mathematical method, is a very effective means for increasing the interaction between engineers and physicians. Catastrophe theory is a mathematical method for describing the evolution of forms in nature. It was created by René Thom, who wrote a revolutionary book "Structural stability and morphogenesis" in 1972 expanding the philosophy behind the concepts. (For more detail, refer to Zeeman (1977) and Saunders (1980)).

3. Combined (FD/SSA) Long-term Predictor

In this section, we first present a statement on the problems involved in the long-term prediction of daily time series of solar and geomagnetic activity indices. We then briefly describe the characteristics of singular spectrum analysis and FD models as the extension of neurofuzzy models. Finally, we provide a short description of the procedure of applying the proposed combined algorithm to solar and geomagnetic activity time series.

3.1 Problem statement

The ultimate goal of a space physics model is to be able to predict space weather extreme events. Relative to terrestrial forecasting, space weather forecasting is still in its infancy. One reason for this difference are the difficulties encountered in gathering space weather data. In comparison with terrestrial data, space weather data are sparse because there is only one point outside the magnetosphere (L1) and only several points inside the magnetosphere that record space weather data. In contrast, terrestrial data are measured every 6 h (about 10 different parameters) at 10^4 to 10^5 observing points, which are interpolated onto more than 10^6 points of a three-dimensional (3-D) grid used by the numerical prediction model. Consequently, modern space weather forecasters rely on a great variety of forecast systems, ranging from simple nonlinear models to complex information-based (empirical approaches) physical, and hybrid models (Bothmer and Daglis, 2007). Over the last two decades, which was when real-time data began to be available online, data-driven approaches, such as artificial neural networks (Bothmer and Daglis, 2007), neurofuzzy modeling (Gholipour *et al.*, 2005, 2007), Kalman filtering (Bothmer and Daglis, 2007), among others, have been shown to perform well in space weather forecasting. Space weather forecasting can be classified into five classes according to the forecasting frame-time:

- Nowcasting: 0–2 h
- Short-term forecasting: 2–36 h.
- Mid-term forecasting: 36–20 h.
- Intermediate-term forecasting: 5 days to several solar rotation
- Long-term forecasting: several solar rotations to the solar cycle (Tascione *et al.*, 1988; Bothmer and Daglis, 2007)

Each level of forecasting has its difficulties. Solar and geomagnetic activity indices that belong to the first three frame-times are very agitated time series, showing chaotic characteristics which vary over time (Mirmomeni and Lucas, 2009). Therefore, the long-term prediction of these solar and geomagnetic activity indices is very difficult. Note that the use of "long term" in this paper refers to the multiplicity

of the time steps and not to the standard time frame in the space weather forecasting categorization described above. This multiplicity of the time steps makes the forecasting difficult due to the difficulty in getting closer to the prediction possibility in the chaotic time series.

Various numerical prediction techniques have been used for geomagnetic activity forecasting, such as Fourier analyses, curve fitting, artificial intelligence, and neural networks (Gholipour *et al*., 2003). However, although these methods are suitable for making short-term predictions (here short term means one step or two steps ahead of the prediction), they are not reliable for the long-term predictions that are necessary for a powerful alarm/warning system. Prediction, which refers to a variety of mathematical methods for estimating parameters of a model beyond its proven validity domain, is usually tightly linked with the evolution of stochastic systems, which in turn demonstrate trends and periodicities hidden by the superimposed random noise.

In this paper, we consider the long-term prediction (several days ahead) of daily solar and geomagnetic activity indices that belong to the third frame-time category. The proposed long-term predictor is a data-driven approach based on spectral analysis and FD models as an extended version of regular neurofuzzy modeling (Mirmomeni *et al*., 2006).

3.2 Singular spectrum analysis

SSA has been defined as a new tool to extract information from short and noisy chaotic time series (Vautard *et al*., 1992). It relies on the Karhunen-Loeve decomposition of an estimate of covariance matrix based on M lagged copies of the time series. Thus, as the first step, the embedding procedure is applied to construct a sequence $\{\tilde{X}(t)\}$ of M-dimensional vectors from time series $\{X(t) : t = 1, \ldots, N\}$

$$\tilde{X}(t) = (X(t), X(t+1), \ldots, X(t+M-1)),$$
$$t = 1, \ldots, N', \quad N' = N - M + 1 \quad (5)$$

The $N' \times M$ trajectory matrix (D) of the time series has the M dimensional vectors as its columns. In the second step, the $M \times M$ covariance matrix C_X is calculated, and its eigenelements can be determined by singular value decomposition (SVD).

The eigenelements $\{(\lambda_k, \rho_k) : k = 1, \ldots, M\}$ of C_X are obtained from

$$C_X \rho_k = \lambda_k \rho_k \quad (6)$$

Each eigenvalue, λ_k estimates the partial variance in the direction of ρ_k, and the sum of all eigenvalues equals the total variance of the original time series. In the third step, the time series is projected onto each eigenvector, yielding the corresponding principal components.

$$A_k(t) = \sum_{j=1}^{M} X(t+j-1)\rho_k(j) \quad (7)$$

As the fourth step, the time series is reconstructed by combining the associated principal components

$$R_K(t) = \frac{1}{M_t} \sum_{k \in K} \sum_{j=L_t}^{U_t} A_k(t-j+1)\rho_k(j) \quad (8)$$

The normalization factor (M_t), and the lower (L_t) and upper (U_t) bounds of the reconstruction procedure differ for the center and edges of the time series and are defined by the following formula

$$(M_t, L_t, U_t) = \begin{cases} \left(\frac{1}{t}, 1, t\right), & 1 \leq t \leq M-1 \\ \left(\frac{1}{M}, 1, M\right), & M \leq t \leq N' \\ \left(\frac{1}{N-t+1}, t-N+M, M\right), & N'+1 \leq t \leq N \end{cases}$$
$$(9)$$

The singular spectrum plot (the logarithmic scale plot of singular values of the covariance matrix in decreasing order) can be used to enhance the SNR. The principal components related to the lower singular values can be omitted in the reconstruction stage to obtain adaptive noise cancellation. On the one hand, if all the components are used in reconstructing the time series, no information is lost. On the other hand, using the components that have characteristics like noise, which cause the components to be unpredictable over the long term, reduce the performance of the long-term predictor. Therefore, it is better to consider these components (which provide a little information on the original time series) as noise—and not use them for prediction purposes. (For further details, refer to Vautard *et al*. (1992) and Mirmomeni *et al*. (2006, 2007)).

3.3 Structure of FD models

Here, we consider the mathematical formulation of FD models (Taniguchi *et al*., 1999a, b). The fundamental approach with such systems is to divide the input space into small linear subspace with FD functions and their appropriate linear descriptor systems. The FD model is defined as

$$\text{Rule } i: \text{If } z_1(t) \text{ is } M_{1i} \text{ and } \ldots \text{ and } z_n(t) \text{ is } M_{ni}$$
$$\text{Then: } \begin{cases} E_i \dot{x}(t) = A_i x(t) + B_i u(t) \\ y(t) = C_i x(t) \end{cases}$$
$$(10)$$

where $x(t) \in R^n$, $y \in R^p$, $u(t) \in R^m$. M_{ij}, is the fuzzy set. $x(t) \in R^n$ is the state vector, $u(t) \in R^m$ is the input vector, $y \in R^p$ is the output vector, $E_k \in R^{n \times n}$, $A_i \in R^{n \times n}$, $B_i \in R^{n \times m}$, and $C_i \in R^{p \times n}$. $z_1(t) \sim z_n(t)$ are the premise variables.

The overall fuzzy model is achieved by fuzzy 'blending' of the linear descriptor subsystems. Given a pair of $(x(t), u(t))$, the final output of the fuzzy system is inferred as follows:

$$\sum_{i=1}^{M} \phi_i(z(t))E_i \dot{x}(t) = \sum_{i=1}^{M} \phi_i(z(t))(A_i x(t) + B_i u(t))$$
$$(11)$$
$$y(t) = \sum_{i=1}^{M} \phi_i(z(t))C_i x(t)$$

where M is the number of if-then rules. In Eq. (11), $\phi_i(z(t)) = \omega_i(z(t))/\sum_{i=1}^{M} \omega_i(z(t))$ where $\omega_i(z(t)) = \prod_{j=1}^{n} M_{ji}(z_j(t))$. $\phi_i(z(t))$ can be regarded as the normalized weight of each if-then rule. $M_{ji}(z_j(t))$ is the membership if $z_j(t)$ in M_{ji}. Defining $x^*(t) = [x^T(t) \; \dot{x}^T(t)]^T$, the
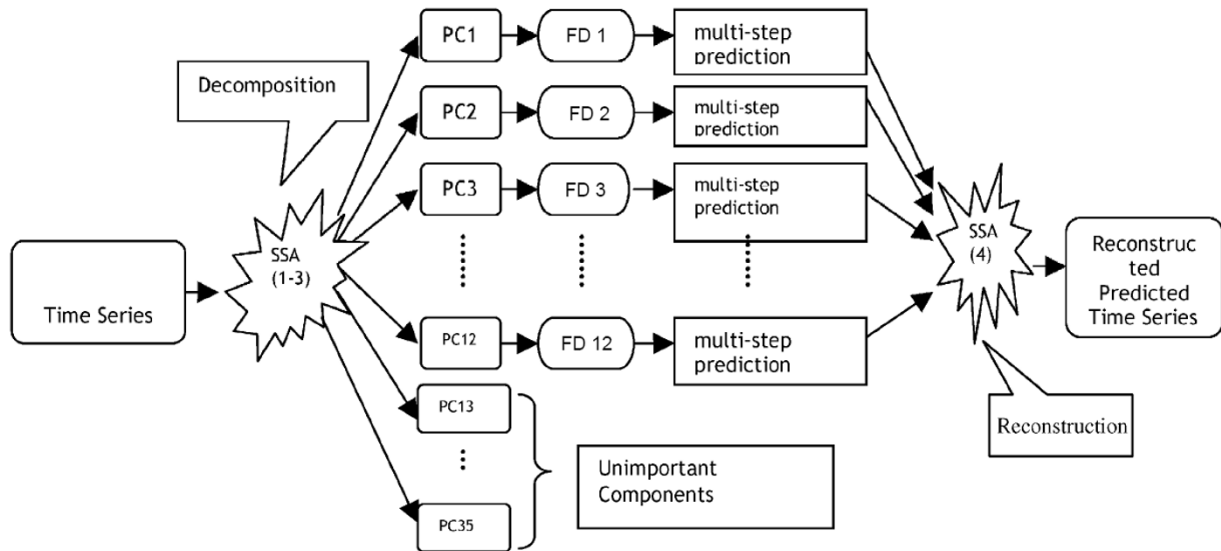
Fig. 2. The proposed method consists of four steps: (1) decomposition of the time series to nonlinear and periodic principal components by SSA stages 1–3; (2) FD modeling of the principal components; (3) multistep prediction of the reconstructed principal components; (4) reconstruction of the prediction of time series by combining the predicted principal components by SSA stage 4.

FD system (Eq. (11)) can be written as

$$E^* \dot{x}^*(t) = \sum_{i=1}^{M} \phi_i(z(t)) \left(A_i^* x^*(t) + B_i^* u(t) \right)$$
$$y(t) = \sum_{i=1}^{M} \phi_i(z(t)) C_i^* x^*(t)$$

(12)

where

$$E^* = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad A_i^* = \begin{bmatrix} 0 & I \\ A_i & -E_i \end{bmatrix}$$
$$B_i^* = \begin{bmatrix} 0 \\ B_i \end{bmatrix}, \quad C_i^* = \begin{bmatrix} C_i & 0 \end{bmatrix}$$

In this paper, the validity functions are chosen as normalized Gaussians. When validity functions are chosen as normalized Gaussians, it is possible to split the input space into only two parameters: the mean and the variance for each Gaussian. Mean parameters are tuned easily by choosing the center of a special region for the mean of the normalized Gaussian, as described in next section, and the variance parameters have to be tuned on the basis of intuition. Therefore, the use of such validity functions increases the simplicity of the algorithm, which is the ultimate goal of the proposed algorithm in this paper. In addition, it has to be noted that the use of other forms for validity functions that are used in different kind of locally linear models (Klir and Folger, 1988) is not ruled out.

Each Gaussian validity function has two parameters, center c_{ij} and standard deviation σ_{ij}. There are $M.m$ parameters for the nonlinear hidden layer. In addition, in this paper, the FD model is used to predict solar activity in terms of natural chaotic dynamics. Therefore, in this application, the dynamic system does not have control input $u(t)$. The

FD model for prediction application can be written as

$$E^* \dot{x}^*(t) = \sum_{i=1}^{r} \phi_i(z(t)) \left(A_i^* x^*(t) \right)$$
$$y(t) = \sum_{i=1}^{r} \phi_i(z(t)) C_i^* x^*(t)$$

(13)

3.4 Combined (FD/SSA) long-term prediction method

Figure 2 shows the block diagram of the proposed long-term prediction method. As can be seen, we first apply SSA to the chaotic solar and geomagnetic activity indices with time-varying characteristics that vary rapidly, especially in storms and sub-storms, with the aim of eliciting the main patterns of the original time series. We then use the selected principal components related to higher singular values (includes the main patterns of the original time series) for long-term prediction because of their simpler structures in comparison with the original time series. To do so, we train a FD model for each selected principal component for the long-term prediction of these simple patterns of the original time series. Optimization of the FD models for each of the principal components is obtained using an incremental learning algorithm on separate training and validation sets. The trained locally linear models are used as long-term predictor models for the corresponding components, and the long-term prediction of solar and geomagnetic activity indices is simply reconstructed from the predicted principal components by the fourth stage of the SSA algorithm. The total model consists of a relatively large number of parameters, but these parameters are optimized independently for the principal components, and over-parameterization is avoided by considering the error indices on separate validation sets. Therefore, all of the boxes denoted FD in Fig. 2 is a complete trainable predictor model and has been validated statistically through the termination condition of its learning procedure.

4. Learning Methodologies

This section is devoted to describing the new learning method for FD models in order to be able to adjust its two kinds of parameters. As mentioned earlier, the consequent part of a FD model is a linear descriptor subsystem that on its own is an improper system. An improper system is a dynamic system in which the order of the numerator of its transfer function is greater than the order of the denominator of its transfer function. In other words, an improper system is a non-causal system, which means that the output of the system is related to the future input of the system (Campbell, 1980). Therefore, to adjust the parameters of the consequent parts, an identification method is needed that is appropriate for improper systems. Unfortunately, there is no such a method to identify the parameters of a descriptor system. Therefore, in this study, we first of all attempted to develop a method that could identify the parameters of descriptor systems. This new method is described below.

4.1 A quasi-static algorithm to identify the parameters of a linear descriptor subsystem

Let us consider a linear descriptor system, such as

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{14}$$

The output and input under zero initial conditions (i.e., $x(0_-) = 0$) are related by the transfer function, $G(s)$, as follows:

$$G(s) = C(sE - A)^{-1}B \tag{15}$$

It is found that the transfer function $G(s)$ may no longer be strictly proper, in which case it may be written as the sum of a strictly proper part, $\overline{G}(s)$, and a polynomial part, $D(s)$. Therefore, we have:

$$G(s) = \overline{G}(s) + D(s) \tag{16}$$

where

$$\overline{G}(s) = \overline{C}\left(sI - \overline{A}\right)^{-1}\overline{B}, \quad \text{strictly proper} \tag{17a}$$

and

$$\begin{aligned} D(s) &= \tilde{C}\left(I - s\tilde{E}\right)^{-1}\tilde{B} \\ &= \tilde{C}\left(I + s\tilde{E} + \cdots + s^v\tilde{E}^v\right)\tilde{B}, \quad \text{polynomial} \end{aligned} \tag{17b}$$

$\overline{A}$, $\overline{B}$, $\overline{C}$, $\tilde{B}$, $\tilde{C}$, and $\tilde{E}$ result from a restricted standard equivalence of the system (Eq. (14)). Here, v is less than
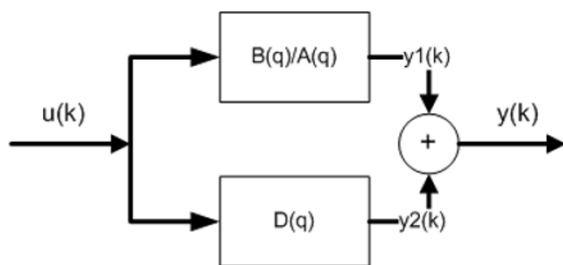

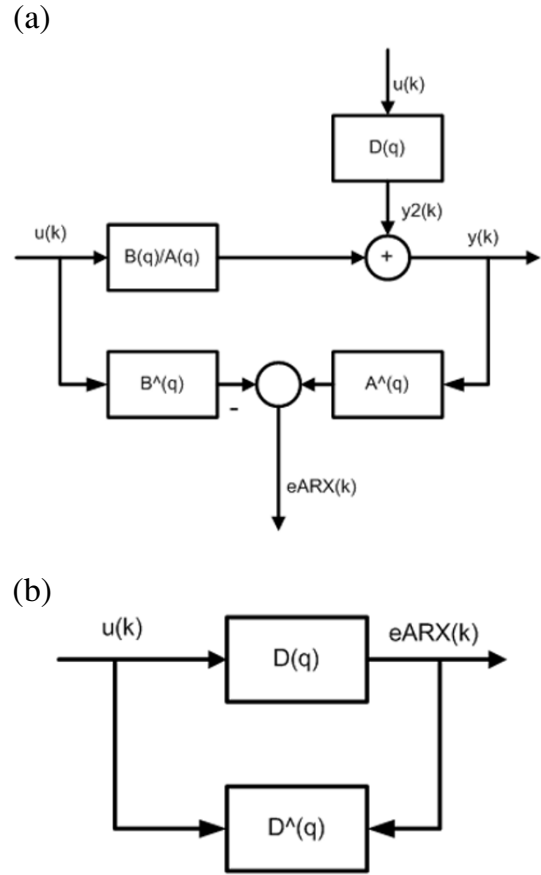Fig. 3. A decoupled descriptor system.

(a)

(b)

Fig. 4. Descriptor system identification by decoupling method. (a) Approximating the parameters of the strictly proper subsystem, (b) approximating the parameters of the polynomial part after identifying the strictly proper subsystem.

the size of $\tilde{E}$, since $\tilde{E}$ is *nilpotent* (i.e., all eigenvalues = 0) (Verghese *et al.*, 1981). Figure 3 shows a linear descriptor system that is decoupled to a strictly proper subsystem and a polynomial part (in discrete domain).

It has to be emphasized that the variable k in Fig. 3 (and in the following equations) is a discrete variable that demonstrates the sampling time (in comparison with the continuous variable t, which demonstrates the time). It is obvious that the polynomial subsystem in a discrete domain will be a moving average subsystem. Fortunately, we were able to identify each subsystem using classical identification methods. Therefore, it was possible to adjust the parameters of a descriptor system by decoupling it into two subsystems, and then adjust these parameters simultaneously. The quasi-static algorithm to identify parameters of linear descriptor system is as follows:

- To identify the parameters of the strictly proper subsystem, consider the output of the polynomial part (which has not been identified yet) as a measurement noise to the strictly proper part.
- Estimate an ARX (Auto-Regressive with eXogenous input) model $A(q)y(k) = B(q)u(k) + y_2(k)$ from the data $\{\underline{u}(k), \underline{y}(k)\}$ by

$$\hat{\theta}_{\text{ARX}} = \left(\underline{X}^T\underline{X}\right)^{-1}\underline{X}^T\underline{y} \tag{18}$$

- Calculate the prediction error of this ARX model

$$e_{\text{ARX}}(k) = \hat{A}(q)y(k) - \hat{B}(q)u(k) \qquad (19)$$

whose $\hat{A}(q)$ and $\hat{B}(q)$ are determined by $\hat{\theta}_{\text{ARX}}$.
- To identify the parameters of the polynomial part, consider the output of the strictly proper part (which is identified in this iteration) as a measurement noise to the polynomial part.
- Estimate the d_i parameters of the following FIR (Finite Impulse Response) model by well-known least squares technique (Ljung, 1987; Nelles, 2001)

$$e_{\text{ARX}}(k) = D(q)u(k) \qquad (20)$$

This algorithm can be iterated until the convergence is reached. This quasi-static algorithm is given in Fig. 4.

This method yields the linear descriptor systems in the frequency domain. This model has to be converted to the state space model in order to be useful in FD models. This can be done using the Silverman-Ho algorithm, which obtains the state space model of the descriptor system through its transfer function (Dai, 1989; Wang *et al.*, 2004).

4.2 GLoLiMoT learning algorithm

In this subsection, the GLoLiMoT algorithm is introduced to adjust the parameters of both the validity functions and the locally linear descriptor systems (Halfmann *et al.*, 1999; Nelles, 2001).

This algorithm is simple and intuitive, but to achieve a good performance one should tune some of the parameters, such as the "splitting ratio" and "standard deviation". The number of neurons is also important and should be optimized to attain the most accurate model with maximum generalization and to avoid over-fitting. A model over-fits if it fits the particularities of the training set (noise, bias, etc). Models which lose the generalization property show a low training error and high testing error. In other words, the operational performance measure for the trained model is the error on future data outside the training set, also known as the generalization error. This error may be undesirably large when, for example, the size of the available training data set is too small compared to the size of the model parameter set (Karystinos and Pados, 2000). Therefore, a complex model has more possibilities to over-fit data. In the literature on artificial neural networks and nonlinear system identification, the decrease in the generalization error exhibited during the first few successive passes through the same set of examples (usually called epochs) may be followed by a steady increase. This phenomenon is usually referred to as "over-fitting" (Brown and Harris, 1995; Karystinos and Pados, 2000; Nelles, 2001). One approach to avoiding over-fitting is to divide the available data set into three subsets: the training data set, the testing data set, and the validation data set. By using the validation data set to estimate the generalization error, it is then possible to detect the over-fitting. In this paper, an appropriate fitness function or generalization error index is calculated on multiple validation sets during training, and when the average error index starts to increase, the algorithm is terminated to prevent over-fitting.

The splitting ratio determines how one locally linear region should be divided into two new regions.

The GLoLiMoT algorithm is described in five steps:

1. Start with an initial model: start with a single LLDM (Locally Linear Descriptor Model), which is a global linear model over the whole input space with $\phi_1(\underline{z}) = 1$ and set $M = 1$.

2. Find the worst LLDM: Calculate a local loss function, such as the MSE (mean squared error) for each of the $i = 1, ..., M$ LLDMs, and find the worst performing LLDM.

3. Check all divisions: The worst LLDM is considered for further refinement. Divisions in all dimensions are tried, and following steps are carried out:

a. construction of the multi-dimensional membership functions for both generated hyper rectangles;
b. construction of all validity functions;
c. system identification of linear descriptor systems for both generated hyper rectangles by decoupling the method introduced in the previous subsection;
d. construction of a new FD system according to new linear descriptor systems in the state space form, which is produced by Silverman-Ho algorithm;
e. calculations of the loss function for current overall model.

4. Find the best division: The best of the p alternatives checked in step 3 is selected, and the related validity functions and LLDMs are then constructed. The number of LLDM neurons is incremented to $M = M + 1$.

5. Test the termination condition: If the termination condition is met, then stop, otherwise go to step 2.

This algorithm overcomes the curse of dimensionality (which refers to the exponential growth of the hypervolume as a function of dimensionality (Bellman, 1961)) and provides a fast and efficient training procedure when there are restrictions in computational resources and memory capacity. The curse of dimensionality causes networks with lots of irrelevant inputs to be behave relatively badly: the dimension of the input space is high, and the network uses almost all of its resources to represent irrelevant portions of the space. Even if there is a network algorithm that is able to focus on important portions of the input space, the higher the dimensionality of the input space, the more data may be needed to find out what is important and what is not. Therefore, the use of the GLoLiMoT algorithm with this property can be useful for modeling solar and geomagnetic storms for which the number of data sets is limited.

For further details, refer to Mirmomeni *et al.* (2006).

5. Simulation Results

Among the many solar and geomagnetic activity indices, those of the disturbance storm time (D_{st}), geomagnetic *aa*, and SWS are extremely important for forecasting the dangerous consequences of the solar and geomagnetic anomalies. All of these indices reflect the severity of the geomagnetic storm, and a knowledge of these in advance will assist satellite operators and some surface technologies avoid major disasters. Historical data for the training phase should contain the relevant measurements of the related physical parameters of solar flare and of unleashed coronal mass ejection (CME) and the shock of X-ray and gamma-ray radiation, radio bursts, and others. All mentioned parameters
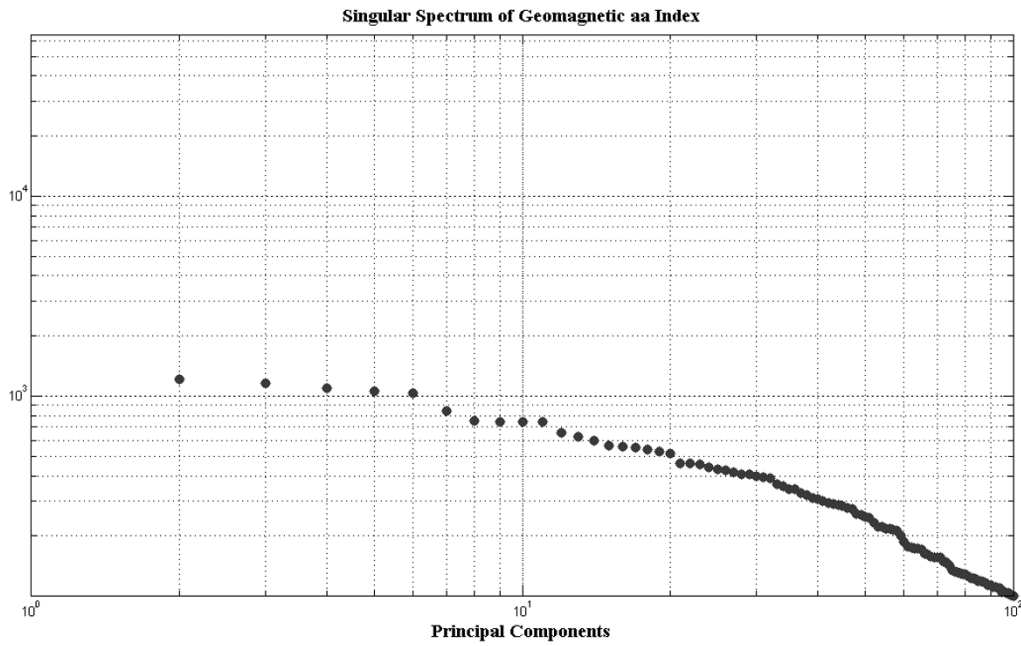
Fig. 5. Singular spectrum of the geomagnetic *aa* index time series by the SSA algorithm.
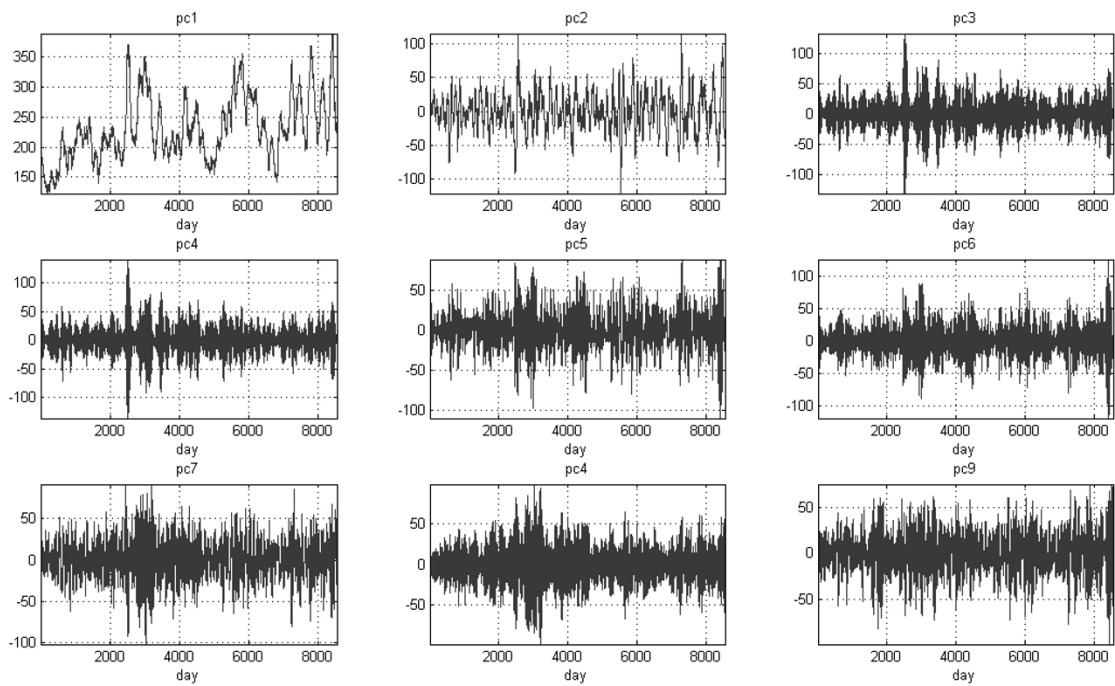

Fig. 6. The first nine principal components of the geomagnetic *aa* index time series obtained by the SSA algorithm.

should be measured after a violent solar explosion that triggers numerous physical processes, a part of which is responsible for the geomagnetic storm. Therefore, historical data for training the geomagnetic *aa* and SWS estimator should contain all of the mentioned parameters of solar flares initiating a significant or less significant geomagnetic storm (GMS).

In this section, we consider the long-term prediction of geomagnetic *aa* and SWS daily time series using a combined (FD/SSA) method. However, it is also possible to use this method to model and predict other geomagnetic in-

Table 1. The NMSE of the proposed FD/SSA model in comparison with the results obtained with some well-known models (MLP, RBF neural networks, and the LLNF model) in terms of predicting the daily average geomagnetic *aa* index.

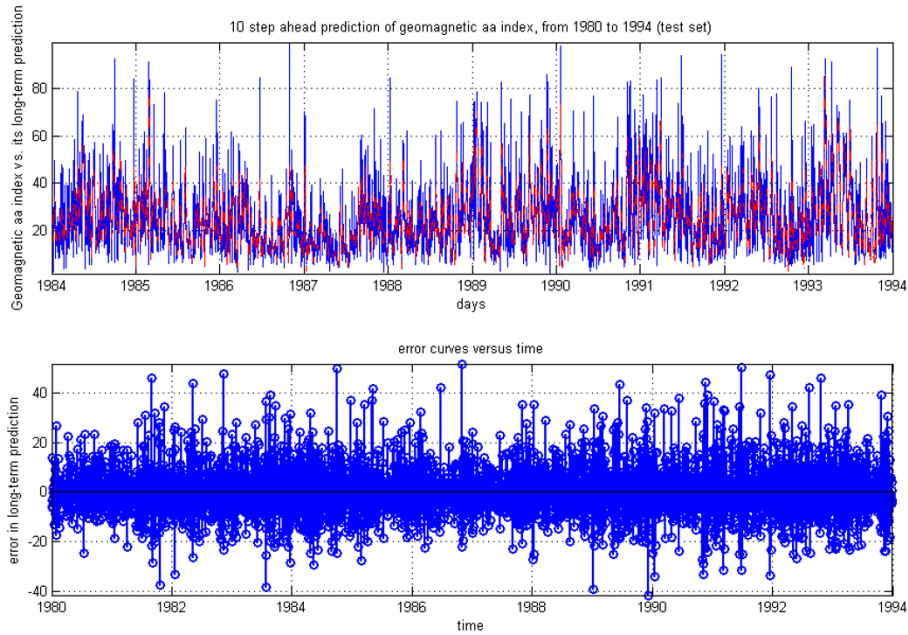| Method	NMSE in predicting geomagnetic *aa* index
MLP	0.1608
RBF	0.2103
LLNF	0.1204
FD/SSA	**0.0836**

Fig. 7. Ten-step-ahead prediction of the geomagnetic *aa* index between 1980 and 1994 by the proposed FD/SSA model. Upper: predicted and real values of test set; red curve shows the predicted *aa* index, and the blue curve shows the real *aa* index. Lower: prediction error.
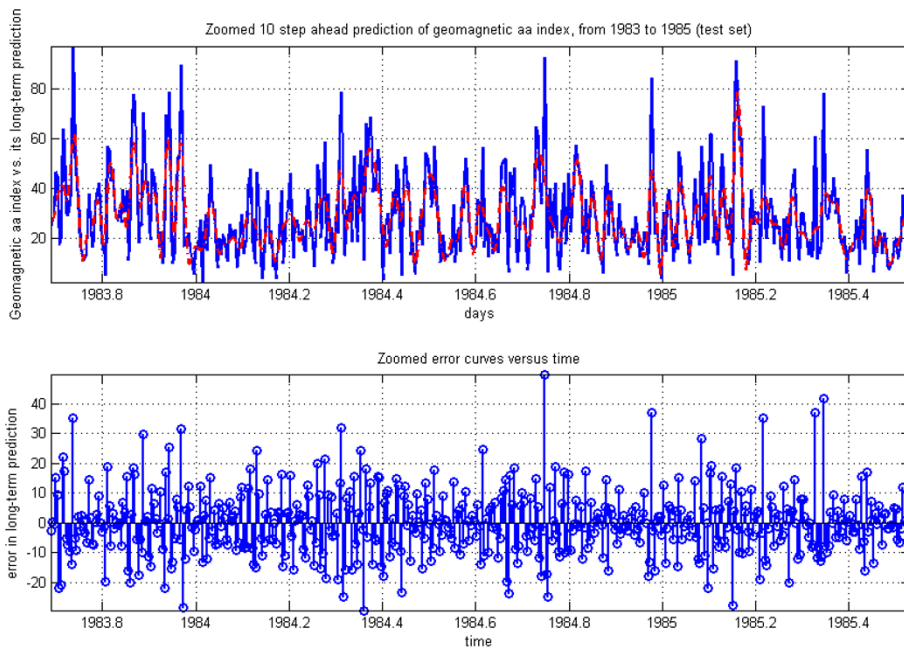

Fig. 8. Zoomed ten-step-ahead prediction of the geomagnetic *aa* index between 1980 and 1994 by the proposed FD/SSA model. Upper: predicted and real values of the test set (zoomed); red curve shows the predicted *aa* index, and the blue curve shows the real *aa* index. Lower: prediction error (zoomed).

dices, such as D_{st}. or AE, in other time scales ranging from minutes to hours.

After the principal components of the geomagnetic *aa* and SWS indices have been extracted by SSA, one FD model is automatically constructed for each of the chosen principal components. Here, we have attempted to predict the daily average of the geomagnetic *aa* index in the period 1963–1994, which contains several important events that are considered to be disastrous in terms of human technologies due to the effects of solar storms. For example, the

super-storm on 13 March 1989 shut down the power supply system in Québec, Canada, and the effects of the CME on 11 January 1997 caused the failure of Telstar 401 satellite.

Figure 5 shows the singular spectrums of the geomagnetic *aa* index. The window length is 100, and the first 20 components, related to the first 20 singular values, are chosen to predict and reconstruct the original geomagnetic *aa* time series. The first nine components are shown in Fig. 6 and the others, with singular spectrums smaller than 0.01 of the first principal component, are eliminated to enhance
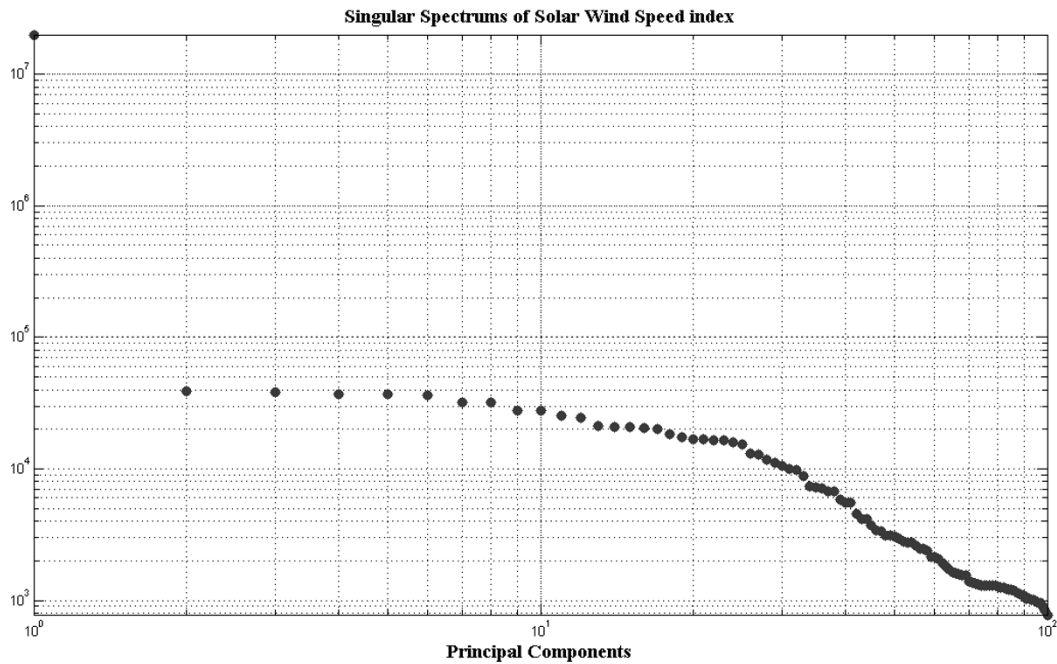
Fig. 9. Singular spectrum of the SWS index time series obtained using the SSA algorithm.
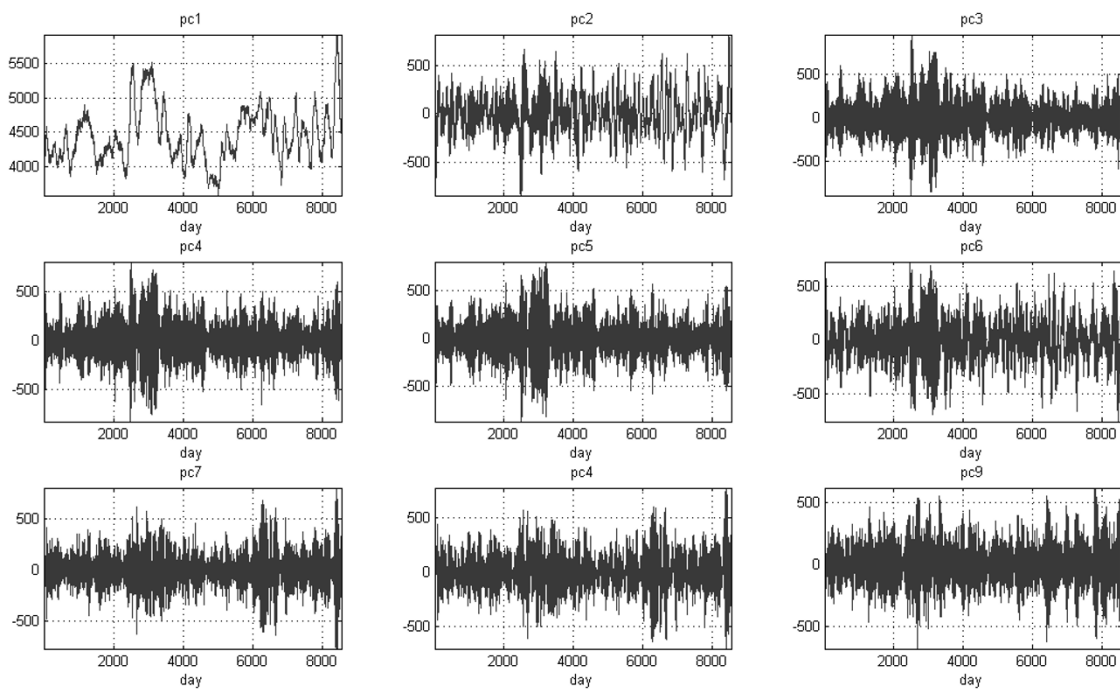

Fig. 10. The first nine principal components of the SWS index time series obtained using the SSA algorithm.

the SNR. A FD model is created via GLoLiMoT learning algorithm for each principal component.

About 4,850 data sets (from 1963 to 1980) are used to adjust the parameters of the FD models and about 3800 (from 1980 to 1994) data sets are kept to test the performance of the (FD/SSA) model in predicting the daily average of the geomagnetic *aa* index. To select the most informative inputs among the lags of each principal component, we apply the correlation analysis to each principal component, and the most correlated lags to the predicted value are chosen

as the inputs of each FD model. For most of the principal components, the validation error increases after eight runs of the GLoLiMoT algorithm. Therefore, for most of the principal components, a FD model with eight locally linear descriptor models is trained. Figures 7 and 8 show the performance of the proposed model in predicting the daily average of the geomagnetic *aa* index in test mode. It is clearly evident that the proposed combined model performs very well for the long-term prediction of the geomagnetic *aa* index. It has to be emphasized that some fluctuations
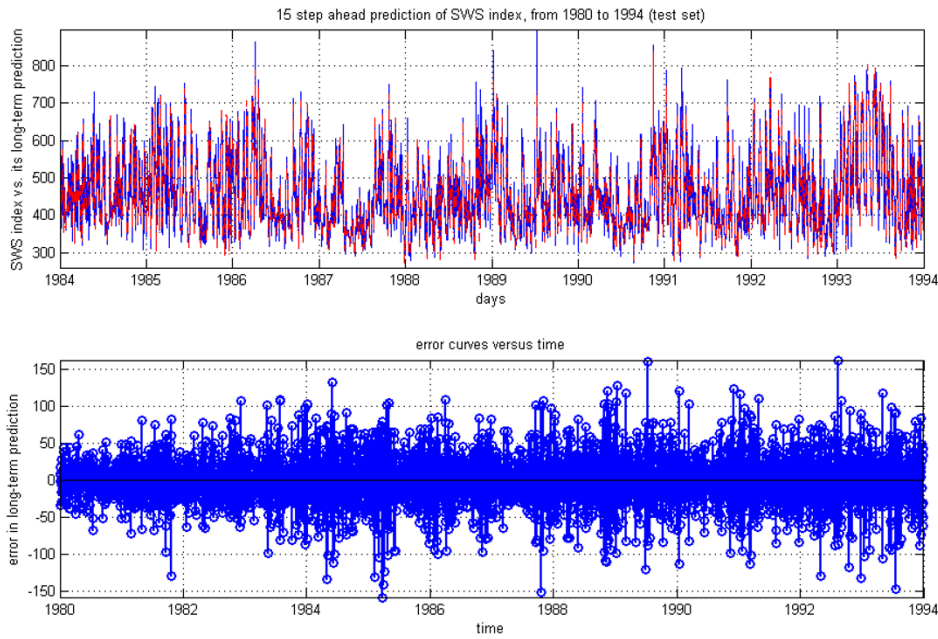
Fig. 11. Fifteen-step-ahead prediction of the SWS index between 1980 and 1994 by the proposed FD/SSA model. Upper: predicted and real values of the test set; red curve shows the predicted *aa* index, and the blue curve shows the real *aa* index. Lower: prediction error.
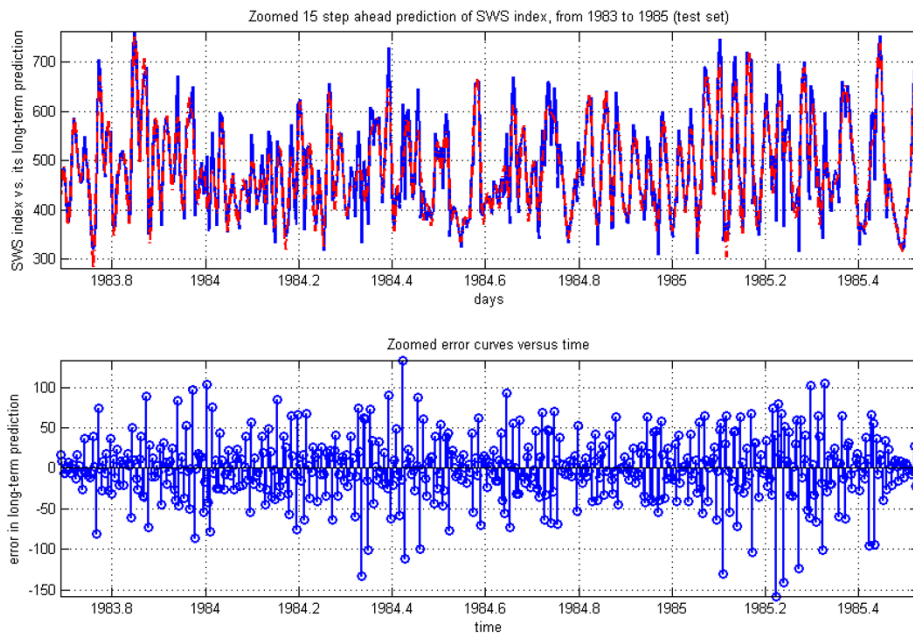

Fig. 12. Zoomed 15-step-ahead prediction of the SWS index between 1980 and 1994 by the proposed FD/SSA model. Upper: predicted and real values of the test set (zoomed); red curve shows the predicted *aa* index, and the blue curve shows the real *aa* index. Lower: prediction error (zoomed).

can be seen within a short period; these may caused by geomagnetic disturbances that did not exist in the training data. However, the overall performance of the proposed method is very good.

Table 1 contains the results of several methods that are used for the long-term prediction of the daily geomagnetic *aa* index: the RBF (radial basis function), MLP (multilayered perception), and LLNF (Locally Linear Neurofuzzy Model) networks. It is once again evident that the performance of the FD/SSA model is superior to that of the other well-known methods tested here. The normalized mean

square error (NMSE) is used as the error index in Table 1 and is defined as

$$
\text{NMSE} = \left(\frac{\sum\limits_{i=1}^{n} (y - \hat{y})^2}{\sum\limits_{i=1}^{n} (y - \overline{y})^2} \right) \tag{21}
$$

y, $\hat{y}$, and $\overline{y}$ are observed data, the predicted data, and the average of observed data respectively.

The same approach used for the geomagnetic *aa* index

was also used in an attempt to predict the daily average of the SWS index between 1963 and 1994.

Figure 9 shows the singular spectrums of the SWS index. As in the first analysis, the window length is 100, and the first 13 components, related to the first 13 singular values, are chosen to predict and reconstruct the original SWS time series. The first nine components are shown in Fig. 10 and the others, with singular spectrums smaller than 0.01 of the first principal component, are eliminated to enhance the SNR. Again, for each principal component, the GLoLiMoT algorithm is applied to give a FD for each principal component.

About 4,850 data sets (from 1963 to 1980) are used to adjust the parameters of the FD/SSA models and about 3,800 data sets (from 1980 to 1994) are kept to test the performance of the FD/SSA model in predicting the daily average of the SWS index. Again, to select the most informative inputs among the lags of each principal component of SWS index, we applied the correlation analysis to each principal component and then chose the most correlated lags to the predicted value as the inputs of each FD model. It should be noted that there are better approaches for input selection, such as the mutual information tool, which gives the most informative lags and which are nonlinearly related to the predicted target; however, for simplicity, we used correlation analysis for input selection. For most of the principal components, the validation error increase after 12 runs of the GLoLiMoT algorithm. Therefore, for most of the principal components, a FD with 12 locally linear descriptor models is trained. Figures 11 and 12 depict the performance of the proposed model in predicting the daily average of the SWS index in test mode. It is evident that the performance of the proposed model in the long-term prediction of the SWS index as well as its performance for *aa* index is very good. The proposed combined model can clearly track the rapid and large fluctuations in the SWS index because the SWS data set is very rich and contains many rapid fluctuations that are used in training.

6. Discussion and Conclusions

The long-term prediction of natural phenomena with a limited number of observations is usually difficult. The reconstruction of dynamics is restricted by limitations in knowledge on physical processes and by the fact that the most powerful tools of chaotic modeling, neural networks and neurofuzzy models, can only be used in short-term predictions. Here, we have used a combined model based on FD models and singular spectrum analysis (FD/SSA) for the long-term prediction of the daily time series of two important solar and geomagnetic activity indices: geomagnetic *aa* and solar wind speed. Geomagnetic *aa* and SWS indices were used in the daily average time scale to depict the performance of the proposed combined long-term predictor (which is used for several step-ahead prediction of these solar and geomagnetic activity indices) in modeling the short-term behavior (daily frame time) of solar and geomagnetic activity. It must be noted that this method can be applied to model other geomagnetic indices, such as D_{st} or AE in other time scales, ranging from minutes to hours. Although forecasting based only on the time series of the

parameter to be predicted has limited success, incorporation of the measurements linked to the forecasted parameter significantly improves the accuracy of the prediction. By optimizing the number of neurons, the splitting ratio, and the standard deviations, an accurate prediction is provided. Due to its high generalization and low prediction error, this method can also be used in predicting the solar and geomagnetic activity several days (or even several years if the original index be a yearly index) in advance of other solar activity indices.

References

Bellman, R., *Adaptive Control Processes: A Guided Tour*, Princeton University Press, 1961.

Bothmer, V. and I. Daglis, *Space Weather Physics and Effects*, Springer Praxis Publishing, Chichester, UK, 2007.

Brown, G. M., The peak of solar cycle 22: predictions in retrospect, *Ann. Geophys.*, **10**, 453–470, 1992.

Brown, M. and C. Harris, *Neurofuzzy Adaptive Modeling and Control*, Prentice Hall, UK, 1995.

Campbell, S. L., *Singular Systems of Differential Equation*, Pitman, London, 1980.

Dai, L., *Singular Control Systems*, Springer, New York, 1989.

Dziurla, B. and R. W. Newcomb, The Drazien inverse and semi-state equations, in *Proc. 4th Int. Symp. on Math. Theory of Networks and Systems. Delft, Netherland*, pp. 283–289, 1979.

Gholipour, A., A. Abbaspour, B. N. Araabi, and C. Lucas, Enhancements in the prediction of solar activity by locally linear model tree, paper presented at *22nd IASTED Int. Con. on Modeling, Identification, and Control, Innsbruck, Austria*, 2003.

Gholipour, A., C. Lucas, B. N. Araabi, and M. Shafiee, Solar activity forecast: spectral analysis and neurofuzzy prediction, *J. Atmos. Sol.-Terr. Phys.*, **67**, 595–603, 2005.

Gholipour, A., C. Lucas, B. N. Araabi, M. Mirmomeni, and M. Shafiee, Extracting the main patterns of natural time series for long-term neurofuzzy prediction, *J. Neural Comput. Appl.*, **16**(4–5), 383–393, 2007.

Halfmann, C., O. Nelles, and H. Holzmann, Modeling and identification of the vehicle suspension characteristics using local linear model trees, in *Proc. of IEEE, Int. Conf. on Control Applications*, Khala-Coast island, Hawaii, USA, pp. 1484–1489, 1999.

Haykin, S. (editor), *Unsupervised Neural networks: A Comprehensive Foundation*, Macmillan, New York, 1994.

Joselyn, J. A. *et al.*, Panel achieves consensus prediction of solar cycle, *Eos Trans. AGU*, **78**, 211–212, 1997.

Kang, J. L. and W. S. Tang, Minimum energy control of 2-D singular systems with constrained control, in *Proc. of the 4th Int. Conference on Machine Learning and Cybernetics*, August 18–21, Guangzhou, 2005.

Karystinos, G. N. and A. Pados, On overfitting, generalization, and randomly expanded training sets, *IEEE Trans. Neural Networks*, **11**(5), 1050–1057, 2000.

Klir, G. J. and T. A. Folger, *Fuzzy sets, uncertainty and information*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

Leung, H., T. Lo, and S. Wang, Prediction of noisy chaotic time series using an optimal radial basis function neural network, *IEEE Trans. Neural Networks*, **12**(5), 1163–1172, 2001.

Lewis, Z. V., On the apparent randomness of substorm onsets, *Geophys. Res. Lett.*, **18**, 1627, 1991.

Lillekjendlie, B., D. Kugimutzis, and N. Christophersen, Chaotic time series, part II: System identification and prediction, *Identification Control*, **15**, 225–243, 1994.

Ljung, L., *System Identification: Theory for the User*, New Jersey, Prentice-Hall, 1987.

Lu, G. and D. W. C. Ho, Generalized quadratic stabilization for perturbated discrete-time singular systems with delayed state, in *4th Int. Conf. on Control and Automation (ICCA'03)*, June 10–12, Montreal, Canada, 2003.

Lucas, C., A. Abbaspour, A. Gholipour, B. N. Araabi, and M. Fatourechi, Enhancing the performance of neurofuzzy predictors by emotional learning algorithm, *Informatica*, **27**(2), 165–174, 2003.

Luenberger, D. G., Dynamic equation in descriptor form, *IEEE Trans. Auto. Contr.*, **AC-22**, 132–321, 1977.

Meng, B. and J. F. Zhang, Reachability conditions for switched linear singular systems, *IEEE Trans. Automatic Control*, **51**(3), 482–488, 2006.

Mirmomeni, M. and C. Lucas, Analyzing the variation of embedding dimension of solar and geomagnetic activity indices during geomagnetic storm time, *Earth Planets Space*, **61**, 237–247, 2009.

Mirmomeni, M. and M. Shafiee, State analysis of discrete-time singular nonlinear systems using a fuzzy neural network, in *13th Iranian Conference on Electrical Engineering (ICEE'05)*, Zanjan University, Zanjan, Iran, 2005a.

Mirmomeni, M. and M. Shafiee, State analysis of time-invariant singular systems via Haar wavelet, in *13th Iranian Conference on Electrical Engineering (ICEE'05)*, Zanjan University, Zanjan, Iran, 2005b.

Mirmomeni, M., M. Shafiee, C. Lucas, and B. N. Araabi, Introducing a new learning method for fuzzy descriptor systems with the aid of spectral analysis to forecast solar activity, *J. Atmos. Sol.-Terr. Phys.*, **68**, 2061–2074, 2006.

Mirmomeni, M., C. Lucas, and B. Moshiri, Long term prediction of chaotic time series with the aid of neuro fuzzy models, spectral analysis and correlation analysis, *IEEE Int. Joint Conference on Neural Networks*, Orlando, Florida, USA, 2007.

Nelles, O., *Nonlinear System Identification*, Springer, Berlin, 2001.

Newcomb, R. W., The semi-state description of nonlinear time-variable circuits, *IEEE Trans. Circuits Systems*, **28**, 62–71, 1981.

Pandolfi, L., On the Regulator Problem for Linear Degenerate Control Systems, *J. Optimiz. Theory Applic.*, **33**, 141–154, 1981.

Raouf, J. and E. K. Boukas, Observer-based controller design for linear singular systems with markovian switching, in *43rd IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas, 2004.

Saunders, P. T., *An Introduction to Catastrophe Theory*, Cambridge University Press, Cambridge, 1980.

Schatten, K. H. and W. D. Pesnell, An early solar dynamo prediction: cycle 23-cycle 22, *Geophy. Res. Lett.*, **20**, 2257–2278, 1993.

Schatten, K. H. and S. Sofia, Forecast of an exceptionally large even numbered solar cycle, *Geophy. Res. Lett.*, **14**, 632, 1987.

Schatten, K. H., P. H. Scherrer, L. Svalgaard, and J. M. Wilcox, Using Dynamo theory to predict the sunspot number during solar cycle 21, *Geophy. Res. Lett.*, **5**, 411, 1978.

Schatten, K. H., D. J. Myers, and S. Sofia, Solar activity forecast for solar cycle 23, *Geophy. Res. Lett.*, **23**(6), 605–608, 1996.

Shafiee, M. and S. Amani, Optimal control for a class of singular systems using neural network, *Iranian J. Sci. Technol. Trans. B, Eng.*, **29**(B1), Shiraz, Iran, 2005.

Sitnov, M. I., A. S. Sharma, and K. Papadopoulos, Modeling substorm dynamics of the magnetosphere: From self-organization and self-organized criticality to non- equilibrium phase transitions, *Phys. Rev. E*, **65**, 106–116, 2001.

Sofia, S., P. Fox, and K. H. Schatten, Forecast update for activity cycle 23 from a Dynamo-based method, *Geophys. Res. Lett.*, **25**(22), 4149–4152, 1998.

Taniguchi, T., K. Tanaka, K. Yamafuji, and H. O. Wang, Fuzzy descriptor systems: Stability analysis and design via LMIs, in *Proc. American Control Conference*, pp. 1827–1831, 1999a.

Taniguchi, T., K. Tanaka, and H. O. Wang, Fuzzy descriptor systems and fuzzy controller designs, in *Proc. of the 8th International Fuzzy System Association World Congress*, pp. 655–659, 1999b.

Taniguchi, T., K. Tanaka, and H. O. Wang, Fuzzy descriptor systems and nonlinear model following control, *IEEE Trans. Fuzzy Systems*, **8**(4), 442–452, 2000.

Tascione, T. F., H. W. Kroehl, R. Creiger, J. W. Freeman Jr., R. A. Wolf, R. W. Spiro, R. V. Hilmer, J. W. Shade, and B. A. Hausman, New ionospheric and mangnetospheric specification models, *Radio Sci.*, **23**, 211–222, 1988.

Thompson, R., A technique for predicting the amplitude of solar cycle, *Sol. Phys.*, **148**, 383, 1993.

Tong, H., *Nonlinear time series: A dynamical system approach*, Oxford University Press, 1996.

Tong, H. and K. Lim, Threshold autoregressive limit cycles and cyclical data, *J. R. Stat. Soc. B*, **42**, 245–292, 1980.

Uluyol, O., M. Ragheb, and S. R. Ray, Local output gamma feedback neural network, *Proc. IEEE Int. Conf. on Neural Networks: IJCNN*, **1**, 337–342, 1998.

Vassiliadis, D., A. J. Klimas, J. A. Valdivia, and D. N. Baker, The nonlinear dynamics of space weather, *Space Weather: Phys. Applic.*, **26**, 197–207, 2000.

Vautard, R. and M. Ghil, Singular spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series, *Physica D*, **35**, 395–424, 1989.

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

Verghese, G. C., B. C. Levy, and T. Kailath, A generalized state-space for singular systems, *IEEE Trans. Auto. Contr.*, **AC-26**(4), 1981.

Wang, J., Q. Zhang, W. Liu, X. Xin, and V. Sreeram, H$^\infty$ model reduction for singular systems, in *Proc. of 2004 American Control Conference*, June 30–July 2, Boston, Massachusetts, pp. 119–124, 2004.

Weigend, A., B. H. Berman, and D. Rumelhart, Predicting sunspots and exchange rates with connectionist networks, *Nonlinear Model. Forecasting J.*, Addison-Wesley, 395–432, 1992.

Xiaoping, L., Solvability of nonlinear singular systems-part II: the case with inputs, in *Proc. of the American Control Conference*, Seattle, Washington, 1995.

Yonchev, A., R. Findeisen, C. Ebenbauer, and F. Allgöower, Model predictive control of linear continuous time singular systems subject to input constraints, in *43rd IEEE Conference on Decision and Control*, December 14–17, Atlantis, Paradise Island, Bahamas, 2004.

Zeeman, E. C., *Catastrophe Theory: Selected Papers: 1972-1977*, Addison-Wesely Publishing Co., Reading, 1987.

M. Mirmomeni (e-mail: mirmomeni@ut.ac.ir), E. Kamaliha, M. Shafiee, and C. Lucas