Journal of Physics: Conference Series **720** (2016) 012001 doi:10.1088/1742-6596/720/1/012001

Forecasting of DST index from auroral electrojet indices using time-delay neural network + particle swarm optimization

J A Lazzús, C H López-Caraballo, P Rojas, I Salfate, M Rivera and L Palma-Chilla

Departamento de Física y Astronomía, Universidad de La Serena, Casilla 554, La Serena, Chile

E-mail: jlazzus@dfuls.cl

Abstract. In this study, an artificial neural network was optimized with particle swarm algorithm and trained to predict the geomagnetic DST index one hour ahead using the past values of DST and auroral electrojet indices. The results show that the proposed neural network model can be properly trained for predicting of $DST(t+1)$ with acceptable accuracy, and that the geomagnetic indices used have influential effects on the good training and predicting capabilities of the chosen network.

1. Introduction

Geomagnetic activity is usually characterized by geomagnetic indices, the most common being the storm DST index and the auroral electrojet indices [1]. Most indices have long records that allow statistical studies of the causes of activity and of related phenomena. Correlations between indices and possible drivers provide the basis for empirical prediction. The $DST(t)$ is defined as the average of the disturbance variation of the H component, divided by the average of the cosines of the dipole latitudes at the observatories for normalization to the dipole equator [2]. DST index can serve as a good measure of the overall strength of the near-Earth electric currents, especially the ring current. It is obtained from selected geomagnetic observatories operating at equatorial regions. It is important to realize that the ultimate goal of space weather research should be to forecast space weather and, at least, predict geomagnetic activity in terms of the geomagnetic DST index as a function of time.

The auroral electrojet index was originally introduced by Davis and Sugiura [3] as a measure of global electrojet activity in the auroral zone. The auroral electrojet index is derived from geomagnetic variations in the horizontal component H observed from selected observatories along the auroral zone in the northern hemisphere [4]. The auroral electrojet index is usually used to represent four indices AU, AL, AE and AO. The AU and AL indices [3], are intended to express the strongest current intensity of the eastward and westward auroral electrojets, respectively [4]. The AE index is defined as $AE = AU - AL$ [3], provides an estimate of the overall horizontal current strength, and, to some extent, a rough measure of the ionospheric energy losses [5]. And the AO index is defined as $AO = (AU + AL)/2$ [3], provides a measure of the equivalent zonal current [6].

In this work, the DST one hour ahead was estimate with a hybrid neural network+particle swarm algorithm using the time series of the past values of DST and auroral electrojet (AE and AO) indices.

2. Computational method

In this study, we consider one of the most successful and frequently used types of neural networks: a multilayer feed-forward neural network with a back-propagation learning algorithm. This consists of one input layer with N inputs, one hidden layer with q units and one output layer with n outputs. The output of this model can be expressed as [7]:

$$y_n = F_n \left(\sum_{j=1}^{q} W_{nj}^{(a)} f_j \left(\sum_{l=1}^{N} W_{jl}^{(b)} x_l + W_{j0}^{(b)} \right) + W_{n0}^{(a)} \right) \tag{1}$$

where $W_{nj}^{(a)}$ are the weights between unit j and unit n of input and hidden layers and $W_{jl}^{(b)}$ are the weights between hidden layer and an output. The activation functions $F_n(x)$ and $f_j(x)$ are linear or nonlinear. We used one hidden layer with $f_j(x)$ as a tangent hyperbolic nonlinear activation functions and $F_1(x)$ as linear function in output layer. For a given set of N inputs, we define the root mean square error (RMSE) by:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{s=1}^{N} \left(y_s^{real} - y_s^{calc} \right)^2}{N^2}} \tag{2}$$

where y^{real} denotes the actual given output and y^{calc} the neural network output. This network was trained to minimize RMSE, replacing the gradient descendent error by a paricle swarm optimization (PSO) [8].

PSO is a population-based optimization tool, where the system is initialized with a population of random particles and the algorithm searches for optima by updating generations [9]. In each iteration, the velocity of each particle j is calculated according to the following formula [10]:

$$v_j^{k+1} = \omega v_j^k + c_1 r_1 \left(\psi_j^k - s_j^k \right) + c_2 r_2 \left(\psi_g^k - s_j^k \right) \tag{3}$$

where s and v denote a particle position and its corresponding velocity in a search space, respectively. k is the current step number, ω is the inertia weight, c_1 and c_2 are the acceleration constants, and r_1, r_2 are elements from two random sequences in the range (0,1). s_j^k is the current position of the particle, ψ_j^k is the best one of the solutions that this particle has reached, and ψ_g is the best solutions that all the particles have reached. In general, the value of each component in v can be clamped to the range $[-v_{\max}, +v_{\max}]$ control excessive roaming of particles outside the search space [9]. After calculating the velocity, the new position of each particle is:

$$s_j^{k+1} = s_j^k + v_j^{k+1} \tag{4}$$

Thus, all the neurons of the ANN have an associated activation value for a give input pattern, and the algorithm continues finding the error that is presented for each neuron, except those of the input layer. After finding the output values, the weights of all layers of the network are actualized $W_{nj} \rightarrow W'_{nj}$ by PSO, using eqs. (3 and 4) [8]. The velocity is used to control how much the position is updated. On each step, PSO compares each weight using the data set. The network with the highest fitness is considered the global best. The other weights are updated based on the global best network rather than on their personal error or fitness [9]. In PSO, the inertial weight ω, the constant c_1 and c_2, the number of particles N_{part} and the maximum speed of particle summary the parameters to syntonize for their application in a given problem. An exhaustive trial-and-error procedure was applied for tuning the PSO parameters. Firstly,

the effect of ω was analyzed for values of 0.1 to 0.9. Next, the effect of N_{part} was analyzed for values of 10 to 100 particles in the swarm. Table 1 shows the selected parameters for this hybrid algorithm.

Table 1. Parameters used in the hybrid ANN+PSO algorithm.

Section	Parameter	Value
ANN	NN-type	*feed-forward*
	Number of hidden layers	1
	Transfer function	*tansig*
	Number of iterations	1500
	Normalization range	[–1, 1]
	Weight range	[–100, 100]
	Bias range	[–10, 10]
	Minimun error	1e–3
PSO	Number of particles in swarm (N_{part})	50
	Number of iterations ($k_{\max}$)	1500
	Cognitive component (c_1)	1.494
	Social component (c_2)	1.494
	Maximum velocity ($v_{\max}$)	12
	Minimum inertia weight ($\omega_{\min}$)	0.5
	Maximum inertia weight ($\omega_{\max}$)	0.7
	Objective function	RMSE

3. Database and training

A data set of geomagnetic DST index and auroral electrojet indices were taken from World Data Center for Geomagnetism of Kyoto [11], and used to train the network. This work used a leave-20%-out cross-validation method to estimate the predictive capabilities of the model. Training and prediction sets were selected with the consideration that the DST variations are present with adequate frequency in the training database. Then, 43848 hourly data points (years 2000 to 2004) were used in the training set, and other 8760 hourly data points (year 2005) were used in the prediction set. From this database, the best input vectors obtained to solve $DST(t+1)$ were:

$$\text{N}[DST(t-3), AE(t-3), DST(t-2), AE(t-2), DST(t-1), AE(t-1), DST(t), AE(t)] \quad (5)$$

$$\text{N}[DST(t-3), AO(t-3), DST(t-2), AO(t-2), DST(t-1), AO(t-1), DST(t), AO(t)] \quad (6)$$

The selected geomagnetic indices cover wide ranges, going from -442 to 67 (nT) for DST index, from 4 to 2241 for AE index, and from -942 to 1071 for AO index. Figure 1 shows the time series used and ranges for these indices.

4. Results and discussion

Several network architectures were tested to select the most accurate scheme. The most basic architecture normally used for this type of application involves a neural network consisting of three layers [8]. The number of hidden neurons needs to be sufficient to ensure that the

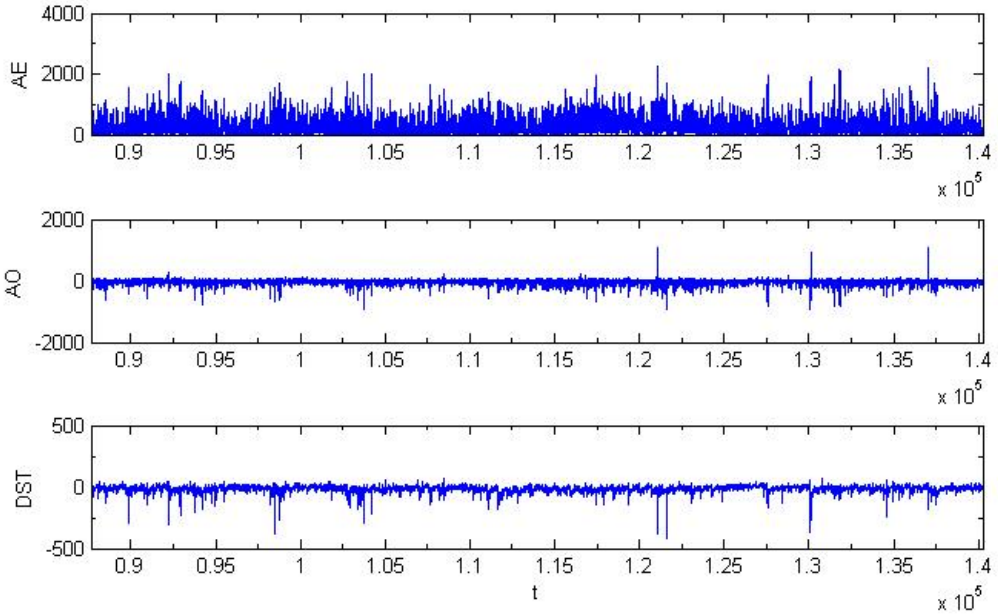

Figure 1. Time series of geomagnetic indices used in this study.

information contained in the data utilized for training the network is adequately represented. There is no specific approach to determine the number of neurons of the hidden layer, but many alternative combinations are possible. The optimum number of neurons was determined by adding neurons in systematic form and evaluating the average absolute deviations of the sets during the learning process [9]. Figure 2 shows the deviation found in the prediction of $DST(t+1)$ as a function of the number of neurons in the hidden layer. As observed in this figure, the optimum number of neurons in the hidden layer is between 6 and 9. Table 2 summarize the statistical results obtained in the estimation of $DST(t+1)$ using the selected input. In this table MAE is the mean absolute error, RMSE is the root mean square error, and R^2 is the correlation coefficient.

Table 2. Statistical results obtained in the estimation of $DST(t+1)$.

Input	Architecture	Step	MAE	RMSE	R^2
Eq. (5)	8-6-1	Training	2.7070	4.2852	0.9829
		Prediction	1.9991	2.8671	0.9661
		Total	2.5654	4.0416	0.9824
Eq. (6)	8-8-1	Training	2.7564	4.3279	0.9826
		Prediction	2.0112	2.8971	0.9653
		Total	2.6074	4.0821	0.9820

Figure 3 shows a comparison between the real (solid line) and calculated values (points) of $DST(t+1)$, obtained by the neural network models. Fig. 3a show the results obtained using

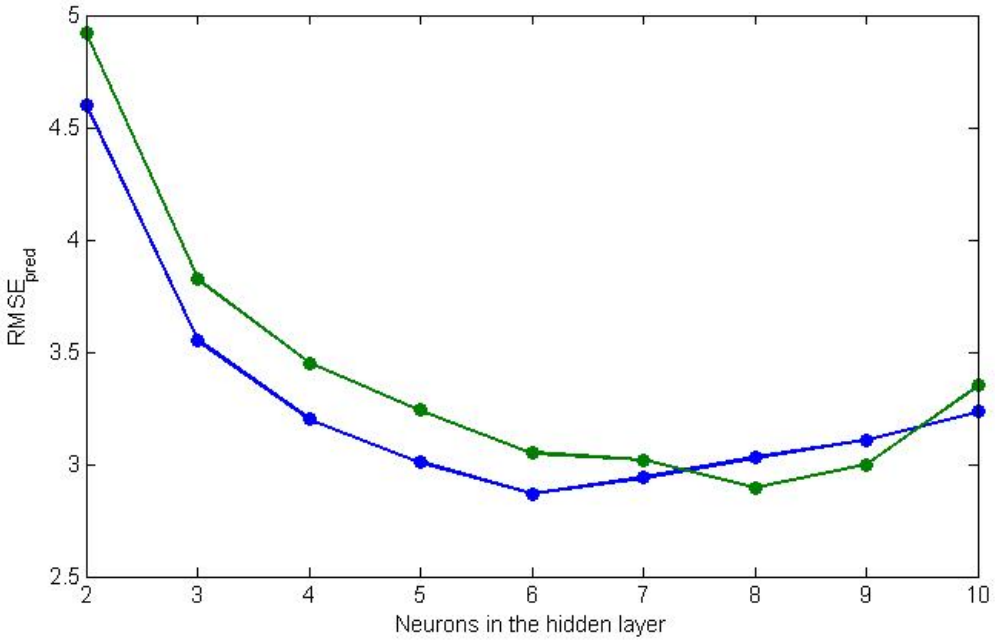

Figure 2. Time series of geomagnetic indices used in this study. $DST + AE$ (blue line), and $DST + AO$ (green line).

the past values of DST and AE (Eq. 5). For the training (years 2000–2004), the correlation coefficient R^2 is 0.9829 and the slope of the curve m is 0.9691 (expected to be 1.0) with RMSE of 4.2852. For the prediction (year 2005), R^2 is 0.9661, m is 1.0102 (expected to be 1.0) and RMSE of 2.8671. Fig. 3b show the results obtained with Eq. 6. In this case, for the training step (years 2000–2004), R^2 is 0.9826, m is 0.9769 and RMSE is 4.3279. And for the prediction step (year 2005), R^2 is 0.9653, m is 1.0095 and RMSE is 2.8971.

Currently, various models have been developed to predict DST [12]. However, comparative studies of artificial neural networks (ANNs) and the traditional regression approaches in modelling DST have also been conducted, and it has been shown that ANN methodology offers a promising alternative to the traditional approach [13]. In this case, a comparison was made with a multiple linear regression (MLR) method, and similar database. The MLR method show errors higher than 500% and accuracy of <50% with R^2 of 0.5 for the prediction of $DST(t+1)$ using $DST + AE$ versus the ANN+PSO method with accuracy of >97% and R^2 >0.96. Similar results were obtained for the prediction of $DST(t+1)$ using $DST + AO$. These results represent a tremendous increase in accuracy for estimating this important geomagnetic index and show that not only the optimum architecture obtained was crucial, also the appropriate selection of the independent parameters (Eq. 5 and 6). This is important because the auroral electrojet indices AE and AO are commonly available parameters [11]. Note that the coefficients of linear correlation of these parameters show a non-linear relationship with the DST [1], and consequently an ANN is the best alternative to model the DST for several applications. The innovative aspect of this study is the use of a neural model with only two input variables and a very limited number of neurons in the hidden layer. This ANN was optimized with a particle swarm algorithm to update the weight of the network. To the best of the authors's knowledge, there is no application for DST forecast such as the one presented here, using a hybrid ANN+PSO algorithm.

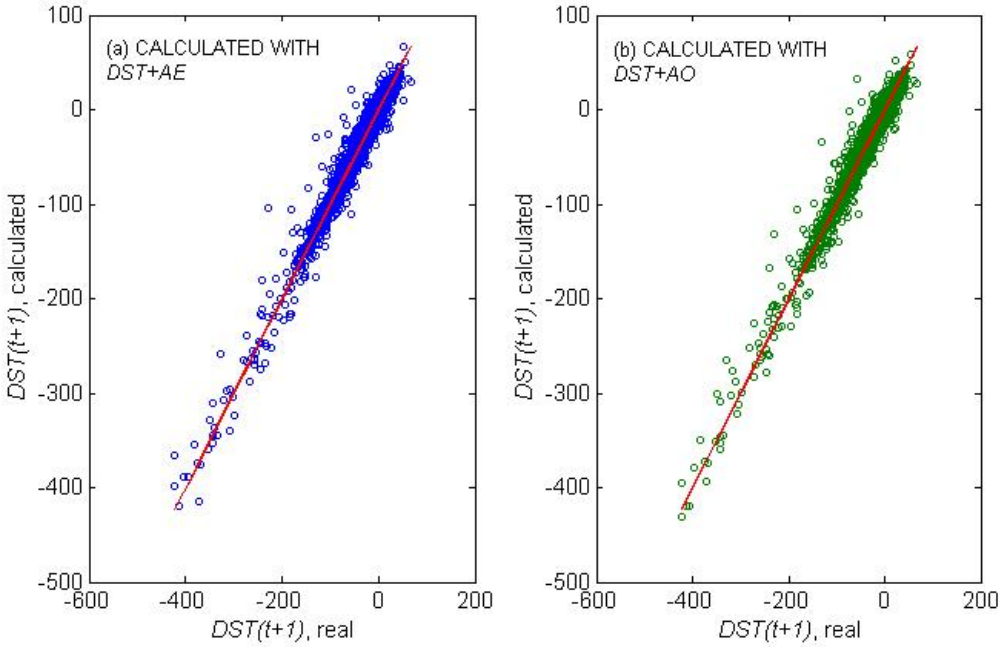

Figure 3. Comparison between calculated and real values for the estimation of $DST(t+1)$. (a) Calculated using Eq. 5, $DST + AE$. (b) Calculated using Eq. 6, $DST + AO$.

5. Conclusions

This study presents a method that includes an artificial neural network (ANN) replacing standard backpropagation with particle swarm optimization (PSO). Then, DST index one hour ahead was estimate by this hybrid algorithm using the time series of the past values of DST and auroral electrojet (AE and AO) indices.

Based on the results presented in this study, the following main conclusions are obtained: i) The results show that the proposed neural network models with eqs. 5 and 6 can be properly trained for predicting the $DST(t+1)$, with acceptable accuracy; ii) the geomagnetic indices used have influential effects, on the good training and predicting capabilities, of the chosen network; iii) The low deviations found with the proposed method indicate that it can predict the future values of DST with better accuracy than other methods available in the literature.

Acknowledgments

The authors thank the Direction of Research of the University of La Serena (DIULS) for the special support that made possible the preparation of this paper.

References

[1] Plotnikov I Y and Barkova E S 2007 *Adv. Space Res.* **40** 1858
[2] Sugiura M and Kamei T 1991 *IAGA Bulletin No. 40*
[3] Davis T N and Sugiura M 1966 *J. Geophys. Res.* **71** 785
[4] Pallocchia G, Amata E, Consolini G, Marcucci M F and Bertello I 2008 *J. Atmos. Sol.-Terr. Phys.* **70** 663
[5] Ahn B H, Akasofu S I and Kamide Y 1983 *J. Geophys. Res.* **88** 6275
[6] Menvielle M, Iyemori T, Marchaudon A and Nosé M 2011 *Geomagnetic Observations and Models*, ed M Mandea and M Korte (New York: Springer) chapter 8 pp 183–228
[7] Lazzús J A 2009 *Chin. J. Chem. Phys.* **22** 19
[8] Lazzús J A, Salfate I and Montecinos S 2014 *Neural Netw. World* **24** 601

[9] Lazzús J A 2013 *Math. Comput. Model.* **57** 2408
[10] Lazzús J A 2010 *Comput. Math. Appl.* **60** 2260
[11] WDC 2014 *World Data Center for Geomagnetism of Kyoto* (Database)
[12] Watanabe S, Sagawa E, Ohtaka K and Shimazu H 2003 *Adv. Space Res.* **31** 829
[13] Stepanova M, Antonova E and Troshichev O 2005 *J. Atmos. Sol.-Terr. Phys.* **67** 1658