Proceedings of the Brazilian Decimetric Array Workshop
São José dos Campos, Brazil - July 28 – August 1, 2008

PRELIMINARY RESULTS WITH NEURAL NETWORK FOR DATA ASSIMILATION TO THE SPACE WEATHER

Haroldo F. de Campos Velho[1] and Fabrício P. Härter[2]

[1]*Laboratory for Computing and Applied Mathematics (LAC)*
National Institute for Space Research (INPE) – São José dos Campos-SP, Brazil
(haroldo@lac.inpe.br)
[2]*National Institute of Meteorology (INMet)*
(fabricio.harter@inmet.gov.br)

ABSTRACT

Data assimilation is an essential step for improving space weather operational forecasting by means of an appropriated combination between observational data and data from a mathematical model. In the present work data assimilation methods based on Kalman filter and artificial neural networks are applied to a three-wave model of auroral radio emissions. A novel data assimilation method is presented, whereby a multilayer perceptron neural network is trained to emulate a Kalman filter for data assimilation by using cross validation. The results obtained render support for the use of neural networks as an assimilation technique for space weather prediction.

INTRODUCTION

Space weather research is the study of the disturbances in the space environment, usually caused by the solar activity and/or interactions of interstellar medium and galactic cosmic rays with the heliosphere. Due to the potential impact of space weather on technological systems on Earth, as well as on human health, space weather forecasting is today an essential task. Nonlinear and chaotic phenomena represented by mathematical models have an intrinsic relationship with the initial conditions (IC). Therefore, from very small discrepancies between two similar ICs, after some time-steps, a disagreement could occur for some systems. In other words, sensitive dependence on the IC could cause the forecasting error to grow exponentially fast with the integration time (Grebogi et al., 1987).

This implies that a better representation for the initial condition will produce a better prediction. The problem for estimating the initial condition is so complex and important for operational prediction system, which it constitutes a science called *Data Assimilation* (Daley, 1993; Kalnay, 2003). Nowadays data assimilation is a research topic in some of the areas of applied physics, such as meteorology, oceanography, and ionospheric weather (for last issue see: Schunk et al., 2004; Scherliess et al., 2004; Hajj et al., 2004).

Many methods have been developed for data assimilation. They have different strategies to combine numerical forecasting and observations, using Kalman filter or variational approaches, for example.

The use of artificial neural network (ANN) for data assimilation is a very recent issue.

The first implementation of the ANNs as a new approach for data assimilation was employed by Nowosad et al. (2000). There are applications in chaotic systems, as well as 1D shallow water equations. An artificial neural network is an arrangement of units characterized by a large number of very simple neuron-like processing units; a large number of weighted connections between the units, where the knowledge of a network is stored; and highly parallel distributed control. Two distinct phases can be devised while using an ANN: the training phase (learning process) and the run phase (activation). The training phase consists of an iterative process for adjusting the weights for the best performance of the network in establishing the mapping of many input/target vector pairs. Once trained, the weights are fixed and new inputs can be presented to the network, which calculates the corresponding outputs based on what had been learned.

In the worked example here, a multilayer Perceptron neural network (MLP-NN) (Haykin, 1994; Nowosad et al.,2000) is trained to emulate a Kalman filter-based data assimilation system. This novel data assimilation strategy is applied to a three-wave model of auroral radio emissions near the electron plasma frequency involving resonant interactions of Langmuir, Alfvén and whistler waves (Chian et al., 1994; Lopes and Chian et al., 2002). Observational evidence of auroral radio emission and nonlinear coupling between Langmuir, Alfvén and whistler waves have been obtained in rocket experiments in the Earth's auroral plasmas (Boehm et al., 1990). These auroral whistler waves may explain the leaked AKR (auroral kilometric radiation), providing the radio signatures of solar-terrestrial connection, and may be used for monitoring space weather from the ground.

Data assimilation is a specialized field of data analysis. The amount of data available today, with the observation system enhancing in quality and quantitative, becomes data analysis a challenge for the science of this new century. Actually, many people are addressing such challenge as *data science*.

NONLINEAR COUPLED WAVE EQUATIONS

A nonlinear analysis of auroral Langmuir, whistler and Alfvén (LAW) events in the planetary magnetosphere was carried out by Lopes and Chian (1996), under the assumption that all three interacting waves are linearly damped. The simplest model for describing the temporal dynamics of resonant nonlinear coupling of three waves can be obtained assuming terms in the wave amplitudes. Moreover, the waves may be assumed monochromatic, with the electric fields $E_\alpha(x,t)$ written in the form: $E_\alpha(x,t) = [A_\alpha(x,t)/2]\exp\{i(k_\alpha x - \omega_\alpha t)\}$, where $\alpha = 1,2,3$ and the time scale of the nonlinear interactions is much longer than the periods of the linear (uncoupled) waves.

In order for three-wave interactions to occur, the wave frequencies ω_α and wave vectors k_α must satisfy the resonant conditions: (i) $\omega_3 \approx \omega_1 - \omega_2$; (ii) $k_3 \approx k_1 - k_2$. Under these circumstances, the nonlinear temporal dynamics of the system can be governed by the following set of three first-order autonomous differential equations written in terms of the complex slowly varying wave amplitude (Meunier et al., 1982):

$$\frac{dA_1}{d\tau} = v_1 A_1 + A_2 A_3 \quad (1a)$$

$$\frac{dA_2}{d\tau} = i\delta A_2 + \bar{v}_2 A_2 - A_1 A_3^* \quad (1b)$$

$$\frac{dA_3}{d\tau} = \bar{v}_3 A_3 - A_1 A_2^* \tag{1b}$$

where the variable $\tau = \chi t$, with χ is a characteristic frequency; $\delta = (\omega_1 - \omega_2 - \omega_3)/\chi$ is the normalized linear frequency mismatch, and $\bar{v}_\alpha = v_\alpha / \chi$ gives the linear wave behaviors on the long time scale. The wave A_1 is assumed linearly unstable ($v_1 > 0$) and the other two waves, A_2 and A_3, are linearly damped ($\bar{v}_2 = \bar{v}_3 \equiv -v < 0$), and it is set $\chi = v_1$ (Meunier et al., 1982; Lopes and Chian, 1996). The system admits both periodic and chaotic waves. Figure 1 shows the strange attractor.

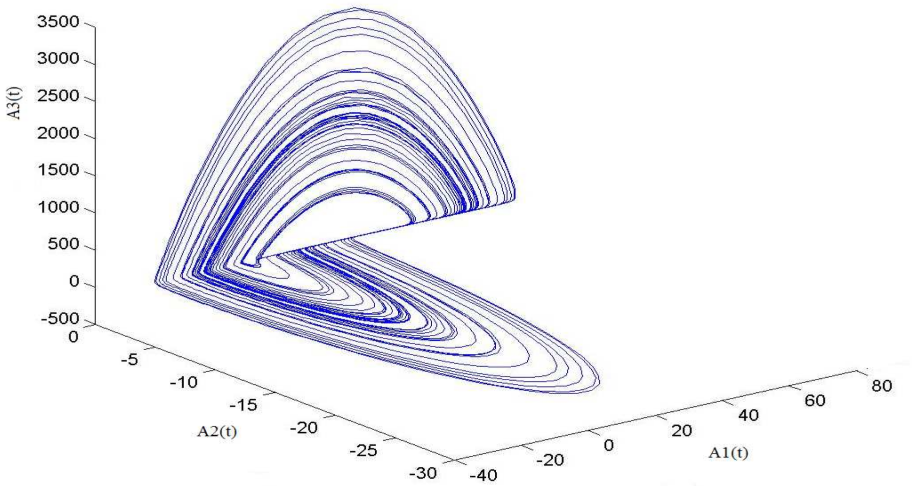

Fig. 1 - The strange attractor for the three-waves: Eqs. (1a)–(1c).

DATA ASSIMILATION EXPERIMENT

Process is illustrated using synthetic experimental data, where synthetic observations are generated by addition of random small level noise on the exact value: $A_\alpha^{Obs}(t_n) = A_\alpha(t_n) + \lambda r_n$, where $\lambda = 10^{-5}$, and r_n is a random value at time t_n. Figure 2 shows observed data inserted after each 5 time-steps on the mathematical model data without any assimilation technique under chaotic regime. The black and blue lines represent the reference model (“model”) and the dynamical system evolution after the data insertion (“corrupted model”), respectively. Clearly, it is noted that the dynamics of the system is lost, even with a small difference in the IC.

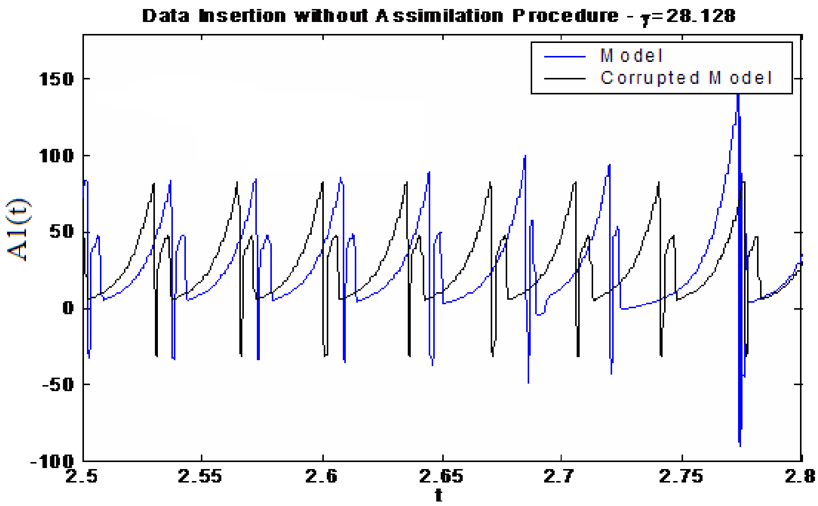

Fig. 2 - Data insertion without assimilation technique with frequency of $5\Delta t$.

Here, the assimilation process will perform by two methods: Kalman filter and neural networks. Denoting the vector $Z^o=[A_1 \quad A_2 \quad A_3]^T$ for the observed data, the extended Kalman filter can be summarized as:

1. Forecasting: $Z^f_{n+1} = F\left(Z^f_n\right) + \mu_n$
2. Evaluation by observation system: $Z^f_{n+1} = H\left(Z^f_n\right) + \nu_n$;
3. Compute the covariance error matrix: $P^f_{n+1} = F_n P^a_n F^{-1}_n + Q_n$;
4. Kalman gain (G): $G_{n+1} = P^f_{n+1} H^T_n \left[R_n + H_n P^f_{n+1} H^T_n\right]^{-1}$;
5. Analysis: $Z^a_{n+1} = Z^f_{n+1} + G_{n+1}[Z^o_{n+1} - H(Z^f_{n+1})]$;
6. Up date the error covariance matrix: $P^a_{n+1} = [I - G_{n+1} H_n] P^f_{n+1}$.

The mathematical model is represented by $F(.)$, μ_n is the stochastic forcing (random modeling noise error) and its covariance matrix is expressed by Q_n. The observation system is modeled by operator (or just a matrix, for linear systems) H_n, and v_n is the noise associated to the observation (covariance matrix denoted by R_n). The typical Gaussian probability density function and zero-mean hypotheses for the noises are adopted. For non-linear dynamical systems, the extended Kalman filter is used (the operator $F(.)$ is expanded into Taylor series, and only linear expansion components are considered). One problem for this approach is to estimate the matrix Q_n. Jazswinski (1970) has proposed an adaptive Kalman filter, where the matrix Q_n is parameterized with these parameters estimated by a secondary Kalman filter. We have applied the Jazswinski's proposal to the data assimilation with good results (Nowosad et al., 2000). The goal here is to design an artificial neural network for emulating a Kalman filter, reducing the computational effort of the assimilation process.

The assimilation procedure using neural network is a non-linear mapping between analysis and data from observation and prediction model:

$$Z^{a}_{n+1} = f_{NN,W}\left[Z^{o}_{n+1}, Z^{f}_{n+1}\right] \tag{2}$$

where $f_{NN,W}$ is the neural network, and W is the matrix of the connection weights. A multilayer perceptron neural network (MLP-NN) was trained with the backpropagation algorithms (Haykin, 1994). The training or learning process is a procedure to identify the best values for the matrix W: the output from the neural network should similar a previous analysis – more details see Härter and Campos Velho (2008a), for a higher dimension system see Härter and Campos Velho (2008b). This target analysis could be the observation, or other acceptable analysis obtained by other method. We follow the second option, and the neural network is designed to emulate the analysis from the Kalman filter.

Numerical example

For simplicity, we assume that all error covariance matrices are diagonal ones. The numerical values for these are given as following:

$$Q_n = 0.1I; \quad R_n = 2I; \quad P^f_0 = \begin{cases} 10Z^f_0 & (i = j) \\ 0 & (i \neq j) \end{cases}. \tag{3}$$

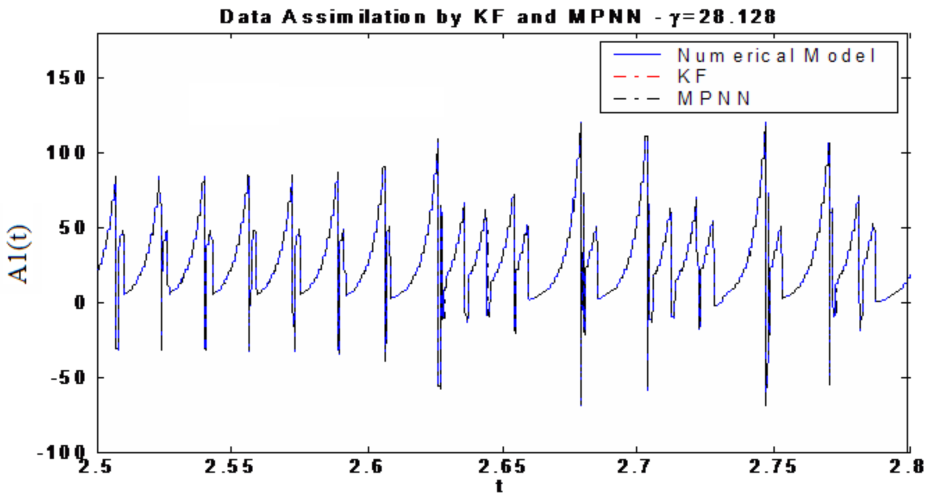

Fig. 3 - Data assimilation using Kalman filter and neural.

The three-wave system is integrated using a fourth-order Runge Kutta scheme, with $\Delta t = 10^{-2}$. After the choice of the best weight set, the 3-wave system is integrated considering data assimilation at each 5 time-steps. Figure 3 depicts the last 10^3 time-steps of a time series of A_α: it is not possible to distinguish the true dynamics, and assimilation obtained with Kalman filter and neural network.

The MLP-NN and KF are effective to carry out the assimilation. Figure 4 shows the mean errors of KF and MLP-NN. Small errors are verified for both schemes.

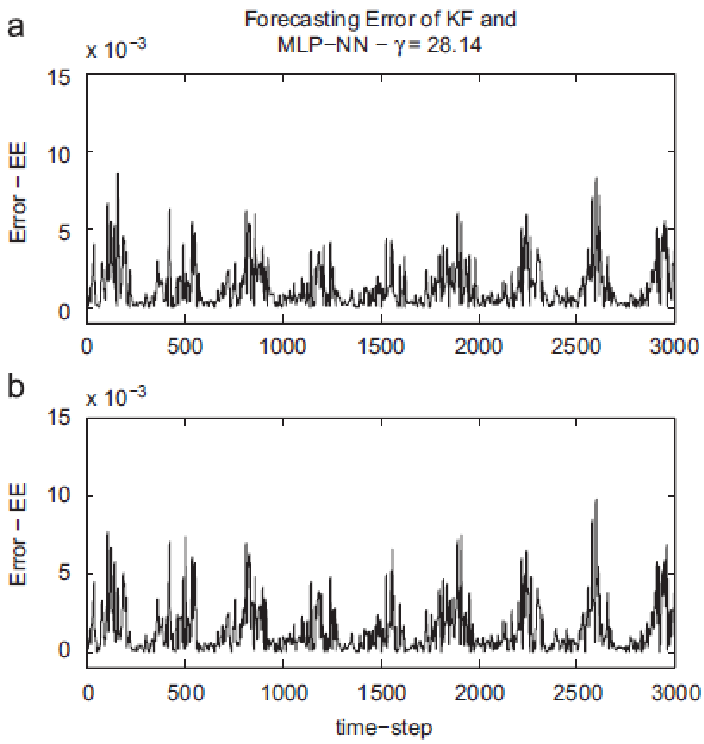

Fig. 4 - Error for data assimilation: (a) Kalman filter, (b) neural network.

REFERENCES

Boehm, M. H., Carlson, C. W., McFadden, J. P., Clemmons, J. H., Mozer, F. S., *Journal of Geophysical Research*, 95, 12157-12171, 1990.

Chian, A. C. -L., Lopes, S. R., Alves, M. V., *Astron. Astrophy.s*, 290, L13-L16, 1994.

Daley, R., *Atmospheric Data Analysis*, Cambridge University Press, 1991.

Garner, T. W., Wolf, R. A., Spiro, R. W., Thomsen, M. F., *Journal of Geophysical Research*, 104, 25145-25152, 1999.

Grebogi, C., Ott, E., Yorke, J. A., *Science*, 238, 585-718, 1987.

Guo, J. S., Shang, S. P., Shi, J. K., Zhang, M. L., Luo, X. G., Zheng, H., *Space Science Reviews*, 107, 229-250, 2003.

Hajj, G. A., Wilson, B. D., Wang, C., Pi, X., Rosen, I. G., *Radio Science*, 1, RS1S05, 2004.

Härter, F. P., Rempel, E. L., Campos Velho, H. F., Chian, A., *Journal of Atmospheric and Solar-Terrestrial Physics*, 70, 1243-1250, 2008.

Härter, F. P., Campos Velho, H. F., *Applied Mathematical Modelling*, 32, 2621-2633, 2008.

Haykin, S., *Neural Networks: A Comprehensive Foundation*. Mcmillan, 1994.

Jazswinski, A. H., *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

Kalnay, E., *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, 2002.

Lopes, S. L., Chian, A.C.-L., *Astronomy and Astrophysics*, 305, 669-676, 1996.

Meunier, C., Bussac, M. N., Laval, G., *Physica*, D4, 236-243, 1982.

Nowosad, A. G., Rios Neto, A., Campos Velho, H. F., *Hybrid Methods in Engineering*, 2(3), 291-310, 2000.
Scherliess, L., Schunk, R. W., Sojka, J. J., Thompson, D. C., *Radio Science*, 1, RS1S04, 2004.
Schunk, R. W., Scherliess, L., Sojka, J. J., Thompson, D. C., Anderson, D. N., Codrescu, M., Minter, C., Fuller-Rowell, T. J., Heelis, R. A., Hairston, M., Howe, B. M., *Radio Science*, 1, RS1S02, 2004.