Identification of nonlinear space weather models of the Van Allen radiation belts using Volterra networks

M. Taylor, I.A. Daglis, A. Anastasiadis and D. Vassiliadis

1Institute for Space Applications and Remote Sensing(ISARS), National Observatory of Athens(NOAA), Metaxa and Vasiliou Pavlou Street, Penteli, Athens 15226, Greece.
2Department of Physics, Hodges Hall, PO Box 615, West Virginia University, Morgantown, WV 26506-6155, USA.

Introduction

Many efforts have been made to develop general dynamical models of the Van Allen radiation belts based on data. Linear time-invariant filter models based on the work of Axford and others [1] to model the electron flux are not adequate for the complex dynamics of the radiation belts. Nonlinear space-variant models [2] have provided model results for the evolution of this complex dynamical system [3,4]. We began with a generalisation of the Wold time series decomposition [5], nonlinearity, and Takens’ Theorem for time-delay embedding combined with the ability of nonlinear autoregressive moving-average (NARMAX) models to incorporate nonlinear time-delay (Volterra) neural networks. We now construct neural networks to model solar and geomagnetic effects on the solar-terrestrial system. In this paper we will show initial results of nonlinear autoregressive modelling of the electron flux based on the NARX process as an example of the network operation and capability.

Theory

We present a spatio-temporal model of the electron flux carried in the radiation belts treated as a function of geostationary orbit parameters and the solar input. The model is based on the Wold time series decomposition [6] with nonlinear autoregressive-moving-average (NARMAX) processes. The network architecture can be converted into equations with known AR and MA coefficients thus providing model equations. In the Results section, we present a spatio-temporal model of the electron flux calculated with a NARX process. The 30 lag steps in the FIR filter makes it impractical to write down the resulting model equations here, but Figures 3b, 4 and 5b show that the simulated model is reproducing fairly well, the physics.

STEP 1: Constructing a taxonomy of nonlinear input-output models

We begin with a generalisation of the Wold time series decomposition [7] having the form:

\[ y_t = \delta + f(z_t) + \varepsilon_t \]

where \( y_t \) is the electron flux time series, \( \delta \) is the prediction function, \( f(z_t) \) is a general function (linear or nonlinear), \( \varepsilon_t \) is the prediction error (zero in the absence of trend), and \( z_t \) is the input. We consider \( f(z_t) \) as an “information matrix” constructed from lagged polynomials \( \Phi, \Theta, \Psi \), lag operators \( L^1, L^2, L^3 \) and differencing indices \( d_1, d_2 \) and have the general form [8,9],

\[ \Psi_t = \Psi_{0,1}(1-L^1)j_{1,t} + \varepsilon_{t+1}(1-L^1)j_{2,t} \]

where \( j_{1,t} \) and \( j_{2,t} \) correspond to a Nonlinear Autoregressive Integrated Moving-Average (NARIMA) process, and \( N_{1,0} \) are the mean and standard deviation of \( J(t) \).

STEP 2: Equivalence between input-output models and Volterra Networks

Equivalence between any NARMAX \( p, q, d_1, d_2, r \) process and its Volterra network representation is guaranteed by a special combination of 2 Theorems:

1. Takens’ Theorem [10]: There is a 1-to-1 mapping between a time series and the underlying dynamical state space

2. Universality Theorem [11]: Nonlinear Multilayer Perceptrons are universal function approximators.

In order to measure the degree of success in reproducing observed values \( J(t) \) from the network model \( \hat{J}(t) \), we used the data-model correlation coefficient:

\[ C = \frac{\sum (J_t - \bar{J})(\hat{J}_t - \bar{\hat{J}})}{\sqrt{\sum (J_t - \bar{J})^2 \sum (\hat{J}_t - \bar{\hat{J}})^2}} \]

where \( \bar{J} \) and \( \bar{\hat{J}} \) are the mean and standard deviation of \( J(t) \).

STEP 3: Construction of Volterra networks

Feedforward MLPs with lagged inputs create short-term memory and incorporate nonlinear dynamics into the network state space. In the case that neural activation functions are linear then they operate as simple infinite impulse-response (FIR) networks [12]. Linear FIR models already exist in the literature [13]. Here we develop nonlinear (noguloar) activation function time-delay FIR networks (Volterra networks) based on the NARMAX \( p, q, \) process, whose general architecture is shown in Figure 2. In this paper we will show initial results of nonlinear autoregressive modelling of the electron flux based on the NARX process as an example of the network operation and capability.

Discussion

All Volterra networks were trained with the Levenberg-Marquardt backpropagation algorithm [14] for 100 epochs and 10 adaptive passes at each step in L-shell altitude (0.1 Earth radii) and over daily-averaged data covering the whole of 1995. These early results suggest that this particular modelling approach is capable of recovering the nonlinear dynamics implicit in the data.

Results

FIG. 3. The best case scenario daily-averaged values of 5-day EAM electron flux at L=1, over the time span 1995-3-109 to 1995-12-10. The data for 1996 used in this study is not available. Note that the DIFFERENT data is available only above L=1.

FIG. 4. (a) The electron flux time series used as the target (1) and 2) final results from the nonlinear autoregressive (NARX) Volterra network. Vertical bars indicate the original data used in this work.

FIG. 5. The angular momentum function obtained directly from the Volterra network ensemble. The main impulsive response region is between 0.17 and 0.2 and has a duration of approximately 12 days.

Acknowledgements

Acknowledgements

References