Simulations of the Interplanetary Magnetic Field conditions with NARX networks

A. G. OLIVEIRA[1], M. R. SILVA[1], E. ECHER[2], A. DAL LAGO[2], C. R. BRAGA[2], R. R. S. MENDONÇA[2], N. J. SCHUCH[3], K. MUNAKATA[4], T. KUWABARA[5]

[1] *Institute of Research and Development, University of the Paraíba Valley, São José dos Campos, São Paulo, Brazil*
[2] *Space Geophysics Department, National Institute for Space Research, São José dos Campos, Brazil*
[3] *Southern Space Research Center, National Institute of Space Research, Santa Maria, Brazil*
[4] *Physics Department, Faculty of Science, Shinshu University, Matsumoto, Japan*
[5] *Bartol Research Institute and Department of Physics and Astronomy, University of Delaware, USA*

a.ger@aol.com

Abstract: Currently most of the data about interplanetary plasma with solar origin comes from space born technology. This technology, although very useful, is also very susceptible to space weather conditions. In this work we proposed simulate the intensity of the Interplanetary Magnetic Field (IMF) using ground observation of the cosmic rays anisotropies of the Global Muon Detector Network (GMDN). We used Non-linear Autoregressive with Exogenous inputs (NARX) neural networks as simulation tool. This neural network model can operate in two modes: the series-parallel mode in which we simulate just one step ahead and must be constantly feeded with the correct value of the IMF, and the parallel mode which uses the value estimated by the network in past iteractions to calculate the next one, producing simulations with longer time range and that just depends of the cosmic rays data. The preliminary results indicate that this is an interesting approach that may produce valuable estimations of the Interplanetary Magnetic Field.

Keywords: cosmic rays, muons, artificial neural networks, interplanetary magnetic field.

1 Introduction

Nowadays, we depend of technological devices, which are often composed by satellites orbiting the Earth. These equipments are suceptible to the solar conditions. Some of these satellites are used to monitor the sun and the interplanetary medium, but they are also suceptible and they can be damaged by energetic protons and/or electrons from the sun. Galactic Cosmic Rays (GCR) detected by the Ground Muon Detector Network (GMDN), can be used to infer the space weather conditions. The ground detection of muons is, however, susceptible to environmental factors that should be corrected. Several mathematical models and iterative and computational methods have been used to obtain space weather information.

The Artificial Neural Networks (ANN) technique is one of the methods that can provide valuable resuts about the relation between gound based GCR data and interplanetary solar wind conditions. We can synthesize the definition of Artificial Neural Network as a computational tool capable to generalize by weighting the interrelationship between the elements assigned to the input vector, via a massive parallel processing system [4] and [6]. This definition makes ANNs interesting to study the relationship between the physical variables in the input vector (GCR data) and in the output vector (IMF data). ANNs are also able to forecast the behavior of the space weather in the interval of some days.

In this paper, we propose to use ANN to estimate the intensity of the IIMF from the components of anisotropy vector of GCR, calculated from GMDN observations. As results from the topology and training process of the neural network we obtained estimatives of the IMF intensity from the components of the anisotropy vector. Data from the Advanced Composition Explorer (ACE) satellite which is located in the Lagrangian point (L1) were used both to training the ANN, as to evaluate the estimation's quality.

2 Galactic Cosmic Rays Data

The GCR data are derived from the muon counting rate detected by the GMDN that is in full operation since March 2006. The detectors which compose the GMDN are located in: Nagoya, Japan; Hobart, Australia; São Martinho da Serra, Brazil and in Kuwait City, Kuwait. The geographic position of each detector is presented in Table 1.

To derive the anisotropy vector components of the GCR, [5] obtained the value $I^{obs}_{i,j}$ that is the percentual variation of the hourly counting rate recorded by the j-th channel of the i-th detector of the GMDN. After remove the barometric effect, they fit $\bar{I}^{fit}_{i,j}$ given by relation:

$$\begin{aligned}\bar{I}^{fit}_{i,j} \quad = \quad & \bar{I}^{0}(t)+\bar{\xi}^{GEO}_{x}(t)(c^{1}_{1i,j}cos\omega.t_i - s^{1}_{1i,j}sin\omega.t_i)+\\ & \bar{\xi}^{GEO}_{y}(t)(s^{1}_{1i,j}cos\omega.t_i - c^{1}_{1i,j}sin\omega.t_i)+ \\ & \bar{\xi}^{GEO}_{z}(t)c^{0}_{1i,j}\end{aligned} \tag{1}$$

where $c^1_{1i,j}$, $s^1_{1i,j}$ and $c^0_{1i,j}$ are coupling coefficients determined assuming a rigidity independent of the anisotropy, t is the universal time, $\omega = \pi/12$, t_i is the local time of i-th detector and $\bar{I}^0$, $\bar{\xi}^{GEO}_x$, $\bar{\xi}^{GEO}_y$ and $\bar{\xi}^{GEO}_z$ are the fitted parameters calculated from the trailing moving average over 12 hour.

The terms $\bar{\xi}^{GEO}_x$, $\bar{\xi}^{GEO}_y$ and $\bar{\xi}^{GEO}_z$ are the anisotropy components, and $\bar{I}^0$ corresponds to GCR density, also named as isotropic component of the Equation (1).

Detector	Latitude	Longitude
Nagoya (77 m)	35.12°N	136.97°E
Hobart (18 m)	−42.90°S	147.33°E
São Martinho da Serra (488 m)	−29.44°S	−53.81°W
Kuwait (59 m)	29.24°N	47.97°E

Table 1: Geographic coordinates of the detectors of Global Muon Detector Network. Beside the identification of detector its altitude above the sea level.

3 Artificial Neural Networks

The concept of Artificial Neural Networks must necessarily pass by three elements: the node, the topology and the learning algorithm.

Nodes or neurons are the simplest components of a neural network. They are information processing units which receive stimulation from the input vector. In each transmission of this stimulation signal, an adjustable weight parameter modifies the intensity of the signal. It is possible to add another adjustable parameter to each node, named *bias*. The mathematical description of the i-th node of an ANN as [8]:

$$u_i = \sum_{j=1}^{n} w_{ij} x_j \tag{2}$$

$$y_i = \varphi\left(u_i + b_i\right) \tag{3}$$

where x_j is the j-th input signal in the node, w_{ij} is the j-th synaptic weight of the node i. The output signal of the node is a result of the activation function φ calculated by the linear combination u_i added of the *bias* value b_i.

There are different topologies (different ways to conect one node to another), that can be cyclic, where we use the output signal of a node as input of the same or some predecessor node. In this paper we use Nonlinear AutoRegressive with eXogenous inputs model of ANN or NARX network. In NARX model the output signal of the network composes the input vector of the network using delay operators. A mathematical formulation to the output final response of a NARX network can be expressed by ([2]):

$$\begin{aligned} y(n+1) \quad = \quad & f[y(n), y(n-1), ..., \\ & y(n-d_y+1); x(n), x(n-1), ..., \\ & x(n-d_x+1; \mathbf{W})] = f[y(n); x(n); \mathbf{W}] \end{aligned} \tag{4}$$

where $x(n) \in R$; $y(n) \in R$ are respectively the components of the input vector and the output of the network in an instant t; $d_x \geq 1$ and $d_y \geq 1$ ($d_x \geq d_y$) respectively represent the memory vector input and the memory vector output . $\mathbf{W}$ is the matrix of the adjustable weights and f is the unknown nonlinear function which we intend simulate.

The Learning Algorithm is the mathematical rule to adjust the weight matrix $\mathbf{W}$ parameters and the *bias* elements during the training of the network. We selected the Levenberg-Maquardt algorithm that may be interpreted as a merging of two another learning algorithms: the Gauss-Newton method and the Gradient Descendent Algorithm. The Levenberg-Maquardt is just partially dependent of the gradient descendent and less susceptible to phenomenon of *vanishing gradient* [7]. This learning algorithm was showed to be faster than others, and did not require a lot of computing power.

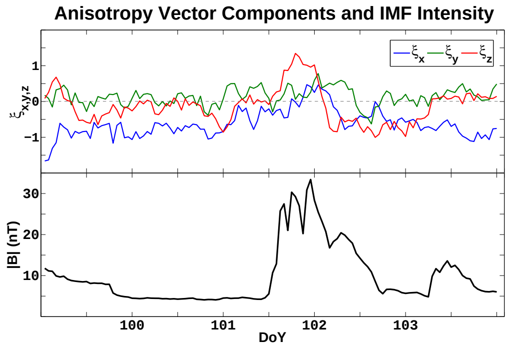

Figure 1: In the first panel from the top the variation of anisotropy vector components of cosmic ray for April 2001: x-component in blue, y-component in green and the z-component in the red. The panel bellow show the variation of IMF intensity in the same period. This period of data was used as reference to the training process.

4 Data and Methodology

We analysed the periods: April, 09 to 14 2001 (April 2001); October,27 2003 to November,01 2003 (October 2003); November, 16 to 24 2003 (November 2003) and December, 12 to 18 2006. We chose periods with intense geomagnetic storms (Dst $\leq$ -100nT), caused by Interplanetary Coronal Mass Ejections (ICMEs) [3]. The anisotropy vector of GCR, calculated by [5], was used as input data of the ANN and the IMF intensity as output data. April 2001 was selected as training set. Figure 1 shows the data of the anisotropy vector components of GCR $\xi_{x,y,z}$, and of the IMF intensity, $|B|$, as function of the day-of-year (DoY), during the period of April 2001. We simulated the periods of October 2003, November 2003 and December 2006, using the ANN trained with April 2001.

We adopted the topology of 5 nodes in the first hidden layer of NARX structure and 20 nodes in the second hidden layer. This topology showed itself a good starting point in previous studies [1]. Also the input vector was feeded with output data from two previous iteractions. We perform simulations with the NARX network operating in two differents modes: the non-cyclic mode and the cyclic mode. The non-cyclic mode is faster and produces more accurate simulations, on the other hand just produces estimatives of one step ahead from the present moment. Differently, the cyclic mode is feeded with the output data produced by the NARX network itself. This enhances the time of forecast but decreases the precision of the results.

As quality indicator of the simulations we used the correlation coeficient between the observed data set and the simulated data set. We also calculated the set of differences between the neural network data and the observed data to establish how distant is the simulation from the reality.

5 Results

The simulations of October 2003 is shown in the Figure 2, where we have , in the first panel, the differences between the long term simulation and measured IMF data, the second panel show the IMF ACE data (black line), and the long term simulation (red line), the next panel show differences of short term simulation and the observed data, and the last panel, we have the short term simulation (green line),

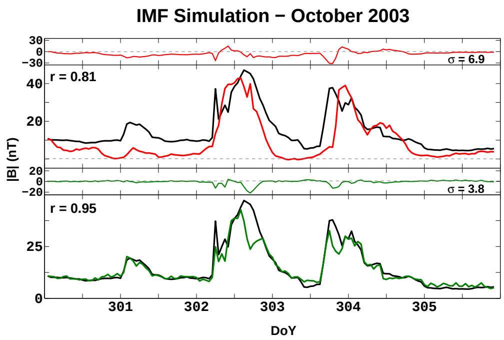

Figure 2: Simulations for the period of October 2003. In the majors panels, the colored line represents the simulation as calculated by the NARX network and the black line represents the actually measured data. The ***r*** value is the correlation coefficient achieved by the simulation. Above each simulation graph it is plotted the differences between the simulation and the observed data where the σ indicates the standard deviation of the differences. In the first two panels from the top, in red, are the difference and the simulation of long term produced by the cyclic network. The lower panels shows, in green, the difference and the simulation of short term produced by the non-cyclic network.

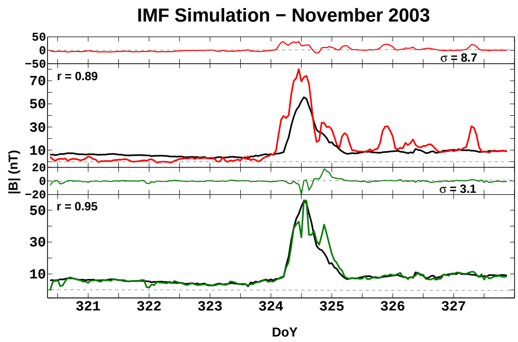

Figure 3: The same as Figure 2, but for the period of November 2003.

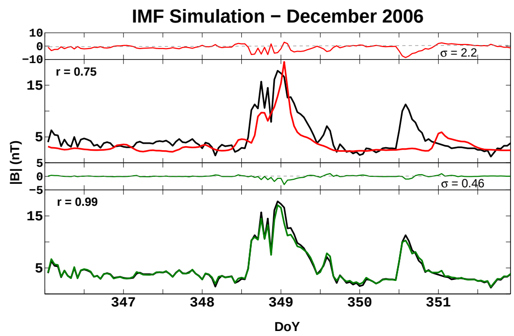

Figure 4: The same as Figure 2, but for the period of December 2006.

and the IMF ACE data (black line). It is possible to see that the majors oscillations are well represented even in the long term simulation. As expected the short term simulation achieved better results but both simulations under-estimate the peak value of the IMF. In the non-cyclic simulation the influence of training set is more evident by the decrease of the simulated IMF after noon of the day 302, that is similar to turbulent region behind the shock exhibited in IMF observed data recorded just before the begining of the day 102 in the Figure 1. It is important to note that during the period of October 2003 occured one of the most intense decreases in the density of cosmic rays recorded by the GMDN: 11%. This peculiarity may justify the low values determined by the neural networks to the peak of the IMF, however it is not clearly perceptible in the general result of the simulations.

The simulations of the period of November 2003 are illustrated in the Figure 3. Despite the higher correlation coefficient than that of October 2003 the long term simulation of the November 2003 was less accurate. The standard deviation of the differences between the neural network simulation and measured data show this clearly. In addition we can note that the values of the differences reached higher values than the long term simulation of October 2003. The peak value was over-estimated and another oscillations appear in the simulation that not corresponding to the behavior of the IMF observed between doy 325 and 326. The short term simulation had a best result. The peak value simulated was very near to the observed by the ACE satellite. However, characteristics of the training set are visible in the short term simulation that presents multiple peaks instead of one single peak expected.

The simulations of December 2006 are showed in Figure 4. The long term simulation presented quality indexes that are contradictory to each other. The correlation coefficient is lower than all other periods, but the standard deviation of the differences is the smallet recorded among the long term simulations. The peak value is delayed and a little higher than the observed. The simulations showed a peak smaller and almost 12 hours after the expected in the doy 350. The quality indexes and the short term simulation are far better than all others here presented. The simulation overlap almost perfectly the IMF variation. It is possible that the similarities between the profile of variation of the IMF in April 2001 and December 2003 have contributed to these impressive results, once the short term simulations of the another periods also showed resemblances with the training data set.

6 Conclusions

It is important to note that we produced simulations in which the IMF is dependent exclusively of the components of the anisotropy vector. The quality indexes obtained in the simulations agrees with the understanding that IMF perform a major role in the modulation of the GCR, producing good simulations. It would be interesting to add in the input vector others parameters that interfere in the GCR modulation, as the solar wind speed.

The fact of the short term simulations sustain some similarities with the training data set suggest that the choice of that training data set is important. It is necessary understand that it is not possible to add more data to the training set, because a huge data set compromises the quality of the simulations and also demands greater computational resources of the hardware. It is necessary determine which data and the amount of them that will produce ANNs capable of the best generalization.

The non-cyclic mode of NARX resulted in simulations with best quality indexes. This may be justified by the inclusion of previous measured data of IMF intensity in the input vector. These data may sistematically correct the estimation of the ANN approaching it to the expect result. Also, NARX networks are always trained in the non-cyclic mode, so this mode of simulation is more similar to training conditions. Despite of its good result is necessary recall that this model of neural network produces estimatives just one step ahead of the present moment.

The preliminary simulations showed in this paper indicate that the NARX model of network deserve more attention of the researchers of cosmic rays field. Despite of assign data of just one single variable in the input vector the quality of the resulting simulations were not negligible.

Acknowledgment: The authors want to thanks Alessandra Abe Pacini (PhD), Oli Luis Dors Jr. (PhD) by the important contributions to this paper. Thanks also to the IP&D (UNIVAP), CAPES, FAPESP and CNPq, for financial support; to the National Science Foundation (NFS) of the USA for a subcontract between Delaware University and FATEC/UFSM from Brazil, for supporting part of the upgrade of the telescope installed in São Martinho da Serra, Brazil.

References

[1] A.G. Oliveira, Study of the relationship between the Magnetic Field and Intensity of Cosmic Rays in the Interplanetary medium via Neural Networks (2013) Master's Dissertation, UNIVAP, São José dos Campos.
[2] C. Jiang and F. Song, Journal of Computers 6 (2011) 1424-1429 doi:10.4304/jcp.6.7.1424-1429.
[3] E. Echer, W.D. Gonzalez, B.T. Tsurutani and A.L.C. Gonzalez, Journal of Geophysical Research 113 (2008), A05221 doi:10.1029/2007JA012744.
[4] H. Simon, Neural Networks: A Comprehensive Foundation, Second Edition. Branche: Pearson Education (1999).
[5] M.R. Silva, A. Dal Lago, W. D. Gonzalez, K. Munakata. C. Kato, J. Bieber, N. J. Schuch, M. L. Duldig, J. E. Humble, H.K. Al Jassar, M.M. Sharma and I. Sabbah, Geophysical Research Letters 38 (2011), doi:10.1029/2011GL048556.
[6] R.G. Tagliaferri, G. Longo, L. Milano, F. Acernese, F. Barone, A. Ciaramella, R. de Rosa, C. Donalek, A. Eleuteri, G. Raiconi, S. Sessa, A. Staino and A. Volpicelli, Neural Networks 16 (2003), 297-319 doi:10.1016/s0893-6080(03)00028-5.
[7] Y. Bengio, P. Simard and P. Fransconi, IEEE Transactions on Neural Networks 5 (1994), 2 157-166 doi:10.1109/72.279181.
[8] Z.L. Kovacs, Artificial Neural Networks: Fundamentals and Applications, São Paulo: Livraria da Física (2006).