AGU PUBLICATIONS

Space Weather

RESEARCH ARTICLE
10.1002/2017SW001608

Key Points:
- For the first time, an ANN was optimized with PSO to forecast the *Dst* index
- A forecast model was built using only prior values of *Dst*
- The model performed well for forecasts of up to 3 h

Supporting Information:
- Supporting Information S1

Correspondence to:
J. A. Lazzús,
jalazzus@gmail.com;
jlazzus@dfuls.cl

Citation:
Lazzús, J. A., P. Vega, P. Rojas, and I. Salfate (2017), Forecasting the *Dst* index using a swarm-optimized neural network, *Space Weather*, *15*, 1068–1089, doi:10.1002/2017SW001608.

Received 4 FEB 2017
Accepted 7 AUG 2017
Accepted article online 14 AUG 2017
Published online 31 AUG 2017

Forecasting the *Dst* index using a swarm-optimized neural network

J. A. Lazzús[1], P. Vega[1], P. Rojas[1], and I. Salfate[1]

[1]Departamento de Física y Astronomía, Universidad de La Serena, La Serena, Chile

Abstract A hybrid technique that combines an artificial neural network with a particle swarm optimization (ANN+PSO) was used to forecast the disturbance storm time (*Dst*) index from 1 to 6 h ahead. Our ANN was optimized by PSO to update ANN weights and to predict the short-term *Dst* index using past values as input parameters. The database used contains 233,760 hourly data from 1 January 1990 to 31 August 2016, considering storms and quiet period, grouped into three data sets: learning set (with 116,880 hourly data points), validation set (with 58,440 data points), and testing set (with 58,440 data points). Several ANN topologies were studied, and the best architecture was determined by systematically adding neurons and evaluating the root-mean-square error (RMSE) and the correlation coefficient (*R*) during the training process. These results show that the hybrid algorithm is a powerful technique for forecasting the *Dst* index a short time in advance like $t + 1$ to $t + 3$, with RMSE from 3.5 nT to 7.5 nT, and *R* from 0.98 to 0.90. However, $t + 4$ to $t + 6$ predictions become slightly more uncertain, with RMSE from 8.8 nT to 10.9 nT, and *R* from 0.86 to 0.79. Additionally, an exhaustive analysis according to geomagnetic storm magnitude was conducted. In general, the results show that our hybrid algorithm can be correctly trained to forecast the *Dst* index with appropriate precision and that *Dst* past behavior significantly affects adequate training and predicting capabilities of the implemented ANN.

1. Introduction

Geomagnetic storms are perturbations of the Earth's magnetic field, lasting several hours or even days [*Gonzalez et al.*, 1999]. They are usually defined by ground-based, low-latitude geomagnetic field horizontal component (H) variations [*Sckopke*, 1966]. Several studies have been conducted to understand the interplanetary causes of geomagnetic storms; *Piddington* [1963], for example, proposed a relation between geomagnetic perturbations and the orientation and intensity of the interplanetary magnetic field. Many studies established that the southward component of the interplanetary magnetic field (IMF) presents the main cause of geomagnetic storms [e.g., *Burton et al.*, 1975; *Gonzalez et al.*, 1994; *Echer et al.*, 2005]. This magnetic field orientation allows magnetic reconnection [*Akasofu*, 1981] and energy transfer from the solar wind to the Earth's magnetosphere [*Echer et al.*, 2008]. This way, enhanced fluxes of particles get injected, producing a depression of the Earth's magnetic field horizontal (*H*) component caused by the diamagnetic effect which is generated by the azimuthal circulation of the ring current particles. This depression has the effect of intensifying the westward ring current circulating the Earth [*Gonzalez et al.*, 1994; *Kamide et al.*, 1998; *Echer et al.*, 2005], as evidenced by a decrease of the geomagnetic field measured at near-equatorial magnetic stations [*Sugiura and Kamei*, 1991]. Depending on the storm intensity, the edges of the ring current will move closer or farther away from the ground. The internal border of the ring current is placed at four Earth radius or less from the Earth's surface during severe storms. But when the storms lose intensity, the ring current is located farther from the Earth [*Gonzalez et al.*, 1994, 1999; *Echer et al.*, 2005, 2008] (for a historical analysis of geomagnetic storms, see *Lakhina and Tsurutani* [2016]).

On the other hand, various mechanisms have been proposed to explain geomagnetic storms as the result of an enhanced solar-magnetosphere energy coupling process. For instance, viscous interaction [*Axford and Hines*, 1961], magnetic reconnection [*Dungey*, 1961], and cross-field diffusion from the magnetosheath to the magnetopause boundary layer by resonant wave-particle interactions [*Tsurutani and Thorne*, 1982].

The influence of solar, ionospheric, and other sources render this geomagnetic activity highly variable [*Gonzalez et al.*, 1999], and thus geomagnetic indices come into use to characterize geomagnetic activity,

of which disturbance storm time (*Dst*) represents the most common index [*Rostoker*, 1972]. Additionally, these indices keep long data records that enable a statistical study of geomagnetic storms [*Burch*, 1974]. Thus, correlations obtained from these statistical studies, which relate the geomagnetic indices and their potential drivers, can provide the basis for the prediction of geomagnetic storms [*Joselyn*, 1995].

Dst index is calculated as an average of the magnetic disturbance of the *H* component variation, divided by the average of dipole low-latitude cosines of four magnetic observatories [*Sugiura*, 1964] that operate in the low-latitude regions around the geomagnetic equator, such as San Juan, SJG; Kakioka, KAK; Honolulu, HON; and Hermanus, HER (coordinates in *Sugiura and Kamei* [1991]). Generally, this index provides a quantitative measure of the variations of ground-based measurements of the magnetic field on the increase of energetic particles in the ring current during a geomagnetic storm [*Sugiura and Kamei*, 1991]. For the *Dst* index, a geomagnetic storm develops in three phases: the initial phase, the main phase, and the recovery phase [*Gonzalez et al.*, 1994]. During a geomagnetic storm, the *Dst* index value can decrease from close to zero to negative until reaching a minimum value. This process usually takes a few hours. From this moment, *Dst* index value turns less negative, requiring some days to recover values close to zero. This minimum value is considered as the magnitude for storm intensity. From the physical point of view, negative or positive disturbances in the *Dst* index reflect the westward/eastward drift of the energetic particles during the storm. Negative value results signal a weakened magnetic field of the Earth [*Karinen and Mursula*, 2005].

Therefore, to better understand the various forms of solar and geomagnetic activity, the forecast of space climate should take priority in space weather research. From these effects, the prediction of the geomagnetic disturbances in terms of *Dst* index can be used to describe the geomagnetic storm phenomenon and to differentiate geomagnetic storm phases [*Burton et al.*, 1975; *O'Brien and McPherron*, 2000; *Pallocchia et al.*, 2006]. Very intense geomagnetic storms and their effects can affect the space technology system and human activities [*Hapgood*, 2011], and increasing evidence suggests that geomagnetic field changes also affect biological systems [e.g., *Blanch et al.*, 2013]. But the main danger of geomagnetic storms is their potential to damage indispensable technological systems, such as satellites, power grids, gas, oil and water conduction systems, and transportation systems [e.g., *Kasinskii et al.*, 2007]. It has been reported that historical geomagnetic storm events such as the March 1989 and October 2003 involved global damages in ground-based infrastructures, causing significant economic losses [*Kappernman and Albertson*, 1990; *Veselovsky et al.*, 2004; *Kappenman*, 2005]. It is, therefore, important to have methods that predict geomagnetic storms and thus allow a quick response to such events [*Singh et al.*, 2010].

For decades, several methods have been proposed to forecast geomagnetic disturbances (see examples in *Baker* [1986]). Nowadays, a greater number of data-driven techniques for the prediction of geomagnetic storms making use of the *Dst* index are available [*Kamide et al.*, 1998; *Joselyn*, 1995]. Most of these methods deal with geomagnetic studies using cause-and-effect parameters [*Kugblenu et al.*, 1999]. *Burton et al.* [1975] introduced an empirical linear relationship for estimating the *Dst* index by using data such as solar wind velocity, density, and the southward component in the interplanetary magnetic field. On the other hand, *Iyemori et al.* [1979] applied a linear filtering prediction method for predicting the *Dst* index. Yet methods based on the supposition that the solar wind-magnetosphere interaction follows a linear behavior are inadequate for predicting the *Dst* index. Thus, the use of nonlinear models could provide the possible nonlinear response which these interactions would require [*Kugblenu et al.*, 1999; *Sharifi et al.*, 2006a].

Over the last decades, several heuristic algorithm types have been developed to solve problems related to nonlinear systems. The main exponent of these algorithms is called artificial neural networks (ANNs) [*Haykin*, 1998], and now-a-day, different architectures and learning algorithms based on several ANN types, such as feed-forward artificial neural networks, radial basis function networks, recurrent neural networks, neuro-fuzzy networks, and hybrid neural networks [for a complete description, see *Zomaya*, 2006], have been proposed for forecasting the *Dst* index [e.g., see *Kugblenu et al.*, 1999; *Lundstedt*, 2005; *Sharifi et al.*, 2006a]. A feed-forward neural network is the most simple ANN-type devised [*Haykin*, 1998]. In this ANN, information flows in one direction through layers of neurons and it is usually trained by using the backpropagation algorithm [*Rumelhart and McClelland*, 1986]. *Dst* forecasting by means of feed-forward ANN was introduced by *Lundstedt* [1993]. Later, *Lundstedt and Wintoft* [1994] proposed an ANN in order to predict the *Dst* index 1 h in advance using the B_z component, density, and solar wind velocity. Next, *Gleisner et al.* [1996] used an ANN to compute the *Dst* index 1 h ahead using as input data the solar wind velocity and density, and IMF southward component. *Kugblenu et al.* [1999] developed an ANN for reproducing the storm recovery phase via *Dst* index by

using the last 3 h of the *Dst* index values during the main phase, the solar wind data of IMF southward component, total strength, and the square root of the dynamic pressure for the minimum *Dst* as input parameters. *Stepanova and Pérez* [2000] used an ANN for predicting the *Dst* index 3 h in advance making use of previous *Dst* values. *Jankovičová et al.* [2002] constructed an ANN to forecast *Dst* index using a multivariate method for determination of the input parameters so-called principal components. *Stepanova et al.* [2005] proposed an ANN for forecasting *Dst* index 1 h ahead using as input parameters the hourly average values of the polar cap index. *Bala and Reiff* [2012] improved *Dst* index prediction by using an ANN to generate 1–6 h in advance predictions by using the solar wind and magnetospheric data from the Advanced Composition Explorer (ACE) spacecraft as inputs. Most recently, *Revallo et al.* [2014] proposed a model to forecast the *Dst* index 1 h ahead based on an ANN combined with an analytical method for the solar wind-magnetosphere interaction utilizing hourly solar wind data such as the IMF B_z component, and the velocity, particle density, and temperature of solar wind.

On the other hand, a recurrent neural network is an ANN type with bidirectional data flow where the connections between neurons form a cycle [*Mandic and Chambers*, 2001]. *Wu and Lundstedt* [1996] introduced the application of recurrent neural networks Elman type [*Elman*, 1990] for forecasting the *Dst* index 1 h ahead from solar wind data such as IMF B_z component, solar wind speed, and density. Later, *Wu and Lundstedt* [1997] used an Elman-ANN to describe the solar wind-magnetosphere coupling by predicting *Dst* index 1–5 h in advance utilizing the interplanetary duskward electric field as the best input parameter. Afterward, *Munsami* [2000] applied an Elman-ANN to correlate the storm-substorm relationship via the ring current index and auroral electrojet indices. *Barkhatov et al.* [2000] used an Elman-ANN for forecasting the *Dst* index using the magnetic field and solar wind plasma parameters acquired from the OMNI database. *Lundstedt et al.* [2002] again implemented an Elman-ANN for predicting the *Dst* index by using averages of the IMF B_z component, and solar wind velocity and particle density. Similarly, *Watanabe et al.* [2002] applied an Elman-ANN for forecasting the *Dst* index by adding the IMF *x-y-z* components. *Pallocchia et al.* [2006] also used an Elman-ANN for the *Dst* forecast called EDDA (Empirical *Dst* Data Algorithm) and based on IMF inputs only. In other studies, *Vörös and Jankovičová* [2002] implemented a recurrent ANN type based on feedback connection to predict the *Dst* index using solar wind inputs taken from the principal component analysis of the IMF, proton density, and bulk velocity. *Xue and Gong* [2006] applied a fully connected neural network [*Williams and Zipser*, 1989] to generate a new model for forecasting the *Dst* index by employing the measured parameters from ACE spacecraft such as the IMF southward component, and the solar wind plasma density and velocity. Moreover, recurrent ANNs can appear as nonlinear versions of the response filters [*Puskorius and Feldkamp*, 1994] and as a nonlinear autoregressive model [*Chen and Billings*, 1989]. In this line of work, *Boaghe et al.* [2001] implemented an approach based on the Nonlinear AutoRegressive Moving Average model with eXogenous inputs [*Chen and Billings*, 1989] to analyze the *Dst* index and the solar wind parameter VB_s (the product of the solar wind speed V and the IMF southward component B_s), and the merging rates of the interplanetary magnetic field and the geomagnetic field. *Drezet et al.* [2002] used a kernel method based on the Volterra series and a Nonlinear Auto-Regressive eXogenous model [*Siegelmann et al.*, 1997] using the past values of VB_s as inputs. *Andrejková and Levický* [2003] applied an ANN using Bayesian training to the problem of predictions of the *Dst* index by depending on four quantities such as the IMF B_z component, the mean square error of B_z that characterize the swings of this quantity, and the solar wind velocity and plasma density. *Ouarbya et al.* [2012] proposed a Bayesian recurrent neural filter for online forecasting of the *Dst* index.

The neuro-fuzzy network is a fuzzy logic system into neural networks [*Buckley and Hayashi*, 1994]. *Andrejková et al.* [1997] applied a fuzzy ANN type to predict geomagnetic storms. Also, *Sharifi et al.* [2006a] introduced a method to estimate *Dst* index based on an analysis of singular spectrum for extracting the main components with a locally linear neuro-fuzzy network. Subsequently, *Sharifi et al.* [2006b] proposed another locally linear neuro-fuzzy model with a progressive tree construction learning algorithm to describe the *Dst* index using the past values of *Dst* index and southward interplanetary magnetic field, the *Dst* derivative value, and the root of dynamic pressure as input data. Next, *Mirmomeni et al.* [2006] introduced a learning method based on fuzzy descriptor systems combined with spectral analysis for predicting solar activity, while *Mirmomeni et al.* [2010] also proposed an adaptive neuro-fuzzy modeling with an online learning method for forecasting the *Dst* index. *Mirmomeni et al.* [2011] suggested an online multistep-ahead prediction of the *Dst* index using another adaptive neuro-fuzzy model combined with a recursive spectral analysis. In a related work, *Lotfi and Akbarzadeh-T* [2014] compared the results of Mirmomeni's method [*Mirmomeni et al.*, 2010] with an adaptive

brain emotional algorithm for online prediction of *Dst* index. Regarding self-organized maps [*Kohonen*, 1982], *Wintoft and Lundstedt* [1993] used self-organized maps for identifying structures in geoeffective solar winds. More related methods based on the ANN forecast of the *Dst* index can be revised in *Kugblenu et al.* [1999], *Lundstedt* [2005], and *Sharifi et al.* [2006a].

As mentioned above, solar wind data have been preferably used as ANN inputs for forecasting the *Dst* index, using mainly solar wind velocity and density, and IMF southward component. But the dynamic pressure of the solar wind, the electric field, and other components of the magnetic field have also served as input data in some methods [*Lundstedt*, 2005]. Note that only in few cases, past values of *Dst* index have been used as inputs but as a complement to other solar wind parameters [*Kugblenu et al.*, 1999; *Stepanova and Pérez*, 2000]. We posit that the disturbances caused by variations of the ring current over the solar wind-magnetosphere system can be described by the *Dst* index itself (through their past values) and obtain thus a better autocorrelation among its own values used as input-output in an ANN type. In this study, *Dst* index 6 h in advance is estimated by a feed-forward neural network using the past values of the *Dst* index as input parameters. We propose an improved method to predict the *Dst* variation based on measurements at ground level. Here the designed ANN was optimized using a particle swarm algorithm (PSO) [*Kennedy and Eberhart*, 1995] to update ANN weights. Note that PSO has been recently applied in the optimization of ANN weights with promising results and performance [*Lazzús et al.*, 2014; *López-Caraballo et al.*, 2015]. For ANN, commonly applied optimization algorithms such as backpropagation learning algorithm [*Rumelhart and McClelland*, 1986] can determine the network connection weights but do not have the capacity of controlling the parameter optimization in the absence of gradient information, while PSO can solve this problem [*Lazzús et al.*, 2014]. The task of PSO in ANN is to obtain the optimum weights (i.e., best particle position) where several particles (i.e., possible problem solutions) move in a search space to obtain the best solution to the problem [*López-Caraballo et al.*, 2015]. Note that the application of PSO to solve optimization problems in geosciences is scarce [*Yuan et al.*, 2009]. To our knowledge, no application to *Dst* index forecast exists using a hybrid ANN+PSO technique similar to the one presented here.

2. Computational Method

2.1. Artificial Neural Network

Neural networks are simple, adaptive sets of computational elements, interfaced in parallel and hierarchically organized which allows computational elements to interact with a system akin to a biological nervous system. The unit of an artificial neural network is a basic processor called "neuron," which has the limited capability of calculating a weighted sum of its inputs and then applying an activation function to obtain a signal that will be transmitted to the next neuron [*Rumelhart and McClelland*, 1986]. These artificial neurons are arranged in layers or levels and have a high degree of connectivity between them, which is weighted by their weights. By means of a supervised or unsupervised learning algorithm, ANNs adjust their architecture and parameters to minimize error functions suggesting the level of adjustment be applied to data and the generalization capacity of ANN [*Haykin*, 1998]. Note that in a supervised learning algorithm the network patterns and the desired output for those patterns are displayed and a mathematical error minimization formula is used to adjust the weights to give the desired output, while in an unsupervised learning algorithm, the network does not need to know the target patterns for the desired output, since this algorithm and its connection modification rule produce consistent output patterns [*Rumelhart and McClelland*, 1986; *Haykin*, 1998].

The performance of an ANN has two phases: training and prediction. The training phase (including learning and validation) is the most important process during which the ANN adjusts its weights in response to the input data for obtaining the corresponding output values. Here the neural network should learn from a training pattern set composed of input-output data, and the prediction phase is used for testing the computed solution to confirm the predictive power of the network [*Haykin*, 1998]. Then, the objective of learning is to adjust ANN weights (w) so that the output generated by the ANN is as similar as possible to the true output for a given input [*Lazzús et al.*, 2014].

As mentioned in section 1, the most commonly ANN type used can be characterized as a feed-forward network using a backpropagation algorithm based on a descending gradient error [*Rumelhart and McClelland*, 1986]. This ANN type consists of one input layer containing N-input units (for each network input parameter),

one hidden layer containing q-processing units (i.e., neurons), and one output layer of n-output units for each network output parameters [*Kang et al.*, 1993]. The output of an ANN can be calculated as follows:

$$y_n = f_n\left(\sum_{j=1}^{q} w_{nj} f_j\left(\sum_{i=1}^{N} w_{ji} x_i + b_j\right) + b_0\right), \tag{1}$$

where x_i is the normalized input vector $x_i = \left(X_i - X_i^{\min}\right)\frac{2}{X_i^{\max} - X_i^{\min}} - 1$, and $X_{\min}$ and $X_{\max}$ represent the smallest and largest data values. Note that prior to a training process, the input data normalization step is crucial in order to obtain optimum results during the ANN performance, as it may speed up the training and reduce the chances of getting stuck in a local minimum of the surface error [*Sola and Sevilla*, 1997]. In our case, the normalization employed is commonly used in neural networks applications [*Haykin*, 1998]. w_{nj} is the weights between unit j and unit n of input, and hidden layers and w_{ji} are the weights between hidden layer and an output; b_j and b_0 constitute the bias of the hidden layer and the output layer, respectively [*Jankovičová et al.*, 2002]. Additionally, the used activation functions f can be nonlinear or linear type. Commonly a tangent hyperbolic nonlinear activation function f_j is applied on the hidden layer, while a linear function f_n is applied on the output layer [*Haykin*, 1998].

In the training phase, the ANN biases and weights are adjusted in an iterative form in order to minimize the root-mean-square error (RMSE) between target outputs and network outputs. For a given set of N-input data, RMSE is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(Y_{\mathrm{calc}} - Y_{\mathrm{real}}\right)_i^2}{N}}, \tag{2}$$

where Y_{calc} is the neural network output value, Y_{real} the given output value, N the total data point, and i constitutes a particular data point. Furthermore, the accuracy of the optimum ANN can be checked by using the mean absolute error MAE and the correlation coefficient R between the calculated values (calc) and the real data (real). These statistical parameters are calculated as follows:

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|Y_{\mathrm{calc}} - Y_{\mathrm{real}}\right|_i \tag{3}$$

$$\mathrm{R} = \frac{\sum_{i=1}^{N}\left(Y_{\mathrm{calc}} - \overline{Y}_{\mathrm{calc}}\right)\left(Y_{\mathrm{real}} - \overline{Y}_{\mathrm{real}}\right)_i}{\sqrt{\sum_{i=1}^{N}\left(Y_{\mathrm{calc}} - \overline{Y}_{\mathrm{calc}}\right)_i^2}\sqrt{\sum_{i=1}^{N}\left(Y_{\mathrm{real}} - \overline{Y}_{\mathrm{real}}\right)_i^2}} \tag{4}$$

During this entire operation, the implementation of backpropagation learning algorithm updates the network weights and biases to the direction in which the merit function decreases most rapidly, i.e., the negative gradient [*Haykin*, 1998]. Note that our proposed ANN was trained for minimizing the RMSE (objective function) but substituting the backpropagation learning algorithm [*Rumelhart and McClelland*, 1986] with a particle swarm algorithm (PSO) [*Kennedy and Eberhart*, 1995] to optimize the weight updates in the ANN [*Lazzús et al.*, 2014].

2.2. Particle Swarm Optimization (PSO)

PSO is based on social behavior observed in groups of individuals such as swarms of insects [*Kennedy and Eberhart*, 1995]. A swarm intelligence can be defined as a structured collection of interacting organisms. Intelligence resides not only in individuals but also in the whole group moving together. Such social behavior is based on transmitting the success of each individual to the rest of the group. Each organism (i.e., particle) is processed as a point in a N-dimensional search space, which adjusts its "flight" according to its experience and the experience of the rest in the "band". Then the particles move through the search space by locating promising regions or particles [*Eberhart and Kennedy*, 1995].

Thus, PSO is a population-based algorithm initialized with a random particle population using a search strategy based on updating generations [*Lazzús et al.*, 2014]. For each iteration, the algorithm calculates each particle velocity j as follows [*Eberhart and Shi*, 2000]:

$$v_j^{k+1} = \omega v_j^k + c_1 r_1\left(\psi_j^k - s_j^k\right) + c_2 r_2\left(\psi_g^k - s_j^k\right), \tag{5}$$

where s represents the particle position and v the particle velocity in the search space, ω is the swarm inertia weight, c_1 and c_2 are two acceleration constants, r_1 and r_2 two elements each obtained from a random value in the range [0,1], and k is the current iteration; s_j^k is the current particle position, ψ_j^k and ψ_g denote the best solution that the particle has reached, and the best solution that all particles have reached, respectively [*Lazzús et al.*, 2014]. Note that the particle velocity value can be fitted into the range $[-v_{max}, v_{max}]$ in order to restrict the particles roaming outside the search space [*Clerc and Kennedy*, 2002]. Once the velocity has been calculated, the new position of each particle is computed as follows [*Eberhart and Kennedy*, 1995]:

$$s_j^{k+1} = s_j^k + v_j^{k+1}, \tag{6}$$

where s_j^{k+1} is the new position for a particle moving from a position in the search space to another one only by adding the new particle velocity v_j^{k+1} to the current particle position s_j^k.

Equations (5) and (6) show the update for the position and velocity vectors of each particle j in each iteration k, respectively. The cognitive component is modeled by factor $c_1 r_1 \left(\psi_j^k - s_j^k\right)$ and represents the distance between the current position and the best position known by such particle—the decision of the particle will be based on its experience throughout its life. The social component is modeled by factor $c_2 r_2 \left(\psi_g^k - s_j^k\right)$ and represents the distance between the current position and the best neighborhood position—the decision of the particle depends on the influence the rest of the cluster has on it [*Lazzús et al.*, 2014]. The inertial factor ω is multiplied by the current velocity v_j^k in the velocity updating equation. This inertial factor can be performed for each iteration as follows [*Eberhart and Shi*, 2000]:

$$\omega = \omega_{max} - \left(\frac{\omega_{max} - \omega_{min}}{k_{max}}\right) k, \tag{7}$$

where ω_{max} and ω_{min} denote the maximum and minimum inertia weight values, respectively; k_{max} is the maximum number of iterations, and k is the current iteration epoch. Note that high values for ω cause a thorough (more diversified) search, while low values for ω cause a more localized (more intensified) search [*Poli et al.*, 2007].

2.3. Hybrid ANN+PSO

As previously mentioned, our ANN uses a PSO algorithm for adjusting the weights and biases values (using the particle velocity and position updates) so that the given fitness function is minimized. Then a position represents the weight and bias values, and here the velocity contains the constraints limit for the weight and bias values. This hybrid method is called ANN+PSO. Recent studies have shown that PSO has been successfully applied to ANN training [e.g., *Grimaldi et al.*, 2004; *Lazzús et al.*, 2014; *López-Caraballo et al.*, 2015]. So once the ANN output values are calculated, the ANN weights and biases are adjusted $w_{i,j} \rightarrow w'_{i,j}$ by the PSO algorithm via equations (5) and (6) [*Lazzús et al.*, 2014; *López-Caraballo et al.*, 2015]. Here PSO modifies each ANN weight using the training data set. Thus, the network that computes the highest fitness will be considered the global best [*Lazzús et al.*, 2014].

Note that backpropagation based on error gradients is essentially a hill-climbing algorithm, and its major disadvantage is the susceptibility to premature convergence on local minima. Another disadvantage is its dependence on the starting point of the search, which would be the initial weights in case of ANNs [*Grimaldi et al.*, 2004]. The comparable advantages of PSO, however, are as follows: (i) less dependence on initial weight values, since multiple starting particles are used in the search process; (ii) not use of derivative information of the activation functions and the fitness function; and (iii) computationally more efficient, and more robust on rugged surfaces, because population-based search is less prone to premature convergence on local minima than backpropagation [*Zhang et al.*, 2007].

The step-by-step approach of ANN+PSO can be summarized thus [*Lazzús et al.*, 2014]:

Step 1. Randomly initialize positions (weights and biases) and velocities of a group of particles. Particles represent weight vectors, including biases, in an ANN. Here the overall number of weights and biases number denote the dimension of the search space.

Step 2. ANN is trained using the initial position of particles in PSO. A learning error in ANN can be treated as particles fitness value according to initial weights and biases. The best current fitness value recorded by particle j is archived as ψ_j^k, and the ψ_j^k with the best value is archived as ψ_g, then this value is stored.

Step 3. Evaluates the required value for the merit function (equation (2)) via the given data set.

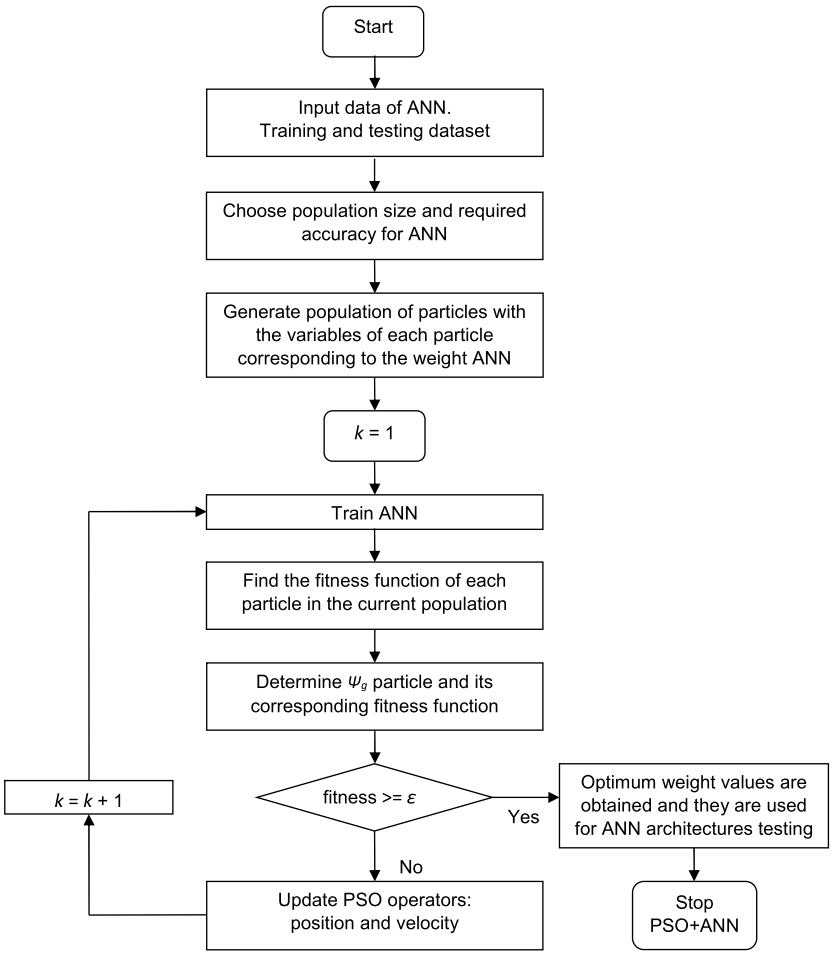

Figure 1. Block diagram for the ANN training via PSO algorithm.

Step 4. Compares the estimated merit value for each particle (F_j). If $F_j < \psi_j^k$ then $\psi_j^k = s_j^k$ is the corresponding coordinate for the best particle so far [*Lazzús et al.*, 2014].

Step 5. Calculates the merit function value for the new particle positions. If a better position is reached, ψ_j^k value will be substituted with the new value, and ψ_g value will be selected as in Step 2. If the new ψ_g value is better than the previous one, it is stored. When $F_j < \psi_g$ then $\psi_g = s_j^k$ will be the particle with the best merit values over the other particles in the swarm [*Lazzús et al.*, 2014].

Step 6. Reduces the learning error at the current epoch by changing particle velocity and position according to movement equations for the search space (equations (5) and (6) to update the ANN weights and biases.

Step 7. Repeat the process until the algorithm reaches a stopping criterion, such as minimum learning error or maximum iteration number. Otherwise, loop to Step 3.

Step 8. Finally, the optimum training process is successfully finished for the network, then the optimum ANN weights and biases are obtained via PSO algorithm.

A flow diagram for the ANN+PSO algorithm developed in this study is presented in Figure 1. The PSO specifications such as the number of particles j_{max}, constants c_1 and c_2, inertial weight ω, and maximum velocity v_{max} of the particles were tuned using a comprehensive trial-and-error process, as in *Lazzús et al.* [2014, 2016]. Thus, this procedure was used to adjust these values by analyzing the behavior of the objective function. First, the effect of the number of particles from 5 to 500 individuals in the swarm was analyzed. Note that for other PSO applications some authors have proposed that a large swarm increases the number of function evaluations to converge to a global minimum [*Lazzús et al.*, 2016], whereas Shi and Eberhart [2000] have shown that the number of particles has no effect on the performance of the algorithm. Next, the effect of inertial weight for values from 0.1 to 0.9 in increments of 0.1 was studied. In this case, PSO showed good convergence for a

Table 1. Specifications for the ANN+PSO Algorithm

Section	Parameter	Value
ANN	NN type	Feed-forward
	Number of hidden layers	1
	Maximum number of hidden neurons	40
	Normalization range	[−1, 1]
	Weight range	[−10, 10]
	Bias range	[−5, 5]
	Transfer function	*tansig*
	Number of iterations	1500
	Minimum error	1e−3
PSO	Number of particles in swarm (j_{max})	50
	Maximum velocity (v_{max})	12
	Minimum inertia weight (ω_{min})	0.5
	Maximum inertia weight (ω_{max})	0.7
	Cognitive component (c_1)	1.494
	Social component (c_2)	1.494
	Number of iterations (k_{max})	1500
	Objective function	RMSE

linearly decreasing inertial weight beginning at 0.7 and ending at 0.5. Afterward, for selecting constants c_1 and c_2, we studied the effect of the variation of these constants between the values commonly used in PSO applications, such as 1.250, 1.494, 1.750, 1.800, and 2.000 [*Lazzús et al.*, 2014, 2016]. For this case, $c_1 = c_2 = 1.494$ presented a better convergence than the other values. Also, the maximum velocity value was analyzed from 1 to 20 in increments of 1, determining the value with the best convergence was 12. The tuned ANN+PSO algorithm specifications are presented in Table 1.

3. Forecasting With ANN+PSO

3.1. Database

A set of 233,760 hourly data of geomagnetic *Dst* index was taken from the World Data Centre for Geomagnetism, Kyoto [*World Data Center* (WDC), 2016], and it was used to train the proposed ANN. Figure S1 shows the time series used (see supporting information). This data set contains hourly *Dst* index data from 00 UT on 1 January 1990 to 23 UT on 31 August 2016.

Our study used a leave-25%-out cross-prediction method to estimate the prediction capabilities of the ANN model. Training set (by containing 116,880 hourly data points from 00 UT on 1 January 1990 to 23 UT on 2 May 2003 for the learning process, and 58,440 hourly data points from 00 UT on 3 May 2003 to 23 UT on 31 December 2009 for the validation process), and testing set (with 58,440 hourly data points from 00 UT on 1 January 2010 to 23 UT on 31 August 2016) were selected considering that they contained an adequate frequency of the *Dst* variations to ensure correct learning with the training database (≥50% of data points of the overall database). Based on the largest decay values for *Dst* index, geomagnetic storms can be categorized as low ($Dst > -20$nT), medium (-20nT $> Dst > -50$nT), high (-50nT $> Dst > -100$nT), and intense ($Dst < -100$nT) storms [*Jankovičová et al.*, 2002]. Note that $Dst > -20$nT variations could be caused by other transient phenomena not connected directly with geomagnetic storms; however, these data are more present in this index and the network should learn to recognize them and decide accordingly. Table 2 shows the levels of the geomagnetic storms present in our database and the number of data points contained in each level. From this analysis, Figure S2 shows the occurrence of geomagnetic storm events by level classification for each year studied (see supporting information). As shown, the occurrence of very intense geomagnetic storms ($Dst < -100$nT) is lower than other storm levels; however, intense geomagnetic storms have a major impact on modern society, as outlined above (see section 1). More details on the trends seen in the occurrence of very intense geomagnetic storms emerge from analyzing the corresponding solar cycles. Note that solar cycle 22 (1986–1996) registered its maximum phase of sunspot activity between 1989 and 1991, while from 1992 to 1996 a period of minimum phase of sunspot activity was registered [*Dubey and Mishra*, 2000].

Table 2. *Dst* Levels Present in the Database Used

Geomagnetic Storm Levels	Learning Set (h)	Validation Set (h)	Testing Set (h)	Total Set (h)
$Dst > -20$nT	71961	45377	47007	164345
$-20\text{nT} > Dst > -50$nT	34951	11378	9774	56103
$-50\text{nT} > Dst > -100$nT	8383	1382	1519	11284
$Dst < -100$nT	1585	303	140	2028

For solar cycle 22, the occurrence frequency shows that very intense recorded geomagnetic storms occurred during its maximum phase (see Figure S2) and also that storms with sudden and gradual commencement were recorded during its maximum solar activity phase [*Dubey and Mishra*, 2000]. On the other hand, solar cycle 23 (1996–2007) registered its maximum solar activity phase in 2000. During this cycle, the occurrence frequency of storms of sudden and gradual commencement do not show a significant correlation with both the minimum and the maximum solar activity phases [*Pandey and Dubey*, 2009]. A large occurrence of very intense geomagnetic storms appears during the final phase of solar cycle 23 [*Pandey and Dubey*, 2009] (see Figure S2). Last, during the current solar cycle 24 (2008 – present), it registered its maximum solar activity phase at ~2014 (2014.5 ± 0.5) [*Pesnell*, 2014] and records an occurrence of intense geomagnetic storms during 2015, which would correspond to its maximum solar activity (see Figure S2). In order to clarify the occurrence of storm with $Dst < -100$nT, a histogram for these events was created. Figure S3 presents the number of *Dst* data points with the corresponding thresholds of very intense storms contained in the training and testing sets (see supporting information). These data show that this study contains a great of very intense storms with levels of $Dst < -100$nT are contained in this study. A complete analysis of *Dst* values and the occurrence rate for geomagnetic storms over the period 1963–2012 can be found in *Yermolaev et al.* [2013].

Normally, databases created from time series are subject to persistence analysis from statistical parameters with associated current and past (or future) series values that indicate which time series values are most useful for predicting future values, such as the autocorrelation function [*Palit and Popovic*, 2005]. Autocorrelation is a correlation of a time series with a delayed copy of itself as a function of time delay. In this way, a persistence analysis of the *Dst* index assumes that data conditions at the time of forecast remain stable. In other words, there is similarity between observations as a function of a time interval between them [*Podladchikova and Petrukovich*, 2012]. But in practice, geomagnetic storm forecasts are done mostly on the basis of persistence and recurrence that are not always strong [*Tsagouri et al.*, 2005]. Figure 2 shows the obtained linear correlation coefficient between current *Dst* values and future *Dst* values of the overall series to evaluate a cutoff value and to detect significant correlation between time series for a correct *Dst* forecast. This figure shows a cutoff value of $t + 3$ to obtain a correct prediction for the *Dst* index; i.e., *Dst* index values above 3 h in advance are difficult to obtain using the internal correlation of its time series. Note that a similar result appears in *Stepanova and Pérez* [2000]. To get a better forecast, we assume that the best inputs for our ANN should be optimally selected from the time series of the *Dst* index and thus obtain an improved forecast for more hours in advance [*Joselyn*, 1995; *Thomson*, 2000].

3.2. Inputs and Training

We used the time series data for forecasting *Dst* as illustrated in Figure S1 (see supporting information). The main challenge of time series study is to predict their next values for a known series up to a specific time, by using their known values from the time series. For the time series prediction, this is generally represented using delay coordinates as follows:

$$x(t+d) = F\left[x(t), x(t-d_1), x(t-d_2), \ldots, x(t-d_{M-1})\right], \tag{8}$$

where $x(t)$ denotes the value for the time series at time t, d represents the delay factor between the individual observations, and M is the number of observations considered. Vector F is used for predicting the next value $x(t + d)$ for the time series [*Palit and Popovic*, 2005]. Many of these models have been proposed to predict future values associated with various geophysical systems (see examples in *van der Baan and Jutten* [2000] and *Poulton* [2002]). In our case, the goal of this method is to use past values for the time series of geomagnetic indices in order to predict future values for the geomagnetic *Dst* index: $Dst(t + 1)$ to $Dst(t + 6)$. Note that the actual advance time for accurate predictions (before the minimum *Dst* value is reached) has been estimated

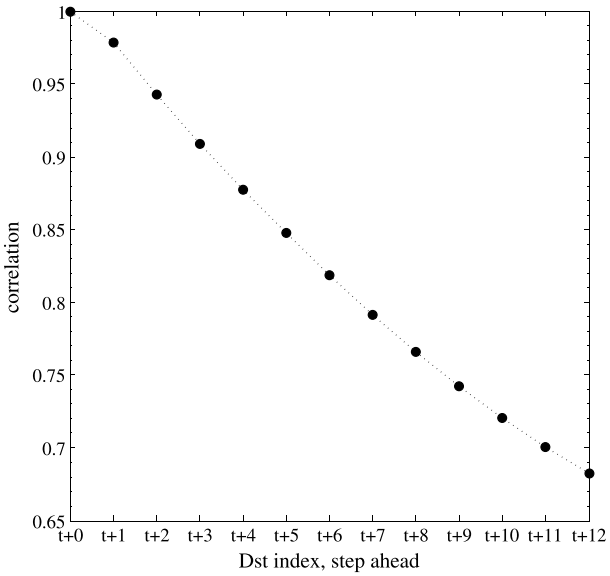

Figure 2. Correlogram for *Dst* index.

5–6 h on average [*Podladchikova and Petrukovich*, 2012], but the current best predictions available in the literature are only obtained for 2–3 h ahead of the current time [*Bala and Reiff*, 2012], as evidenced in Figure 2. Thus, hourly data from the current value to 48 past hours $(t, t-1, \ldots, t-47, t-48)$ were considered in order to select the best input parameters. These inputs were added systematically, by evaluating the statistical parameters RMSE, MAE, and R for each case studied ($Dst(t+1)$ to $Dst(t+6)$). Based on the minimization of equation (2), an exhaustive procedure was applied in order to obtain the best input vector, starting from the current value t to the past value $t-48$, in steps of 1 h per time. The following input vectors were selected from this methodology:

$$Dst(t+1) = F[Dst(t, t-1, t-2)] \tag{9}$$

$$Dst(t+2) = F[Dst(t, t-1, t-2, t-3, t-4)] \tag{10}$$

$$Dst(t+3) = F[Dst(t, t-1, t-2, t-3)] \tag{11}$$

$$Dst(t+4) = F[Dst(t, t-1, t-2, t-3)] \tag{12}$$

$$Dst(t+5) = F[Dst(t, t-1, t-2, t-3, t-4, t-5, t-6)] \tag{13}$$

$$Dst(t+6) = F[Dst(t, t-1, t-2, t-3, t-4, t-5, t-6)], \tag{14}$$

where for each case the input vector was selected independently. Note that similar input vectors for $t+3$ and $t+4$, and $t+5$ and $t+6$ are only an unintentional coincidence since they correspond to different cases studied and at this point have no relation with each other. Once input vectors were selected, the next step was to normalize their values (see description of equation (1). Thus, *Dst* index min and max values for the corresponding studied period were 67 nT and −422 nT, respectively.

Additionally, a different number of neurons in the hidden layer were tested in order to select the most accurate topology. Note that the number of neurons in each layer are forming the so called network topology or architecture for each input vector [*Haykin*, 1998]. Usually, for a supervised learning a fully connected three layer feed-forward ANN is used [*Kang et al.*, 1993]. To guarantee adequate representation of the information contained in the input data for an optimum learning process, we must provide a sufficient number of neurons in the hidden layer to ensure correct training. No specific methods for determining the optimum number of neurons in the hidden layer (NHL) exist, giving rise to many potential alternate combinations [*Lazzús et al.*, 2014]. We opted to determine this number of neurons by adding neurons in a systematic manner from 1 to 40 units and by evaluating the RMSE, MAE, and R for each neural network architecture during the training phase process [*Lazzús et al.*, 2014; *López-Caraballo et al.*, 2015]. Figure S4 shows the correlation coefficient (R) found in the prediction of $Dst(t+d)$ as a function of the number of neurons in the hidden layer NHL (see supporting information). Note that in all the cases studied, the best inputs were obtained in the first past hours (but not more than $t-6$), while the optimum number of neurons in the hidden layer was found between 20 and 30 units.

3.3. Results

Figure 3 shows a comparison between real (solid line) and calculated values (dots) for the *Dst* index that were obtained by using the proposed ANN+PSO method. It also shows a comparison during the training phase

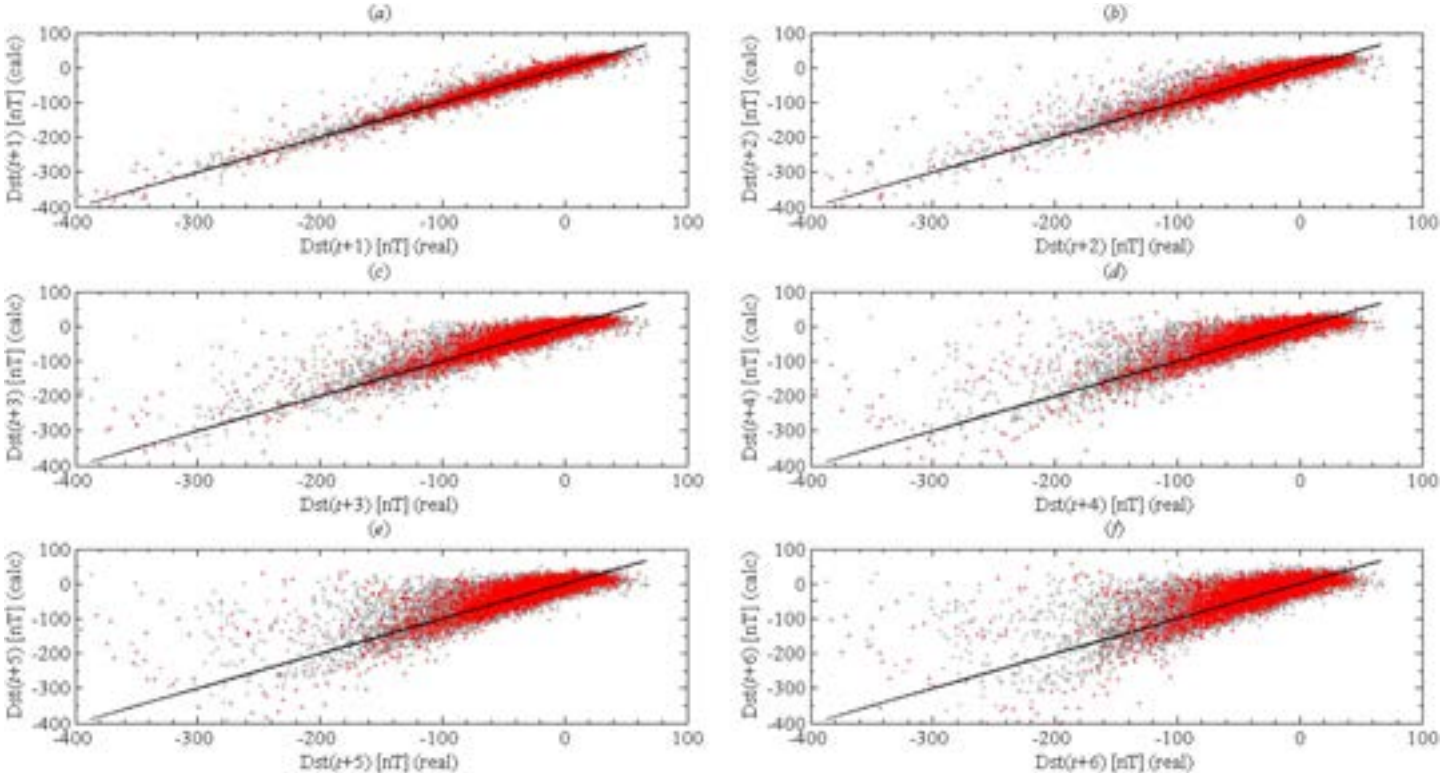

Figure 3. Comparison between calculated and real values for forecasting $Dst(t + d)$ during the training phase. Cases: (a) $Dst(t + 1)$; (b) $Dst(t + 2)$; (c) $Dst(t + 3)$; (d) $Dst(t + 4)$; (e) $Dst(t + 5)$; and (f) $Dst(t + 6)$. In these graphs grey dots show the learning process and red dots show the validation process.

between correlated values and real values for $Dst(t + d)$ (by containing 116,880 hourly data points from 00 UT on 1 January 1990 to 23 UT on 2 May 2003 for the learning process, and 58,440 hourly data points from 00 UT on 3 May 2003 to 23 UT on 31 December 2009 for the validation process). Figures 3a to 3f show the cases from $Dst((t + 1)$ to $(t + 6))$, respectively. In these cases the correlation coefficient R (equation (4) decreases from 0.98 to 0.89 (expected to be 1.0), while the slope m for the curve varies from 1.1 to 0.9 (also expected to be 1.0). Note that the high values of R show that the proposed method trained the ANN correctly.

Once ANN training was complete, the optimum ANN weights were obtained. ANN weights represent the network solutions for the studied problem [*Lazzús et al.*, 2014]. Tables S1 to S6 show the optimum weights determined by the hybrid algorithm ANN+PSO for forecasting the *Dst* index (see supporting information).

Figure 4 presents a similar comparison with Figure 3 between real and calculated values for the *Dst* index obtained via ANN+PSO method during the prediction phase by the testing set (with 58,440 hourly data points from 00 UT on 1 January 2010 to 23 UT on 31 August 2016). To clarify, all data used by the testing set were not used during the training step. Figures 4a to 4f show the ANN prediction capability for cases $Dst((t + 1)$ to $(t + 6))$, respectively. These cases show similar results to Figure 3 and show that the proposed technique can forecast the *Dst* index with good accuracy.

Table 3 summarizes the statistical results obtained in the forecast of $Dst((t + 1)$ to $(t + 6))$ using the selected input vectors (equations (9) to (14), respectively) via performance analysis of ANN+PSO (see Figure S4 in the supporting information). In this table, RMSE is the root-mean-square error (equation (2), R is the correlation coefficient (equation (4), and MAE is the average absolute relative deviation (equation (3). Additionally, the last column in this table contains the correlation coefficient obtained by the persistence method (R_{pers}). Clearly, RMSE and MAE increase with the time ahead, while R decreases when the time ahead progress from 1 to 6 h. Note that the learning set has a greater number of geomagnetic storms than the validation set, as well as a greater number of storms than the test set. Similarly, the learning set has higher Dst_{peak} than the validation set and the test set (see Figure S2 in the supporting information). The results show that the increasing or decreasing values of the statistical deviations agree with the characteristics of each set such as the number of intense geomagnetic storms and the number of minimum Dst_{peak} (see Tables 2 and 3).

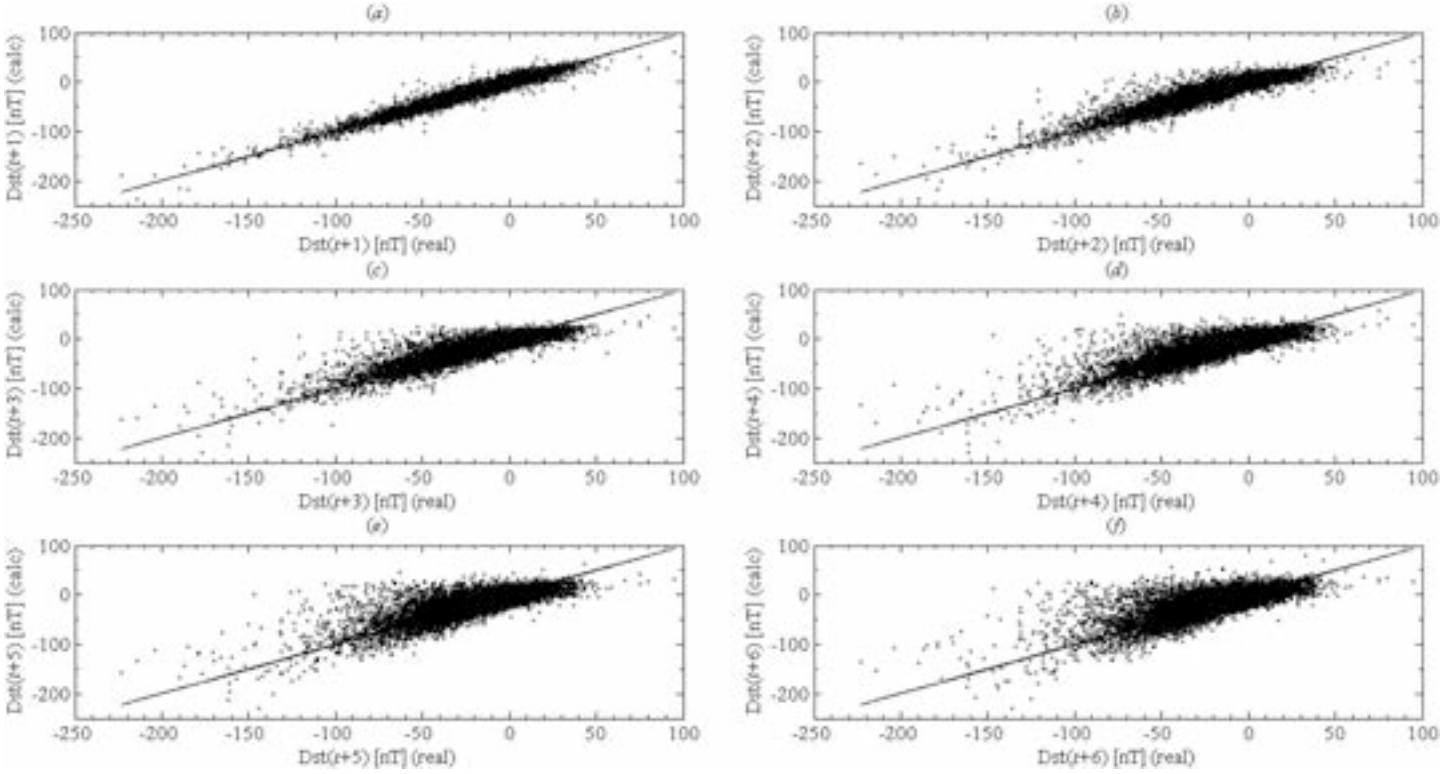

Figure 4. Comparison between calculated and real values for forecasting $Dst(t + d)$ during the prediction phase. Cases: (a) $Dst(t + 1)$; (b) $Dst(t + 2)$; (c) $Dst(t + 3)$; (d) $Dst(t + 4)$; (e) $Dst(t + 5)$; and (f) $Dst(t + 6)$.

The forecast of *Dst* index was compared by considering the prediction for each year used in the database. Figure 5 shows the forecast of *Dst* index per year from 1990 to 2016, obtained by ANN+PSO method. Figures 5a to 5f show these analyses for the cases from $Dst((t + 1)$ to $(t + 6))$, respectively. As shown, the effectiveness of the forecast decreases as time ahead increases. Additionally, 2 years show a lower effectiveness for all cases, such as 1996 and 2009. Note that these years present a lower occurrence of geomagnetic storms with levels of $Dst < -20$ nT, while they have a higher occurrence of geomagnetic storms with levels of $Dst > -20$ nT (see Figure S2 in the supporting information). This is due to solar cycle 22 finalizing at the end of 1995 and solar cycle 23 starting in early May 1996, causing a lower Sun-magnetosphere interaction. A similar effect appears in 2009 as solar cycle 24 started at the end of 2008 [*Hathaway*, 2015].

Moreover, an exhaustive analysis of geomagnetic storm types was conducted to distinguish the predictive capability of ANN+PSO among the different storm levels of the *Dst* index. Figure 6 shows the statistical results obtained using the proposed ANN+PSO method according to *Dst* levels taken from *Jankovičová et al.* [2002]. As in Figure 6, the best results in terms of *R* were for *Dst* $((t + 1)$ to $(t + 3))$ at all *Dst* levels (1 to 4). Yet the

Table 3. Summary of Statistical Results Obtained in the Forecast of $Dst((t + 1)$ to $(t + 6))$

			Learning			Validation			Prediction			Total			
			RMSE	MAE		RMSE	MAE		RMSE	MAE		RMSE	MAE		
Ahead	Input	NHL	(nT)	(nT)	*R*	(nT)	(nT)	*R*	(nT)	(nT)	*R*	(nT)	(nT)	*R*	R_{pers}
$t + 1$	Equation (9)	30	4.74	3.02	0.982	3.78	2.34	0.981	3.57	2.37	0.978	4.24	2.68	0.982	0.978
$t + 2$	Equation (10)	25	7.85	5.02	0.950	6.33	3.89	0.946	5.97	4.02	0.936	7.05	4.48	0.949	0.942
$t + 3$	Equation (11)	28	9.87	6.27	0.920	7.96	4.88	0.913	7.54	5.08	0.895	8.87	5.62	0.918	0.908
$t + 4$	Equation (12)	26	11.57	7.27	0.890	9.53	5.66	0.876	8.82	5.88	0.857	10.44	6.52	0.887	0.865
$t + 5$	Equation (13)	19	12.93	8.04	0.862	10.63	6.17	0.845	9.75	6.44	0.825	11.65	7.17	0.858	0.837
$t + 6$	Equation (14)	28	14.54	8.99	0.830	11.99	6.85	0.808	10.89	7.15	0.788	13.09	7.99	0.826	0.808

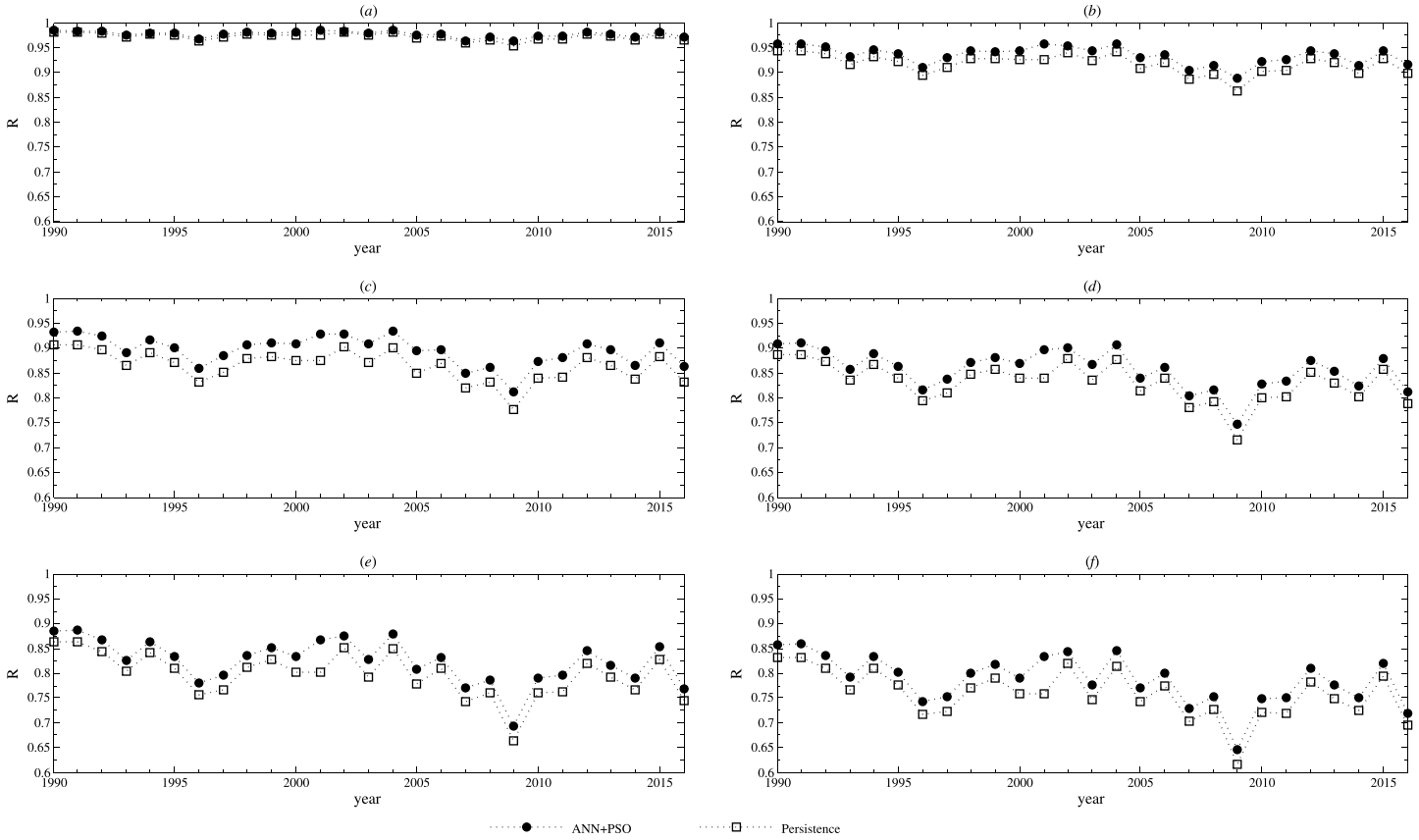

Figure 5. Forecast of *Dst* index per year from 1990 to 2016. Cases: (a) $Dst(t+1)$; (b) $Dst(t+2)$; (c) $Dst(t+3)$; (d) $Dst(t+4)$; (e) $Dst(t+5)$; and (f) $Dst(t+6)$.

effectiveness of the forecast decreases as the time ahead increases. In addition, geomagnetic storms of level 2 and 3 obtained a lower R for shorter times in advance ($t+1$ to $t+3$), while geomagnetic storms of levels 3 and 4 obtained a lower R for long times in advance ($t+4$ to $t+6$). These results show that the proposed hybrid algorithm is a powerful technique for forecasting the *Dst* levels for shorter times in advance $t+1$ to $t+3$. In this analysis, we calculated the *Dst* index variations over 352 geomagnetic storm events reported by *Kakioka Magnetic Observatory (KMO)* [2016] from 1990 to 2016. Specifically, these events correspond to 218 geomagnetic storms included in the learning set, 60 storms included in the validation set, and 74 storms in our testing set. These intense storms show geomagnetic events with peaks from −503 to −59 nT in H component and with many hours in duration. All geomagnetic storms were extracted from the results obtained via the ANN+PSO method, and their correlation coefficients were calculated as a further test for our proposed method. Tables S7 to S9 show the results obtained for forecasting the *Dst* index over 352 geomagnetic storms reported by *KMO* [2016] and using our proposed method from $t+1$ to $t+3$ hours ahead (see supporting information). These results show that our proposed method can forecast the *Dst* index relatively accurately (such as RMSE = 6.57 nT, MAE = 4.46 nT, and $R = 0.939$ for $t+1$; RMSE = 11.32 nT, MAE = 7.53 nT, and $R = 0.854$ for $t+2$; and RMSE = 13.93 nT, MAE = 9.28 nT, and $R = 0.802$ for $t+3$), as well as predict the complete behavior of a geomagnetic storm. To illustrate this fact, we randomly selected six geomagnetic storms for a graphical analysis. Figure 7 shows the results from forecasting the *Dst* index during the selected geomagnetic storms. Figures 7a and 7b represent geomagnetic storms that were part of the learning set, while Figures 7c and 7d were part of the validation set and Figures 7e and 7f show storms that were part of the prediction set. As can be observed, the characteristic evolution of a geomagnetic storm causes an abrupt variation in the *Dst* index, whose values descrease (negative) until reaching a minimum value and then starting recovery until reaching again a *Dst* value close to zero [*Sugiura and Kamei*, 1991]. Crucially, this whole characteristic behavior was correctly predicted by our proposed method, and for these six examples the ANN+PSO method shows excellent results for $t+1$ (with RMSE = 8.26 nT, MAE = 5.51 nT, and $R = 0.977$), $t+2$ (with RMSE = 15.39 nT, MAE = 9.72 nT, and $R = 0.937$), and $t+3$ (with RMSE = 19.48 nT, MAE = 11.79 nT, and $R = 0.900$). Additionally,

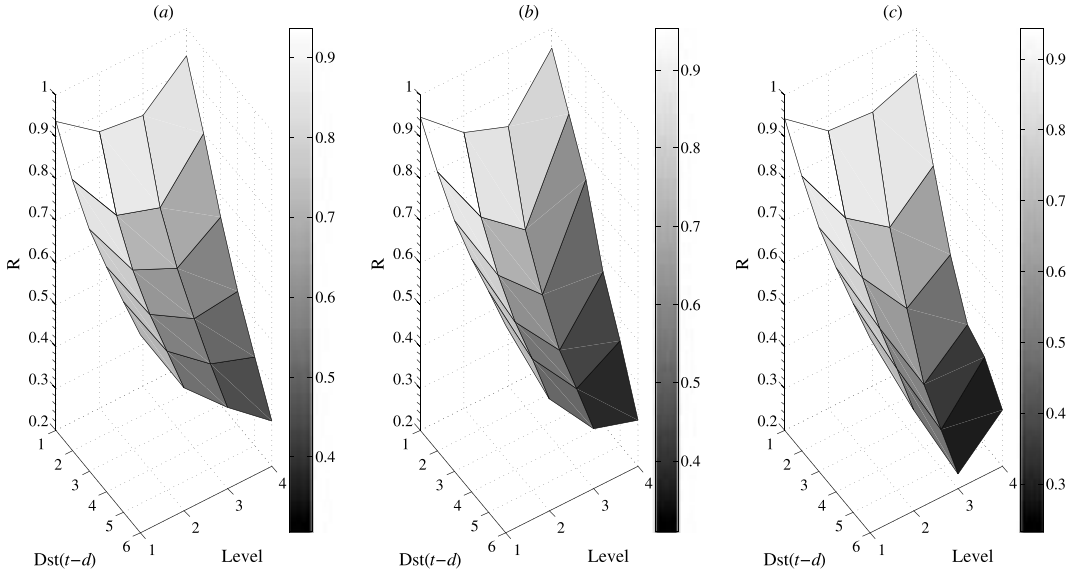

Figure 6. Statistical results obtained using the proposed ANN+PSO method according to *Dst* levels. (a) Learning set, (b) validation set, and (c) testing set. In this figure the *Dst* levels were labeled as follows: [1] low storm (*Dst* > −20 nT), [2] medium storm (−20 nT > *Dst* > −50 nT), [3] high storm (−50 nT > *Dst* > −100 nT), and [4] intense storm (*Dst* < −100nT).

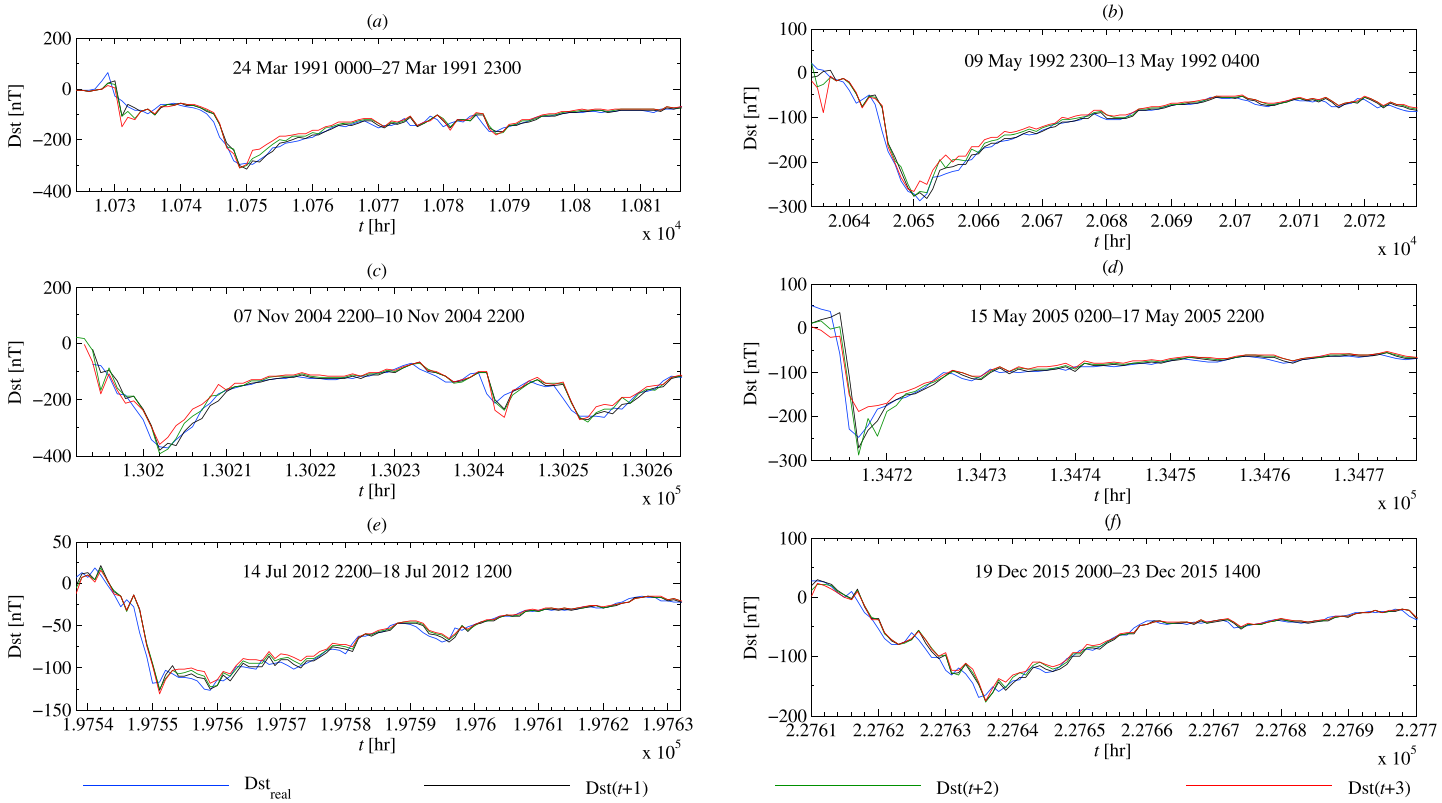

Figure 7. Examples of accuracy of the proposed method for forecasting the *Dst* index from 1 to 3 h ahead. (a) and (b) Geomagnetic storms that were part of the learning set, with a Dst_{peak} of −298 nT and −288 nT, respectively; (c) and (d) geomagnetic storms that were part of the validation set, with a Dst_{peak} of −374 nT and −108 nT, respectively; and (e) and (f) geomagnetic storms that were part of the testing set, with a Dst_{peak} of −127 nT and −170 nT, respectively.

Table 4. Summary of Results Obtained for Geomagnetic Events Taken From *KMO* [2016], and Used as an External Test for the Predictive Capacity of the Proposed Method for $Dst(t+1$ to $t+3)$ According to the Storm Phases

	Main Phase				Recovery Phase			
Ahead	RMSE (nT)	MAE (nT)	R	R_{pers}	RMSE (nT)	MAE (nT)	R	R_{pers}
$t+1$	10.87	8.62	0.925	0.908	4.98	3.64	0.944	0.929
$t+2$	19.14	14.93	0.860	0.826	8.09	5.94	0.879	0.847
$t+3$	25.18	20.25	0.816	0.789	9.87	7.22	0.840	0.802

Figure 7 presents the three standard phases of a geomagnetic storm, namely, initial phase, main phase, and recovery phase (see *Gonzalez et al.* [1994], for an analysis of the causes of the different phases of geomagnetic storms). We analyzed the forecast capabilities of the proposed method for main and recovery phases data for the 352 collected geomagnetic storms from *KMO* [2016]. Tables S10 to S12 show this complementary analysis (see supporting information). Clearly, our proposed method can also forecast all geomagnetic storm phases with higher accuracy; see Table 4. Moreover, the initial phase was also correctly predicted for the storms in the database that exhibited this behavior. In addition, this table contains the correlation coefficient obtained by the persistence method (R_{pers}) for an objective comparison.

4. Discussion

In general, the proposed method performed well accurately for forecasting the *Dst* index from 1 to 6 h in advance, but the best results were obtained for 1 to 3 h in advance.

Our selection of the input vector was systematic from $t-0$ to $t-48$ h backward, considering that the number of neurons in the hidden layer increases from 1 to 40. After applying this methodology, we found that the forecast of the *Dst* index shows a hidden persistence since (at intervals) predicting with a time in advance, it needs (approximately) the selfsame time backward but as an input vector (see Figure S4 in the supporting information).

As mentioned in section 1, various models have been developed to forecast the *Dst* index (mainly for 1 h ahead) [*Iyemori et al.*, 1979; *Burton et al.*, 1975; *Stepanova et al.*, 2005; *Sharifi et al.*, 2006a; *Wei et al.*, 2007; *Revallo et al.*, 2014]. These methods are up to 90% accurate but the time of alert (∼1 h) is too short to prevent damage caused by geomagnetic storms [*Khabarova*, 2007]. In our case, the proposed method can forecast the *Dst* index for more than 1 h ahead (see Table 3). Table 5 presents the best and recommended cases for using the proposed method from the obtained results. All cases marked with "✓" signal superior forecasting capacity, while those marked with "±" imply that the method should be used with caution. It is necessary to highlight the forecasting capacity of $Dst(t+1)$ to $Dst(t+3)$ over other methods found in the literature. In order to demonstrate this fact, Table 6 summarizes the best results and limitations reported in the literature for various artificial neural network-based methods for forecasting the *Dst* index. From Table 6 some artificial neural network methods can be highlighted. For example, *Wu and Lundstedt* [1997] obtained an RMSE from 14 nT to 24 nT with R from 0.9 to 0.7 in the forecast of *Dst* index from 1 to 8 h in advance using selected storms between years 1963 and 1992. *Stepanova and Pérez* [2000] obtained R between 0.95 and 0.72 from selected storms taken since 1983. *Jankovičová et al.* [2002] obtained R of 0.93, 0.73, 0.69, and 0.66 in the forecast of *Dst* index from 1, 6, 12, and 18 h in advance, respectively, using selected storms taken from 1998 to 1999. *Sharifi et al.* [2006b] obtained RMSE from 4.4 nT to 12 nT with R from 0.98 to 0.87 for the prediction of *Dst* index from 1 to 4 h ahead using data from 1995 to 1999. *Bala and Reiff* [2012] obtained R of 0.86, 0.84, and 0.77 in the forecast of the *Dst* index from 1, 3, and 6 h in advance, respectively, using data from 1998 to 2009. On the other hand, the classical models such as Burton's model [*Burton et al.*, 1975] and its revised version [*O'Brien and McPherron*, 2000] obtain an accurate prediction of geomagnetic storms for $Dst > -150$ nT, but these methods cannot appropriately predict cases with $Dst < -150$ nT [*O'Brien and McPherron*, 2000; *Watanabe et al.*, 2002; *Wei et al.*, 2007]. However, the best results presented in Table 6 were obtained by studies based on small databases containing selected storms, whereas our model uses a large database, incorporating all levels of geomagnetic storms that occurred between 1990 and 2016. All results presented in Table 6 were obtained from different methodologies and databases, so the results cannot be directly compared

Table 5. Recommended Cases for Using the ANN+PSO Method

	Storm's Level				Storm's Phases		
	Low	Medium	High	Intense	Initial	Main	Recovery
Ahead	$Dst > -20$ nT	-20 nT $> Dst > -50$ nT	-50 nT $> Dst > -100$ nT	$Dst < -100$nT	Phase	Phase	Phase
$t+1$	✓	✓	✓	✓	✓	✓	✓
$t+2$	✓	✓	✓	✓	✓	✓	✓
$t+3$	✓	✓	✓	✓	✓	✓	✓
$t+4$	✓	✓	✓	±	±	±	✓
$t+5$	✓	✓	±	±	±	±	±
$t+6$	✓	✓	±	±	±	±	±

between them. Nevertheless, based on the statistical analysis of these different ANN methods, our ANN+PSO method can be considered as a very accurate method. Still, a comparison can be made for some selected data sets; thus, we compared our method versus other six *Dst* forecast models for 1 h in advance (called B [*Burton et al.*, 1975], FL [*Fenrich and Luhmann*, 1998], OM [*O'Brien and McPherron*, 2000], TL [*Temerin and Li*, 2002], W [*Wang et al.*, 2003], NM [*Boynton et al.*, 2011], and persistence P [*Witt and Malamud*, 2013]) using 63 selected storms with *Dst* levels < -100 nT taken between 1998 and 2006, following the methodology and database proposed by *Ji et al.* [2012]. Then, we statistically compared these six methods using the correlation coefficient *R*, the root-mean-square error, RMSE, the difference in the minimum value of forecasted and measured *Dst* values $|\Delta Dst_{\min}|$, and the absolute timing difference between the observed and measured *Dst* values $|\Delta t_{Dst}|$ to test the accuracy and performance of these methods. Table S13 shows the details of this analysis for each individual storm (see supporting information). In addition, Figure 8 presents this quantitative comparison. From these results, we find that for all parameters our proposed method is better than the others.

In order to quantify the accuracy of forecasts generated by our proposed method, we introduced a root-mean-square error skill score ($SS = 1 - \frac{RMSE_{forecast}}{RMSE_{reference}}$) as a measure of the prediction efficiency relative to persistence [*Wilks*, 2011]. For this, we used the forecast of $Dst(t+1)$, where the persistence method was employed as the reference forecast model. In this analysis we applied all cases previously studied, that is, analysis per year (1990 to 2016), and analysis per geomagnetic storm levels. These results appear in Tables S14 and S15, respectively (see supporting information). From the generic form of a skill score, when $SS > 0$ the evaluated method is better than reference, while if $SS < 0$ it is less skillful than reference. As shown in Tables S14 and S15, for all cases we method is better than the persistence method in the forecast of $Dst(t+1)$. In general, the proposed method presents a $SS = 0.096$ or an accuracy of ~10% over the persistence method via RMSE.

On the other hand, once the ANN+PSO was trained and its optimal weights were obtained, we could use them to forecast the rest of the *Dst* time series. Historical data for the *Dst* time series are available from 1963 to present [*World Data Center (WDC)*, 2016]. We can use as an additional testing set the available data from years 1963 to 1989. The results of this new test appear in Table S16 (see supporting information). In this analysis, the proposed method also obtains good results, with an identical accuracy as obtained for the initial interval studied (1990–2016).

Note that one of the innovative aspects of our work is the application of a hybrid algorithm composed of an artificial neural network plus a particle swarm in order to optimize the weights update process in the ANN, thus improving performance and accuracy for forecasting the *Dst* index. All our results show a significant increase in the accurate forecast of this important geomagnetic index and also show that the application of the hybrid algorithm with the appropriate selection of the input parameters was crucial. This study thus proposes an improved method for forecasting the *Dst* index based on measurements at ground level, and our inputs were based only on past values of *Dst* index to exclusively emphasize the disturbances caused by variations of the symmetric ring current over the solar wind-magnetosphere system and to obtain a better autocorrelation among its own values used as input-output. However, we recognize that forecasting geomagnetic storms should not be addressed exclusively using a ring current analysis, since the other magnetosphere regions are very closely related and all of them are involved when disturbances break the balance in the Sun-magnetosphere system [*Cid et al.*, 2013]. Therefore, geomagnetic indices that are also capable of measuring disturbances in auroral regions (such as *AL*, *AU*, *AE*, and *AO*) and indices that can also measure disturbances

Table 6. Some Methods Reported in the Literature for Forecasting the *Dst* Index

Ahead	ANN technique	*R*	RMSE [nT]	Database	Reference
$t+1$	Backpropagation Neural Network	0.92	15	1963–1983	*Gleisner et al.* [1996]
$t+1$	Elman Neural Network	0.91	16	1963–1987	*Wu and Lundstedt* [1996]
$t+1$	Elman Recurrent Neural Network	0.91	14.5	1963–1992	*Wu and Lundstedt* [1997]
$t+2$	Elman Recurrent Neural Network	0.89	16.3	1963–1992	*Wu and Lundstedt* [1997]
$t+3$	Elman Recurrent Neural Network	0.86	18.2	1963–1992	*Wu and Lundstedt* [1997]
$t+4$	Elman Recurrent Neural Network	0.83	19.9	1963–1992	*Wu and Lundstedt* [1997]
$t+5$	Elman Recurrent Neural Network	0.82	20	1963–1992	*Wu and Lundstedt* [1997]
$t+6$	Elman Recurrent Neural Network	0.82	20.8	1963–1992	*Wu and Lundstedt* [1997]
$t+7$	Elman Recurrent Neural Network	0.80	21.8	1963–1992	*Wu and Lundstedt* [1997]
$t+8$	Elman Recurrent Neural Network	0.77	23.1	1963–1992	*Wu and Lundstedt* [1997]
$t+1$	Backpropagation Neural Network	0.95	11	1972–1982	*Kugblenu et al.* [1999]
$t+1$	Elman Neural Network	0.88		1968–1987	*Munsami* [2000]
$t+1$	Backpropagation Neural Network	0.95		1983	*Stepanova and Pérez* [2000]
$t+2$	Backpropagation Neural Network	0.93		1983	*Stepanova and Pérez* [2000]
$t+3$	Backpropagation Neural Network	0.88		1983	*Stepanova and Pérez* [2000]
$t+4$	Backpropagation Neural Network	0.85		1983	*Stepanova and Pérez* [2000]
$t+5$	Backpropagation Neural Network	0.82		1983	*Stepanova and Pérez* [2000]
$t+6$	Backpropagation Neural Network	0.78		1983	*Stepanova and Pérez* [2000]
$t+7$	Backpropagation Neural Network	0.75		1983	*Stepanova and Pérez* [2000]
$t+8$	Backpropagation Neural Network	0.72		1983	*Stepanova and Pérez* [2000]
$t+1$	Backpropagation Neural Network	0.93		1998–1999	*Jankovičová et al.* [2002]
$t+6$	Backpropagation Neural Network	0.73		1998–1999	*Jankovičová et al.* [2002]
$t+12$	Backpropagation Neural Network	0.69		1998–1999	*Jankovičová et al.* [2002]
$t+18$	Backpropagation Neural Network	0.66		1998–1999	*Jankovičová et al.* [2002]
$t+1$	Backpropagation Neural Network	0.70		1997	*Stepanova et al.* [2005]
$t+1$	Elman Neural Network	0.83	13.9	1995–2005	*Pallocchia et al.* [2006]
$t+1$	Locally Linear Neuro-Fuzzy Model	0.98	4.38	1995–1999	*Sharifi et al.* [2006b]
$t+2$	Locally Linear Neuro-Fuzzy Model	0.95	7.43	1995–1999	*Sharifi et al.* [2006b]
$t+3$	Locally Linear Neuro-Fuzzy Model	0.95	10	1995–1999	*Sharifi et al.* [2006b]
$t+4$	Locally Linear Neuro-Fuzzy Model	0.87	11.83	1995–1999	*Sharifi et al.* [2006b]
$t+1$	Radial Basis Function Network		18.45	1998	*Wei et al.* [2007]
$t+1$	Backpropagation Neural Network	0.86	8.83	1998–2009	*Bala and Reiff* [2012]
$t+3$	Backpropagation Neural Network	0.84	9.6	1998–2009	*Bala and Reiff* [2012]
$t+6$	Backpropagation Neural Network	0.77	11.09	1998–2009	*Bala and Reiff* [2012]
$t+1$	Backpropagation Neural Network	0.77		1998–2005	*Revallo et al.* [2014]
$t+1$	Relevance Vector Machine	0.96	10	1996–2007	*Andriyas and Andriyas* [2015]

in ring current regions (such as *SYM-H* and *ASY-H*) could be used in future studies related to the forecast of geomagnetic storms based on measurements at ground level by using the ANN+PSO algorithm.

As it is well known, the *Dst* index is used for characterizing the energy transfer of solar wind that is strongly influenced by the *z* component in the interplanetary magnetic field identified by the *H* component behavior in the terrestrial magnetic field and the ring current [*Cid et al.*, 2013]. The ring current is caused by energetic ion particles penetrating into the magnetosphere due to the solar wind interaction [*Williams*, 1995; *Smith et al.*, 2004]. During this interaction, the particle drifts generated in the geomagnetic storm give rise to a westward current which results in a negative *Dst* index [*Gonzalez et al.*, 1994]. Our method could reproduce the whole *Dst* index dynamic behavior caused by the Sun-magnetoshpere interaction during a geomagnetic storm, including the three characteristic phases. The initial phase has been considered by some authors as a herald for geomagnetic storms [*Malville*, 1960; *Gonzalez et al.*, 1994]. Note that our method can reproduce

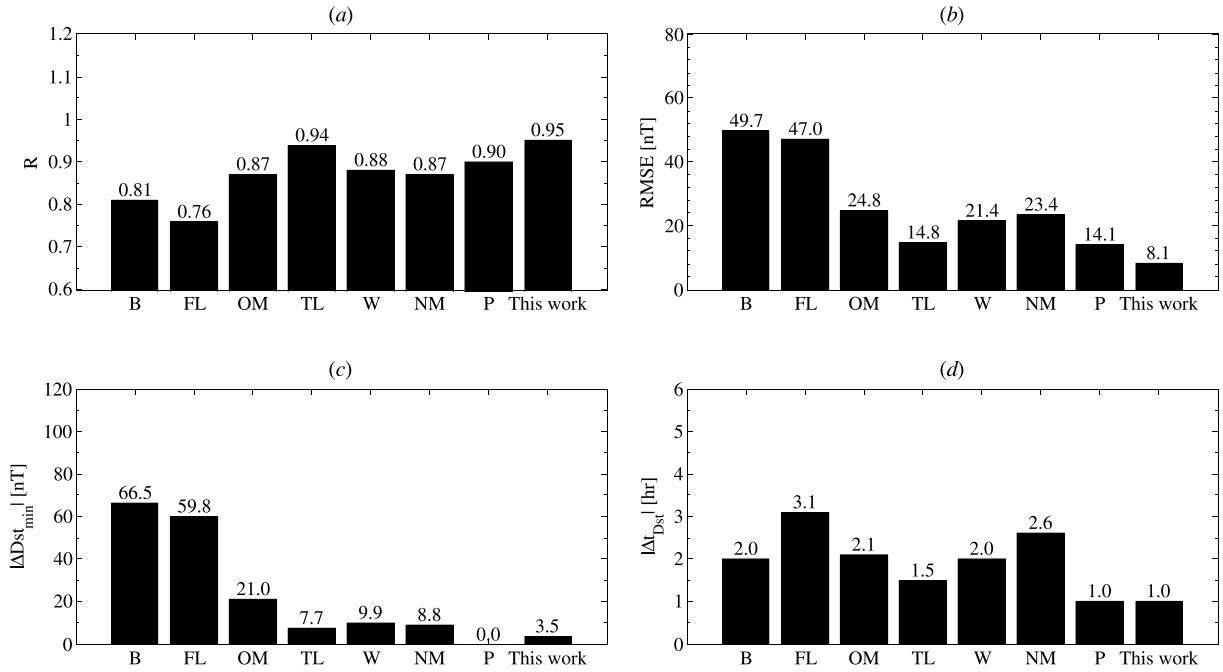

Figure 8. Comparison between the proposed method and other methods reported in the literature for forecasting the *Dst* index (called B [*Burton et al.*, 1975], FL [*Fenrich and Luhmann*, 1998], OM [*O'Brien and McPherron*, 2000], TL [*Temerin and Li*, 2002], W [*Wang et al.*, 2003], NM [*Boynton et al.*, 2011], and persistence P [*Witt and Malamud*, 2013]). (a) Correlation coefficient *R*; (b) root-mean-square error RMSE; (*c*) difference in the minimum value of forecasted and measured *Dst* $|\Delta Dst_{\min}|$; and (d) absolute timing difference between observed and measured *Dst* $|\Delta t_{Dst}|$.

a slight increase (caused in some storms) of *Dst* index positive values before the main phase, due to strong interaction of the solar wind plasma. The main phase begins when the energy passes from solar wind to the magnetosphere by intensifying the ring current and by producing a strong decrease of *Dst* index in increasingly negative value until reaching a minimum *Dst* value (Dst_{peak}) [*Akasofu and Chapman*, 1963; *Gonzalez et al.*, 1994]. For the main phase, our method can reproduce the energy transfer into the magnetosphere, which intensifies the ring current. Finally, the recovery phase, when the magnetosphere returns to the value of a quiet time due to the loss of the energy that entered the ring current during the main phase [*Cid et al.*, 2013]; and where our method can reproduce the Dst index variations from the Dst_{peak} to a new quiet time during this phase (see examples in Figure 7). Thus, the ANN+PSO method applied in the present study performs excellently, reproducing at least 96% of *Dst* index dynamic behavior for the initial, main, and recovery phases, e.g., $R_{initial} = 0.975$, $R_{main} = 0.925$, and $R_{recovery} = 0.944$ for $Dst(t+1)$. Note that this accuracy of 96% is obtained from the corresponding value of R^2=0.9647. The decay in *Dst* index at a minimum value is due to loss processes as a charge exchange [*Jordanova et al.*, 1996], particle-wave interaction [*Kozyra et al.*, 1997], and Coulomb interaction [*Kozyra et al.*, 1998]. We used this minimum value (Dst_{peak}) for classifying the levels of storm data intensity that occurred in the time interval from 1990 to 2016, as low (164,345 data points), medium (56,103 data points), high (11,284 data points), and intense (2028 data points) magnitudes (see Table 2). And all of the magnitude levels are correctly predicted for $t+1$ to $t+3$ (see Table 3).

We propose a method to forecast the *Dst* variation making use of measurements at ground level by considering that the parameters of interplanetary medium are only related to the injection function and assuming that the main loss processes which cause the decay of ring current, such as the conversion of energetic ions into neutrals, the direct particle precipitations into the auroral ionosphere, among others, have an intrinsically magnetospheric origin [*Stepanova and Pérez*, 2000]. In addition, long temporal correlation often occurs in the solar wind drivers, and so this model is also exploiting those correlations for forecasts. In general, forecasts made with the proposed method for $Dst(t+1)$ to $Dst(t+3)$ are quite accurate and better than those obtained by other authors as previously demonstrated (see Tables 5 and 6). However, the generated forecasts for $Dst(t+4)$ to $Dst(t+6)$ should be used with caution when estimating Dst_{peak} for storms of high to intense levels although it is worth highlighting that predictions generated by our method for $t+4$ to $t+6$ may fare

better than those generated by other studies for forecasting of long times in advance [e.g., *Wu and Lundstedt*, 1997; *Stepanova and Pérez*, 2000; *Jankovičová et al.*, 2002; *Sharifi et al.*, 2006b; *Bala and Reiff*, 2012]. But based on the analyzed, the *Dst* values obtained by our method could not be used as a basis for reproducing the variations in the solar wind and interplanetary magnetic field (or other geomagnetic interactions) [*Vichare et al.*, 2005; *Bhaskar and Vichare*, 2013; *Vichare et al.*, 2014], if these do not exhibit direct effects on the ring current.

5. Conclusions

Based on the results and discussion presented in this study, the following main conclusions are obtained:

1. The results show that our proposed hybrid algorithm can be properly trained for forecasting the *Dst* index with acceptable accuracy. Note that the strategy proposed for input selection and for the optimization of the architecture has been important for an efficient performance of our ANN+PSO algorithm.
2. The proposed method uses the measurements at ground level in order to forecast the *Dst* variation by considering that the parameters of interplanetary medium are only related to the injection function and that the main loss processes which cause the decay of ring current have an intrinsically magnetospheric origin.
3. The first past values of *Dst* index significantly influence the good training and predicting capabilities of our ANN; however, we recognize that forecasting geomagnetic storms should not be addressed exclusively using a ring current analysis from *Dst* values, since the other magnetosphere regions are closely related and all of them are involved when disturbances break the balance in the Sun-magnetosphere system. To improve the predictive capacity of the proposed method parameters such as solar wind velocity and pressure, components of interplanetary magnetic field, auroral electrojet (*AE*) indices, and longitudinally asymmetric (*ASY*) and symmetric (*SYM*) disturbance indices, among others could be used in future studies related to the forecast of intense geomagnetic storms using the ANN+PSO algorithm.
4. The low deviations obtained via our method show that the hybrid ANN+PSO algorithm can forecast the future values for the *Dst* index with more accuracy than other methods/models available in the literature. Forecasts generated by the proposed method for $Dst(t+1)$ to $Dst(t+3)$ are very accurate and better than those obtained by other authors, while the forecast generated for $Dst(t+4)$ to $Dst(t+6)$ should be used with caution when estimating the minimum *Dst* index for storms of high to intense levels.

Acknowledgments
The authors thank the Direction of Research and Development of the University of La Serena (DIDULS) and the Postgraduate Program of the Department of Physics of the University of La Serena for the special support that enabled the preparation of this paper. The geomagnetic data were obtained from the World Data Center (WDC) for Geomagnetism, Kyoto.

References

Akasofu, S. I. (1981), Energy coupling between the solar wind and the magnetosphere, *Space. Sci. Rev.*, *28*, 121–190, doi:10.1007/BF00218810.

Akasofu, S.-I., and S. Chapman (1963), The development of the main phase of magnetic storms, *J. Geophys. Res.*, *68*, 125–129, doi:10.1029/JZ068i001p00125.

Andrejková, G., H. Tóth, and K. Kudela (1997), Fuzzy neural networks in the prediction of geomagnetic storms, in *Proceedings of the Artificial Intelligence in Solar-Terrestrial Physics*, pp. 173–179, European Space Agency, Lund, Sweden.

Andrejková, G., and M. Levický (2003), Neural networks using Bayesian training, *Kybernetika*, *39*, 511–520.

Andriyas, T., and S. Andriyas (2015), Relevance vector machines as a tool for forecasting geomagnetic storms during years 1996–2007, *J. Atmos. Sol. Terr. Phys.*, *125–126*, 10–20, doi:10.1016/j.jastp.2015.02.005.

Axford, W. I., and C. O. Hines (1961), A unifying theory of high-latitude geophysical phenomena and geomagnetic storms, *Can. J. Phys.*, *39*, 1433–1464, doi:10.1139/p61–172.

Baker, D. N. (1986), Statistical analyses in the study of solar wind-magnetosphere coupling, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 17–38, Terra Sci., Tokyo, doi:10.1007/978-90-277-2303-1_2.

Bala, R., and P. Reiff (2012), Improvements in short-term forecasting of geomagnetic activity, *Space Weather*, *10*, S06001, doi:10.1029/2012SW000779.

Barkhatov, N. A., N. S. Bellustin, A. E. Leviti, and S. Y. Sakharov (2000), Comparison of efficiency of artificial neural networks for forecasting the geomagnetic activity index *Dst*, *Radiophys. Quantum Electron.*, *43*, 347–355, doi:10.1007/BF02677150.

Bhaskar, A., and G. Vichare (2013), Characteristics of penetration electric fields to the equatorial ionosphere during southward and northward IMF turnings, *J. Geophys. Res. Space Phys.*, *118*, 4696–4709, doi:10.1002/jgra.50436.

Blanch, E., S. Marsal, A. Segarra, J. M. Torta, D. Altadill, and J. J. Curto (2013), Space weather effects on Earth's environment associated to the 24–25 October 2011 geomagnetic storm, *Space Weather*, *11*, 153–168, doi:10.1002/swe.20035.

Boaghe, O. M., M. A. Balikhin, S. A. Billings, and H. Alleyne (2001), Identification of nonlinear processes in the magnetospheric dynamics and forecasting of *Dst* index, *J. Geophys. Res.*, *106*, 30,047–30,066, doi:10.1029/2000JA900162.

Boynton, R. J., M. A. Balikhin, S. A. Billings, A. S. Sharma, and O. A. Amariutei (2011), Data derived NARMAX *Dst* model, *Ann. Geophys.*, *29*, 965–971, doi:10.5194/angeo-29-965-2011.

Buckley, J. J., and Y. Hayashi (1994), Fuzzy neural networks: A survey, *Fuzzy Set. Syst.*, *66*, 1–13, doi:10.1016/0165-0114(94)90297-6.

Burch, J. L. (1974), Observations of interactions between interplanetary and geomagnetic fields, *Rev. Geophys.*, *12*, 363–378, doi:10.1029/RG012i003p00363.

Burton, R. K., R. L. McPherron, and C. T. Russell (1975), An empirical relationship between interplanetary conditions and *Dst*, *J. Geophys. Res.*, *80*, 4204–4214, doi:10.1029/JA080i031p04204.

Chen, S., and S. A. Billings (1989), Representations of non-linear systems: The NARMAX model, *Int. J. Control*, *49*, 1013–1032, doi:10.1080/00207178908559683.

Cid, C., J. Palacios, E. Saiz, Y. Cerrato, J. Aguado, and A. Guerrero (2013), Modeling the recovery phase of extreme geomagnetic storms, *J. Geophys. Res. Space Phys.*, *118*, 4352–4359, doi:10.1002/jgra.50409.

Clerc, M., and J. Kennedy (2002), The particle swarm—Explosion, stability, and convergence in a multidimensional complex space, *IEEE Trans. Evol. Comput.*, *6*, 58–73, doi:10.1109/4235.985692.

Drezet, P. M., R. F. Harrison, and M. Balikhin (2002), A kernel-based technique for forecasting geomagnetic activity and prediction of *Dst*, *Adv. Space Res.*, *30*, 2181–2188, doi:10.1016/S0273-1177(02)80216-2.

Dubey, S. C., and A. P. Mishra (2000), Characteristics of large geomagnetic storms observed during solar cycle 22, *Indian J. Radio Space Phys.*, *29*, 51–60.

Dungey, J. W. (1961), Interplanetary magnetic field and the auroral zones, *Phys. Rev. Lett.*, *6*, 47–48, doi:10.1103/PhysRevLett.6.47.

Eberhart, R. C., and J. Kennedy (1995), A new optimizer using particle swarm theory, in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43, IEEE Press, Piscataway, N. J., doi:10.1109/MHS.1995.494215.

Eberhart, R. C., and Y. Shi (2000), Comparing inertia weights and constriction factors in particle swarm optimization, in *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 84–88, IEEE Press, Piscataway, N. J., doi:10.1109/CEC.2000.870279.

Echer, E., W. D. Gonzalez, F. L. Guarnieri, A. Dal Lago, and L. E. A. Vieira (2005), Introduction to space weather, *Adv. Space Res.*, *35*, 855–865, doi:10.1016/j.asr.2005.02.098.

Echer, E., W. D. Gonzalez, and B. T. Tsurutani (2008), Interplanetary conditions leading to superintense geomagnetic storms ($Dst \leq -250$ nT) during solar cycle 23, *Geophys. Res. Lett.*, *35*, L06S03, doi:10.1029/2007GL031755.

Elman, J. L. (1990), Finding structure in time, *Cognit. Sci.*, *14*, 179–211, doi:10.1207/s15516709cog1402_1.

Fenrich, F. R., and J. G. Luhmann (1998), Geomagnetic response to magnetic clouds of different polarity, *Geophys. Res. Lett.*, *25*, 2999–3002, doi:10.1029/98GL51180.

Gleisner, H., H. Lundstedt, and P. Wintoft (1996), Predicting geomagnetic storms from solar-wind data using time-delay neural networks, *Ann. Geophys.*, *14*, 679–686, doi:10.1007/s00585-996-0679-1.

Gonzalez, W. D., J. A. Joselyn, Y. Kamide, H. W. Kroehl, G. Rostoker, B. T. Tsurutani, and V. M. Vasyliunas (1994), What is a geomagnetic storm?, *J. Geophys. Res.*, *99*, 5771–5792, doi:10.1029/93JA02867.

Gonzalez, W. D., B. T. Tsurutani, and A. L. Clúa de Gonzalez (1999), Interplanetary origin of geomagnetic storms, *Space Sci. Rev.*, *88*, 529–562, doi:10.1023/A:1005160129098.

Grimaldi, E. A., F. Grimaccia, M. Mussetta, and R. E. Zich (2004), PSO as an effective learning algorithm for neural network applications, in *Proceedings of the 3rd International Conference on Computational Electromagnetics and Its Applications*, pp. 557–560, IEEE Press, Piscataway, N. J., doi:10.1109/ICCEA.2004.1459416.

Hapgood, M. A. (2011), Towards a scientific understanding of the risk from extreme space weather, *Adv. Space Res.*, *47*, 2059–2072, doi:10.1016/j.asr.2010.02.007.

Hathaway, D. H. (2015), The solar cycle, *Living Rev. Sol. Phys.*, *12*, 4, doi:10.1007/lrsp-2015-4.

Haykin, S. (1998), *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, N. J.

Iyemori, T., H. Maeda, and T. Kamei (1979), Impulse response of geomagnetic indices to interplanetary magnetic fields, *J. Geomag. Geoelectr.*, *31*, 1–9, doi:10.5636/jgg.31.1.

Jankovičová, D., P. Dolinský, F. Valach, and Z. Vörös (2002), Neural network-based nonlinear prediction of magnetic storms, *J. Atmos. Sol. Terr. Phys.*, *64*, 651–656, doi:10.1016/S1364-6826(02)00025-1.

Ji, E.-Y., Y.-J. Moon, N. Gopalswamy, and D.-H. Lee (2012), Comparison of *Dst* forecast models for intense geomagnetic storms, *J. Geophys. Res.*, *117*, A03209, doi:10.1029/2011JA016872.

Jordanova, V. K., L. M. Kistler, J. U. Kozyra, G. V. Khazanov, and A. F. Nagy (1996), Collisional losses of ring current ions, *J. Geophys. Res.*, *101*, 111–126, doi:10.1029/95JA02000.

Joselyn, J. A. (1995), Geomagnetic activity forecasting: The state of the art, *Rev. Geophys.*, *33*, 383–401, doi:10.1029/95RG01304.

Kamide, Y., N. Yokoyama, W. Gonzalez, B. T. Tsurutani, I. A. Daglis, A. Brekke, and S. Masuda (1998), Two-step development of geomagnetic storms, *J. Geophys. Res.*, *103*, 6917–6921, doi:10.1029/97JA03337.

Kang, K., J.-H. Oh, C. Kwon, and Y. Park (1993), Generalization in a two-layer neural network, *Phys. Rev. E*, *48*, 4805, doi:10.1103/PhysRevE.48.4805.

Kappenman, J. G. (2005), An overview of the impulsive geomagnetic field disturbances and power grid impacts associated with the violent Sun–Earth connection events of 29–31 October 2003 and a comparative evaluation with other contemporary storms, *Space Weather*, *3*, S08C01, doi:10.1029/2004SW000128.

Kappernman, J., and V. Albertson (1990), Bracing for the geomagnetic storms, *IEEE Spectrum*, *27*, 27–33, doi:10.1109/6.48847.

Karinen, A., and K. Mursula (2005), A new reconstruction of the *Dst* index for 1932–2002, *Ann. Geophys.*, *23*, 475–485, doi:10.5194/angeo-23-475-2005.

Kasinskii, V. V., N. G. Ptitsyna, N. N. Lyahov, M. I. Tyasto, G. Villoresi, and N. Iucci (2007), Effect of geomagnetic disturbances on the operation of railroad automated mechanisms and telemechanics, *Geomagn. Aeron.*, *47*, 676–680, doi:10.1134/S0016793207050179.

Khabarova, O. V. (2007), Current problems of magnetic storm prediction and possible ways of their solving, *Sun Geosphere*, *2*, 32–37.

Kennedy, J., and R. Eberhart (1995), Particle swarm optimization, in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, IEEE Press, Piscataway, N. J., doi:10.1109/ICNN.1995.488968.

Kakioka Magnetic Observatory (KMO) (2016), Japan Meteorological Agency, Ibaraki Prefecture, Japan. [Available at http://www.kakioka-jma.go.jp/.]

Kohonen, T. (1982), Self-organized formation of topologically correct feature maps, *Biol. Cybern.*, *43*, 59–69, doi:10.1007/bf00337288.

Kozyra, J. U., V. K. Jordanova, R. B. Home, and R. M. Thorne (1997), Modeling the contribution of electromagnetic ion cyclotron (EMIC) waves to stormtime ring current erosion, in *Magnetic Storms, Geophys. Monogr. Ser.*, edited by B. T. Tsurutani et al., pp. 187–202, AGU, Washington D. C., doi:10.1029/GM098p0187.

Kozyra, J. U., M.-C. Fok, E. R. Sanchez, D. S. Evans, D. C. Hamilton, and A. F. Nagy (1998), The role of precipitation losses in producing the rapid early recovery phase of the Great Magnetic Storm of February 1986, *J. Geophys. Res.*, *103*, 6801–6814, doi:10.1029/97JA03330.

Kugblenu, S., S. Taguchi, and T. Okuzawa (1999), Prediction of the geomagnetic storm associated *Dst* index using an artificial neural network algorithm, *Earth Planet Space*, *51*, 307–313, doi:10.1186/BF03352234.

Lakhina, G. S., and B. T. Tsurutani (2016), Geomagnetic storms: Historical perspective to modern view, *Geosci. Lett.*, *3*, 5, doi:10.1186/s40562-016-0037-4.

Lazzús, J. A., I. Salfate, and S. Montecinos (2014), Hybrid neural network-particle swarm algorithm to describe chaotic time series, *Neural Netw. World*, *24*, 601–617, doi:10.14311/NNW.2014.24.034.

Lazzús, J. A., M. Rivera, and C. H. López-Caraballo (2016), Parameter estimation of Lorenz chaotic system using a hybrid swarm intelligence algorithm, *Phys. Lett. A*, *380*, 1164–1171, doi:10.1016/j.physleta.2016.01.040.

López-Caraballo, C. H., J. A. Lazzús, I. Salfate, P. Rojas, M. Rivera, and L. Palma-Chilla (2015), Impact of noise on a dynamical system: Prediction and uncertainties from a swarm-optimized neural network, *Comput. Intell. Neurosci.*, *2015*, 145874, doi:10.1155/2015/145874.

Lotfi, E., and M.-R. Akbarzadeh-T (2014), Adaptive brain emotional decayed learning for online prediction of geomagnetic activity indices, *Neurocomputing*, *126*, 188–196, doi:10.1016/j.neucom.2013.02.040.

Lundstedt, H. (1993), A trained neural network, geomagnetic activity and solar wind variation, in *Proceedings of the Workshop on Solar-Terrestrial Prediction*, edited by J. Hruška et al., p. 607, Natl. Oceanic and Atmos. Admin., Space Environ. Lab., Boulder, Colo.

Lundstedt, H. (2005), Progress in space weather predictions and applications, *Adv. Space Res.*, *36*, 2516–2523, doi:10.1016/j.asr.2003.09.072.

Lundstedt, H., and P. Wintoft (1994), Prediction of geomagnetic storms from solar wind data with the use of a neural network, *Ann. Geophys.*, *12*, 19–24, doi:10.1007/s00585-994-0019-2.

Lundstedt, H., H. Gleisner, and P. Wintoft (2002), Operational forecasts of the geomagnetic *Dst* index, *Geophys. Res. Lett.*, *29*, 2181, doi:10.1029/2002GL016151.

Malville, J. M. (1960), The effect of the initial phase of a magnetic storm upon the outer Van Allen belt, *J. Geophys. Res.*, *65*, 3008–3010, doi:10.1029/JZ065i009p03008.

Mandic, D., and J. Chambers (2001), *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, Wiley, New York.

Mirmomeni, M., M. Shafiee, C. Lucas, and B. N. Araabi (2006), 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, doi:10.1016/j.jastp.2006.07.001.

Mirmomeni, M., C. Lucas, B. Moshiri, and B. N. Araabi (2010), Introducing adaptive neurofuzzy modeling with online learning method for prediction of time-varying solar and geomagnetic activity indices, *Exp. Syst. Appl.*, *37*, 8267–8277, doi:10.1016/j.eswa.2010.05.059.

Mirmomeni, M., C. Lucas, B. N. Araabi, B. Moshiri, and M. R. Bidar (2011), Online multi-step ahead prediction of time-varying solar and geomagnetic activity indices via adaptive neurofuzzy modeling and recursive spectral analysis, *Sol. Phys.*, *272*, 189–213, doi:10.1007/s11207-011-9810-x.

Munsami, V. (2000), Determination of the effects of substorms on the storm-time ring current using neural networks, *J. Geophys. Res.*, *105*, 27,833–27,840, doi:10.1029/2000JA000041.

O'Brien, T. P., and R. L. McPherron (2000), Forecasting the ring current index *Dst* in real time, *J. Atmos. Sol. Terr. Phys.*, *62*, 1295–1299, doi:10.1016/S1364-6826(00)00072-9.

Ouarbya, L., D. T. Mirikitani, and E. Martin (2012), The use of sequential recurrent neural filters in forecasting the *Dst* index for the strong magnetic storm of autumn 2003, *Appl. Math. Lett.*, *25*, 1361–1366, doi:10.1016/j.aml.2011.12.002.

Palit, A. K., and D. Popovic (2005), *Computational Intelligence in Time Series Forecasting Theory and Engineering Applications*, Springer-Verlag, London, U. K.

Pallocchia, G., E. Amata, G. Consolini, M. F. Marcucci, and I. Bertello (2006), Geomagnetic *Dst* index forecast based on IMF data only, *Ann. Geophys.*, *24*, 989–999, doi:10.5194/angeo-24-989-2006.

Pandey, S. K., and S. C. Dubey (2009), Characteristic features of large geomagnetic storms observed during solar cycle 23, *Indian J. Radio Space Phys.*, *38*, 305–312.

Pesnell, W. D. (2014), Predicting solar cycle 24 using a geomagnetic precursor pair, *Sol. Phys.*, *289*, 2317–2331, doi:doi:10.1007/s11207-013-0470-x.

Podladchikova, T. V., and A. A. Petrukovich (2012), Extended geomagnetic storm forecast ahead of available solar wind measurements, *Space Weather*, *10*, S07001, doi:10.1029/2012SW000786.

Piddington, J. H. (1963), Theories of the geomagnetic storm main phase, *Planet. Space Sci.*, *11*, 1277–1288, doi:10.1016/0032-0633(63)90232-0.

Poli, R., J. Kennedy, and T. Blackwell (2007), Particle swarm optimization. An overview, *Swarm Intell.*, *1*, 33–57, doi:10.1007/s11721-007-0002-0.

Poulton, M. M. (2002), Neural networks as an intelligence amplification tool: A review of applications, *Geophysics*, *67*, 979–993, doi:10.1190/1.1484539.

Puskorius, G. V., and L. A. Feldkamp (1994), Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks, *IEEE Trans. Neural Networks*, *5*, 279–297, doi:10.1109/72.279191.

Revallo, M., F. Valach, P. Hejda, and J. Bochníček (2014), A neural network *Dst* index model driven by input time histories of the solar wind-magnetosphere interaction, *J. Atmos. Sol. Terr. Phys.*, 9–14, doi:10.1016/j.jastp.2014.01.011. 110-111.

Rostoker, G. (1972), Geomagnetic indices, *Rev. Geophys.*, *10*, 935–950, doi:10.1029/RG010i004p00935.

Rumelhart, D. E., and J. McClelland (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press, Cambridge, Mass.

Sckopke, N. (1966), A general relation between the energy of trapped particles and the disturbance field near the Earth, *J. Geophys. Res.*, *71*, 3125–3130, doi:10.1029/JZ071i013p03125.

Sharifi, J., B. N. Araabi, and C. Lucas (2006a), Multi-step prediction of *Dst* index using singular spectrum analysis and locally linear neurofuzzy modeling, *Earth Planet Space*, *58*, 331–341, doi:10.1186/BF03351929.

Sharifi, J., C. Lucas, and B. N. Araabi (2006b), Locally linear neurofuzzy modeling and prediction of geomagnetic disturbances based on solar wind conditions, *Space Weather*, *4*, S06003, doi:10.1029/2005SW000209.

Siegelmann, H. T., B. G. Horne, and C. L. Giles (1997), Computational capabilities of recurrent NARX neural networks, *IEEE Trans. Syst. Man. Cybern. B*, *27*, 208–215, doi:10.1109/3477.558801.

Singh, A. K., D. Siingh, and R. P. Singh (2010), Space weather: Physics, effects and predictability, *Surv. Geophys.*, *31*, 581–638, doi:10.1007/s10712-010-9103-1.

Smith, Z. K., W. Murtagh, and C. Smithtro (2004), Relationship between solar wind low-energy energetic ion enhancements and large geomagnetic storms, *J. Geophys. Res.*, *109*, A01110, doi:10.1029/2003JA010044.

Sola, J., and J. Sevilla (1997), Importance of input data normalization for the application of neural networks to complex industrial problems, *IEEE Trans. Nucl. Sci.*, *44*, 1464–1468, doi:10.1109/23.589532.

Stepanova, M. V., and P. Pérez (2000), Autoprediction of *Dst* index using neural network techniques and relationship to the auroral geomagnetic indices, *Geofís. Int.*, *39*, 143–146.

Stepanova, M., E. Antonova, and O. Troshichev (2005), Prediction of *Dst* variations from polar cap indices using time-delay neural network, *J. Atmos. Sol. Terr. Phys.*, *67*, 1658–1664, doi:10.1016/j.jastp.2005.02.027.

Sugiura, M. (1964), Hourly values of equatorial *Dst* for the IGY, *Ann. Int. Geophys. Year*, *35*, 9–45.

Sugiura, M., and T. Kamei (1991), Equatorial *Dst* index 1957–1986, in *International Association of Geomagnetism and Aeronomy Bulletin*, vol. 40, edited by A. Berthelier and and M. Menvielle, Int. Serv. of Geomagn. Indices Publ. Off., Saint-Maur-des-Fosses, France.

Temerin, M., and X. Li (2002), A new model for the prediction of *Dst* on the basis of the solar wind, *J. Geophys. Res.*, *107*, 1472, doi:10.1029/2001JA007532.
Thomson, A. W. P. (2000), Evaluating space weather forecasts of geomagnetic activity from a user perspective, *Geophys. Res. Lett.*, *27*, 4049–4052, doi:10.1029/2000GL011908.
Tsagouri, I., A. Belehaki, and L. R. Cander (2005), A dynamic system to forecast ionospheric storm disturbances based on solar wind conditions, *Ann. Geophys.*, *48*, 467–475, doi:10.4401/ag-3210.
Tsurutani, B. T., and R. M. Thorne (1982), Diffusion processes in the magnetopause boundary layer, *Geophys. Res. Lett.*, *9*, 1247–1250, doi:10.1029/GL009i011p01247.
van der Baan, M., and C. Jutten (2000), Neural networks in geophysical applications, *Geophysics*, *65*, 1032–1047, doi:10.1190/1.1444797.
Veselovsky, I. S., et al. (2004), Solar and heliospheric phenomena in October–November 2003: Causes and effects, *Cosmic Res.*, *42*, 435–488, doi:10.1023/B:COSM.0000046229.24716.02.
Vichare, G., S. Alex, and G. S. Lakhina (2005), Some characteristics of intense geomagnetic storms and their energy budget, *J. Geophys. Res.*, *110*, A03204, doi:10.1029/2004JA010418.
Vichare, G., R. Rawat, A. Bhaskar, and B. M. Pathan (2014), Ionospheric current contribution to the main impulse of a negative sudden impulse, *Earth Planet Space*, *66*, 92, doi:10.1186/1880-5981-66-92.
Vörös, Z., and D. Jankovičová (2002), Neural network prediction of geomagnetic activity: A method using local Hölder exponents, *Nonlin. Process. Geophys.*, *9*, 425–433, doi:10.5194/npg-9-425-2002.
Wang, C. B., J. K. Chao, and C. H. Lin (2003), Influence of the solar wind dynamic pressure on the decay and injection of the ring current, *J. Geophys. Res.*, *108*, 1341, doi:10.1029/2003JA009851.
Watanabe, S., E. Sagawa, K. Ohtaka, and H. Shimazu (2002), Prediction of the *Dst* index from solar wind parameters by a neural network method, *Earth Planet Space*, *54*, e1263–e1275, doi:10.1186/BF03352454.
Wei, H. L., D. Q. Zhu, S. A. Billings, and M. A. Balikhin (2007), Forecasting the geomagnetic activity of the *Dst* index using multiscale radial basic function Network, *Adv. Space Res.*, *40*, 1863–1870, doi:10.1016/j.asr.2007.02.080.
Wilks, D. S. (2011), *Statistical Methods in the Atmospheric Sciences*, Elsevier, San Diego, Calif.
Williams, D. J. (1995), Global energy flow in the magnetosphere-ionosphere system, *J. Geomagn. Geoelectr.*, *47*, 1147–1160, doi:10.5636%2Fjgg.47.1147.
Williams, R. J., and D. Zipser (1989), A learning algorithm for continually running recurrent neural networks, *Neural Comput.*, *1*, 270–280, doi:10.1162/neco.1989.1.2.270.
Wintoft, P., and H. Lundstedt (1993), Geomagnetic activity and large-scale solar magnetic field structures with two neural network paradigms, in *Proceedings of the International Workshop on Artificial Intelligence Applications in Solar-Terrestrial Physics*, edited by J. Josylen, H. Lundstedt, and J. Trolinger, pp. 37–43, Natl. Oceanic and Atmos. Admin., Space Environ. Lab., Boulder, Colo.
Witt, A., and B. D. Malamud (2013), Quantification of long-range persistence in geophysical time series: Conventional and benchmark-based improvement techniques, *Surv. Geophys.*, *34*, 541–651, doi:10.1007/s10712-012-9217-8.
World Data Center (WDC) (2016), World Data Center for Geomagnetism: Kyoto, Japan. [Available at http://wdc.kugi.kyoto-u.ac.jp/, accessed date October, 2016.]
Wu, J.-G., and H. Lundstedt (1996), Prediction of geomagnetic storms from the solar wind data using Elman recurrent neural networks, *Geophys. Res. Lett.*, *23*, 319–322, doi:10.1029/96GL00259.
Wu, J.-G., and H. Lundstedt (1997), Geomagnetic storm predictions from solar wind data with the use of dynamic neural networks, *J. Geophys. Res.*, *102*, 14255–14268, doi:10.1029/97JA00975.
Xue, B., and J. Gong (2006), Forecasting *Dst* index with artificial neural network, *Chin. J. Space Sci.*, *26*, 183–186, doi:10.11728/cjss2006.03.183.
Yermolaev, Y. I., I. G. Lodkina, N. S. Nikolaeva, and M. Y. Yermolaev (2013), Occurrence rate of extreme magnetic storms, *J. Geophys. Res. Space Phys.*, *118*, 4760–4765, doi:10.1002/jgra.50467.
Yuan, S., S. Wang, and N. Tian (2009), Swarm intelligence optimization and its application in geophysical data inversion, *Appl. Geophys.*, *6*, 166–174, doi:10.1007/s11770-009-0018-x.
Zhang, J.-R., J. Zhang, T.-M. Lok, and M. R. Lyu (2007), A hybrid particle swarm optimization-back-propagation algorithm for feedforward neural network training, *Appl. Math. Comput.*, *185*, 1026–1037, doi:10.1016/j.amc.2006.07.025.
Zomaya, A. Y. (2006), *Handbook of Nature-Inspired and Innovative Computing*, Springer, New York.