Monograph:
Wei, H.L., Billings, S.A. and Balikhin, M. (2003) Prediction of the DST Index Using Multi-resolution Wavelet Models. Research Report. ACSE Research Report 849 . Department of Automatic Control and Systems Engineering

Prediction of the *Dst* Index Using Multiresolution Wavelet Models

H.L. Wei, S.A. Billings and M. Balikhin

Research Report No. 849

Department of Automatic Control and Systems Engineering
The University of Sheffield
Mappin Street, Sheffield,
S1 3JD, UK

Oct. 2003

Prediction of the *Dst* Index Using Multiresolution Wavelet Models

H.L. Wei, S.A. Billings and M. Balikhin

Department of Automatic Control and Systems Engineering, University of Sheffield
Mappin Street, Sheffield, S1 3JD, UK

A new identification approach is introduced for predicting the *Dst* index using multi-resolution B-spline wavelet models based on an observational data set consisting of *VBs*, the solar wind parameter as the input, and the *Dst* index as the output. The relationship between the input *VBs* and output *Dst* is initially described using a B-spline wavelet model. This model is then simplified using an OLS-ERR (orthogonal least squares and error reduction ratio) algorithm by selecting the significant model terms to produce a parsimonious wavelet model. Forecasts of the *Dst* index are then computed based on this model.

Keywords: B-spline; wavelets; *Dst* index; multi-resolution analysis; orthogonal least squares

1. Introduction

The sun is a source of a continuous flow of charged particles, ions and electrons called the solar wind. The terrestrial magnetic field shields the Earth from the solar wind, and forms a cavity in the solar wind flow that is called the terrestrial magnetosphere. The magnetopause is a boundary of the cavity, and its position on the day side (sunward side) of the magnetosphere can be determined as the surface where there is a balance between the dynamic pressure of the solar wind outside the magnetosphere and the pressure of the terrestrial magnetic field inside. A complex current system exists in the magnetosphere to support the complex structure of the magnetosphere and the magnetopause. Changes in the solar wind velocity, density or magnetic field lead to changes in the shape of the magnetopause and variations in the magnetospheric current system. In addition if the solar wind magnetic field has a component directed towards the south a reconnection between the terrestrial magnetic field and the solar wind magnetic field is initiated. Such a reconnection results in a very drastic modification to the magnetospheric current system and this phenomenon is referred to as magnetic storms. During a magnetic storm, which can last for hours, the magnetic field on the Earth's surface will change as a result of the variations of the magnetospheric current system. Changes in the magnetic field induce considerable currents in long conductors on the terrestrial surface such as power lines and pipe-lines. Unpredicted currents in power lines can lead to blackouts of huge areas, the Ontario Blackout is just one recent example. Other undesirable effects include increased radiation to crew and passengers on long flights, and effects on communications and radio-wave propagation. Forecasting geomagnetic storms is therefore highly desirable and can aid the prevention of such effects.

The magnetosphere can be considered as a complex input-output system. For such a system, the solar wind plays the role of the input and the geomagnetic indeces can be considered as outputs. The *Dst* index is used to measure the disturbance of the geomagnetic field in the magnetic storm. Numerous studies of correlations between the solar wind parameters and magnetospheric disturbances show that the product of the solar wind velocity *V* and the southward component of the magnetic field, quantified by *Bs*, represents the input that can be considered as the input to the magnetosphere [Gonzalez et al, 1994]. This multiplied input will be denoted by *VBs*.

Several approaches have been proposed to forecast the *Dst* index and other geomagnetic activity indices based on input-output observational data, see for example, the review by Klimas et al. [1996]. The analogue modelling approach which aims to build low-dimensional dynamic analogue systems that can be used to interpret the physical interactions in the magnetosphere and to forecast the geomagnetic indices, has been successfully applied by Goertz et al. [1993] and Klimas et al. [1997,1998,1999]. Existing input-output observational data-based modelling approaches which have been applied to forecast these indices include ARMA models [McPherron, 1999; Vassiliadis, 1999; Vassiliadis, 2000 and the references therein], neural networks [Hernandez et al, 1993; Wu and Lundsted, 1997], NARMAX models [Boaghe et al., 2001], and wavelet-NARMAX models [Billings and Wei, 2003; Wei et al., 2003a].

The aim of the present paper is to introduce a new approach for identifying the input-output relationship for the magnetosphere system based on a limited observational input-output data set. The nonlinear dynamics of the magnetosphere system is initially assumed to be described by a wavelet model with over-complete wavelet bases, an OLS-ERR algorithm [Billings et al., 1988, 1989; Korenberg et al., 1988; Chen et al., 1989] is then applied to determine the model structure and to select the most significant model terms. The resulting parsimonious model, which consists of a relatively small number of wavelet basis functions, is then used to predict the *Dst* index.

2. B-spline wavelets

2.1 Why wavelets ?

Among almost all the functions used for the purpose of approximation, few have had such an impact and spurred so much interest as wavelets. Wavelet decompositions outperform many other approximation schemes and offer a flexible capability for approximating arbitrary functions. Wavelet basis functions have the property of localization in both time and frequency. Due to this inherent property, wavelet approximations provide the foundation for representing arbitrary functions economically using only a small number of basis functions. It can be shown that the intrinsic nonlinear dynamics related to real nonlinear systems can easily be captured by an appropriately fitted wavelet model consisting of a small number of wavelet basis functions.

2.2 Multi-resolution wavelet decompositions

Under some assumptions and considerations, an orthogonal wavelet system can be constructed using *multiresolution analysis* (MRA) [Mallat, 1989; Chui, 1992; Daubechies, 1992]. Assume that the wavelet ψ and associated scaling function ϕ constitute an orthogonal wavelet system, then any function $f \in L^2(R)$ can be expressed as a *multiresolution wavelet decomposition*

$$f(x) = \sum_k a_{j_0,k}\phi_{j_0,k}(x) + \sum_{j \ge j_0}\sum_k d_{j,k}\varphi_{j,k}(x) \qquad (1)$$

where $\varphi_{j,k}(x)=2^{j/2}\varphi(2^j x-k)$ and $\phi_{j,k}(x)=2^{j/2}\phi(2^j x-k)$, $j,k\in Z$ are the scale and translation parameters, and j_0 is an arbitrary integer representing the lowest resolution or scaling level.

Using the concept of *tensor products*, the multiresolution decomposition (9) can be immediately generalised to the muti-dimensional case, where a multiresolution wavelet decomposition can be defined by taking the *tensor product* of the one-dimensional scaling and wavelet functions [Mallat,1989]. Let $f\in L^2(R^d)$, then $f(x)$ can be represented by the *multiresolution wavelet decomposition* as

$$f(x_1,\cdots,x_d)=\sum_k \alpha_{j_0,k}\Phi_{j_0,k}(x_1,\cdots,x_d)+\sum_{j\ge j_0}\sum_k\sum_{l=1}^{2^d-1}\beta_{j,k}^{(l)}\Psi_{j,k}^{(l)}(x_1,\cdots,x_d) \quad (2)$$

where $k=(k_1,k_2,\cdots,k_d)\in Z^d$ and

$$\Phi_{j_0,k}(x_1,\cdots,x_d)=2^{j_0 d/2}\prod_{i=1}^{d}\phi(2^{j_0}x_i-k_i) \quad (3)$$

$$\Psi_{j,k}^{(l)}(x_1,\cdots,x_d)=2^{jd/2}\prod_{i=1}^{d}\eta^{(i)}(2^j x_i-k_i) \quad (4)$$

with $\eta^{(i)}=\phi$ or φ (scalar scaling function or the mother wavelet) but at least one $\eta^{(i)}=\varphi$.

Notice that if j_0 is large enough, the approximation representation (2) can be expressed using only the scaling function ϕ, that is, there exists a sufficiently large integer *J*, such that

$$f(x_1,\cdots,x_d)=\sum_k \alpha_{J,k}\Phi_{J,k}(x_1,\cdots,x_d)=\sum_{k_1,k_2,\cdots,k_d}2^{Jd/2}\prod_{i=1}^{d}\phi(2^J x_i-k_i) \quad (5)$$

2.3 B-spline wavelets

Although many functions can be chosen as scaling and/or wavelet functions, most of these are not suitable for system identification applications, especially in the case of multidimensional and multiresolution expansions. An implementation, which has been tested with very good results, involves B-spline and B-wavelet functions in multiresolution wavelet decompositions [Billings and Coca, 1999; Coca and Billings, 2001; Wei and Billings, 2002].

B-splines are piece-wise polynomial functions with good local properties, and were originally introduced by Chui and Wang [1992] to define a class of semi-orthogonal wavelets for representing a signal using multiroselution decompositions.

The B-spline function of m th order is defined by the following recursive formula:

$$N_m(x)=\frac{x}{m-1}N_{m-1}(x)+\frac{m-x}{m-1}N_{m-1}(x-1),\ m\ge 2 \quad (6)$$

with

$$N_1(x)=\chi_{[0,1)}(x)=\begin{cases}1 & if\ x\in[0,1)\\ 0 & otherwise\end{cases} \quad (7)$$

Setting N_m as the scaling function, that is, $\phi(x) = N_m(x)$, then both the wavelet and the scaling functions can be expressed in terms of the scaling function $N_m(x)$ as follows

$$\phi(x) = \sum_{k=0}^{m} c_k N_m(2x-k) \tag{8}$$

$$\varphi(x) = \sum_{k=0}^{3m-2} d_k N_m(2x-k) \tag{9}$$

with the coefficients given by

$$c_k = \frac{1}{2^{m-1}} \binom{m}{k} \tag{10}$$

$$d_k = \frac{(-1)^k}{2^{m-1}} \sum_{j=0}^{m} \binom{m}{j} N_{2m}(k-j+1),\ k = 0,1,\cdots,3m-2 \tag{11}$$

Clearly, the support of the m th order B-spline wavelet and the associated scaling function are

$$\begin{cases} \text{supp}\ \phi = \text{supp}N_m = [0,m] \\ \text{supp}\ \varphi = [0,2m-1] \end{cases} \tag{12}$$

Both the B-spline wavelets and the associated scaling functions are symmetric or anti-symmetric within the supports. The most commonly used B-spline wavelets are the linear $(m=2)$ and cubic $(m=4)$ cases, both of which can be expressed explicitly. In the present study, the 4^{th} order B-spline and the associated wavelet functions will be used for magnetosphere system identification and the *Dst* index prediction. The 4^{th} order B-spline is defined as

$$\phi(x) = N_4(x) = \frac{1}{6} \sum_{j=0}^{4} \binom{4}{j} (-1)^j (x-j)_+^3 \tag{13}$$

where $x_+^n = x^n$ for $x \geq 0$ and $x_+^n = 0$ for $x < 0$.

3. Nonlinear Input-Output Representations

3.1 The NARMAX model

A wide range of nonlinear systems can be described using the NARMAX model proposed by Leontaritis and Billings [1985]

$$y(t) = f(y(t-1),\cdots,y(t-n_y),u(t-1),\cdots,u(t-n_u),e(t-1),\cdots,e(t-n_e)) + e(t) \tag{14}$$

where f is an unknown nonlinear mapping, $u(t)$ and $y(t)$ are the sampled input and output sequences, n_u and n_y are the maximum input and output lags, respectively. The noise variable $e(t)$ with maximum lag n_e, is unobservable but is assumed to be bounded and uncorrelated with the inputs and the past outputs. The model (14) relates the inputs and outputs and takes into account the combined effects of measurement noise, modelling errors and unmeasured disturbances represented by the noise variable $e(t)$. As a general and natural representation for a wide class of linear and nonlinear systems, model (14) includes, as special cases, several model types, including the Volterra and Wiener representations, time-invariant and time-varying AR(X), NARX and ARMA(X) structures, output-affine and rational models, and the bilinear model [Pearson, 1999].

The NARX model is a special case of the NARMAX model and takes the form

$$y(t) = f(y(t-1),\cdots,y(t-n_y),u(t-1),\cdots,u(t-n_u)) + e(t) \tag{15}$$

3.2 The approximation of multivariate functions in high-dimensions

It can be shown that a multivariate nonlinear function can often be decomposed into a superposition of a number of functional components similar to the well known functional analysis of variance (ANOVA) expansions [Chen, 1993]

$$\begin{aligned} y(t) &= f(x_1(t),x_2(t),\cdots,x_n(t)) \\ &= f_0 + \sum_{i=1}^{n} f_i(x_i(t)) + \sum_{1\le i<j\le n} f_{ij}(x_i(t),x_j(t)) + \sum_{1\le i<j<k\le n} f_{ijk}(x_i,x_j,x_k) + \cdots \\ &\quad + \sum_{1\le i_1<\cdots<i_m\le n} f_{i_1 i_2\cdots i_m}(x_{i_1}(t),x_{i_2}(t),\cdots,x_{i_m}(t)) + \cdots + f_{12\cdots n}(x_1(t),x_2(t),\cdots,x_n(t)) + e(t) \end{aligned} \tag{16}$$

where the first functional component f_0 is a constant to indicate the intrinsic varying trend; f_i, f_{ij}, $\cdots$, are univariate, bivariate, etc., functional components. The univariate functional components $f_i(x_i)$ represent the independent contribution to the system output that arises from the action of the ith variable x_i alone; the bivariate functional components $f_{ij}(x_i,x_j)$ represent the interacting contribution to the system output from the input variables x_i and x_j, etc. Notice that the constant term f_0 can often be omitted since it can be combined into other functional components. Although the ANOVA expansion (16) involves up to 2^n different functional components, experience shows that a truncated representation containing the components up to the bivariate functional terms is often sufficient

$$y(t) = f_0 + \sum_{p=1}^{n} f_p(x_p(t)) + \sum_{p=1}^{n}\sum_{q=p+1}^{n} f_{pq}(x_p(t),x_q(t)) + e(t) \tag{17}$$

In practice, many types of functions, such as kernel functions, splines, polynomials and other basis functions can be chosen to express the functional components in the models (16) and (17). In the present study, however, mutiresolution wavelet decompositions will be chosen to describe the functional components. The functional components $f_{i_1 i_2\cdots i_r}(x_{i_1}(t),x_{i_2}(t),\cdots,x_{i_r}(t))$ $(1\le i_1<i_2<\cdots<i_r\le n)$ will be expressed using the multiresolution wavelet decompositions (2) or (5). For example, the univariate and bivariate functional component $f_p(x_p(t))$ (p=1,2,…,n) and $f_{pq}(x_p(t),x_q(t))$ $(1\le p<q\le n)$ can be expressed using the multiresolution wavelet decompositions (2) and (5) as

$$f_p(x_p(t)) = \sum_k \alpha_{j_0,k}^{(p)} \phi_{j_0,k}(x_p(t)) + \sum_{j \ge j_0} \sum_k \beta_{j,k}^{(p)} \varphi_{j,k}(x_p(t)), \; p = 1,2,\cdots,n, \tag{18}$$

$$f_{pq}(x_p(t), x_q(t)) = \sum_{k_1} \sum_{k_2} \alpha_{J;k_1,k_2} \phi_{J,k_1}(x_p(t)) \phi_{J,k_2}(x_q(t)), \; 1 \le p < q \le n, \tag{19}$$

Define

$$x_k(t) = \begin{cases} y(t-k), & 1 \le k \le n_y \\ u(t-k+n_y), & n_y + 1 \le k \le n = n_y + n_u \end{cases} \tag{20}$$

The quasi-ANOVA expansion (16) can then be viewed as a special form of the NARX model (15) for dynamic input and output systems.

3.3 Model term determination

Assume that M bases (scalar mother wavelet or scaling functions or multiplication of some scalar wavelet and scaling functions) are required to expand the NARX model (16) or (17), and for convenience of representation also assume that the M wavelet bases are ordered according to a single index m, that is, the wavelet dictionary $D = \{p_m\}_{m=1}^{M}$, then (16) or (17) can be expressed as a linear-in-the-parameters form as below:

$$y(t) = \sum_{m=1}^{M} \theta_m p_m(t) + e(t) \tag{21}$$

which can be solved using linear regression techniques. Note that for large n_y and n_u, the model (21) might involve a great number of model terms or regressors. Experience shows that very often many of the model terms are redundant and therefore are insignificant to the system output and can be removed from the model. An efficient orthogonal least squares (OLS) algorithm and an error deduction ratio (ERR) criterion [Billings et al., 1988, 1989; Korenberg et al., 1988; Chen et al., 1989] was developed to determine which terms should be included in the model.

4. System Identification

Figure 1 shows 1464 data points of measurement of the solar wind parameter *VBs* (input) and the *Dst* index (output) with a sample period *T*=1hour from 1st March 1979 to 30th April 1979. This data set was separated into the estimation set consisting of 744 input-output data points measured in March 1979, and the validation set consisting of 711 input-output data points measured in April 1979. The objective was to identify an input-output nonlinear model based on the estimation data set. This model was then used to predict the *Dst* index over the next month.

A variable selection algorithm [Wei et al., 2003b] was applied and six significant variables,{*Dst* (t-1),*Dst* (t-2), *Dst* (t-3),*Dst* (t-4), *VBs* (t-1),*VBs* (t-2)} were selected for this data set. These six variables were used to form a wavelet model.

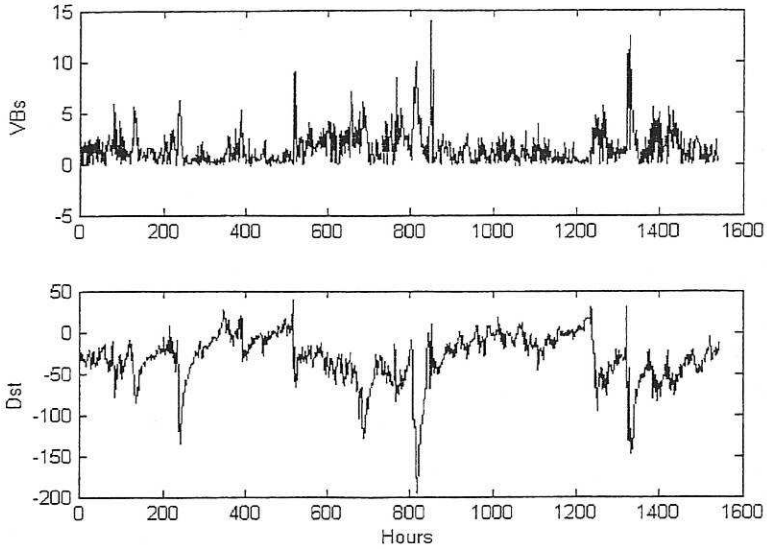

Figure 1 The input (*VBs*) and output(*Dst*) data of the terrestrial magnetospheric dynamic system

4.1 Data pre-processing

The original observational data $\tilde{x}(t) = [\tilde{x}_1(t), \tilde{x}_2(t), \cdots, \tilde{x}_n(t)]^T$ are often normalized into a standard domain, for example the unit hypercube $[0,1]^n$, for the convenience of problem description, where $\tilde{x}_k(t)$ is defined similar to (20), and n=6 for the input-output data set shown in Figure 1. This is especially true when a compactly supported wavelet and/or a scaling function are chosen in the multiresolution decomposition (2) and/or (5). Taking the univariate Haar wavelet (the first-order B-spline wavelet) as an example, it is much easier to select the starting resolution level and the range of the shift parameters if the sample data has been normalized to [0, 1].

Assume that the initial observations $\tilde{x} \in R^n$ fall into the finite hypercube $[a_1, b_1] \times [a_2, b_2] \times \cdots [a_n, b_n]$, then $\tilde{x}(t)$ can be normalized into the unit hypercube $[0,1]^n$ by means of a simple linear transform $x_i(t) = (\tilde{x}_i(t) - a_i)/(b_i - a_i)$, $i = 1,2,\cdots,n$. The modelling can then be performed in the standard hypercube $[0,1]^n$, and the model output can then be recovered to the original system operating domain by taking the inverse transform which converts x back into $\tilde{x}$.

After data nomalisation the truncated multiresolution wavelet models (18) and (19) can be expressed as

$$f_p(x_p(t)) = \sum_{k=-3}^{2^{j_1}-1} \alpha_{j_1,k}^{(p)} \phi_{j_1,k}(x_p(t)) + \sum_{j=j_1}^{J_1} \sum_{k=-2}^{2^j-1} \beta_{j,k}^{(p)} \varphi_{j,k}(x_p(t)), \ p = 1,2,\cdots,n, \tag{22}$$

$$f_{pq}(x_p(t), x_q(t)) = \sum_{k_1=-3}^{2^{J_2}-1} \sum_{k_2=-3}^{2^{J_2}-1} \alpha_{J_2;k_1,k_2} \phi_{J_2,k_1}(x_p(t)) \phi_{J_2,k_2}(x_q(t)), \ 1 \le p < q \le n, \tag{23}$$

4.2 Modelling the *Dst* index

The initial wavelet model was chosen as

$$y(t) = f(y(t-1), y(t-2), y(t-3), y(t-4), u(t-1), u(t-2)) + e(t)$$

$$= \sum_{i=1}^{6} f_i(x_i(t)) + \sum_{i=1}^{5} \sum_{j=i+1}^{6} f_{ij}(x_i(t), x_j(t)) + e(t) \quad (24)$$

where $x_i(t) = y(t-i) = Dst(t-i)$ for i=1,2,3,4 and $x_i(t) = u(t-i+3) = VBs(t-i+3)$ for i=5,6, f_i and f_{ij} are unknown univariate and bivariate functions which can be approximated by one- and two-dimensional wavelet decompositions. In this example, both the input and output data points were initially normalized and the modelling procedure was performed on the standard hypercube $[0,1]^n$, where n=6. The first 744 input-output data points were used for model identification and the remaining 711 data points were used for testing.

Each univariate function component f_i was decomposed using the multiresolution wavelet model (22) with starting resolution scale j_1=0 and the highest resolution scale $J_1 = j_{\max}$=5, and each bivariate functional component f_{ij} was decomposed using the multiresolution wavelet model (23) with the resolution scale J_2=2. The initial model (24) contains 1209 candidate regressors (model terms) after decomposition into the multiresolution wavelet models. An OLS-ERR algorithm [Billings et al., 1988, 1989; Korenberg et al., 1988; Chen et al., 1989] was then used to select the significant model terms. The final was found to be

$$y(t) = \sum_{i=1}^{9} \theta_i B_i(t) + e(t) \quad (25)$$

where $B_i(t)$ (i=1,2, ...,9) are wavelet regressors formed by the 4th-order B-spline function defined by (13), and θ_i (i=1,2,...,9) are the parameters. The terms, parameters and corresponding ERR values are listed in Table 1. Notice again that the variables in the model (24) and (25) were normalised into $[0,1]$, and the model outputs were recovered to the original system operating domain by taking inverse transforms.

4.3 Prediction results

In practice the one-step-ahead (one-hour-ahead) predictions for the *Dst* index are not useful, since it is difficult during a few minutes to collect all the data from both satellite measurements and ground based magnetometers and to feed them into the model (25) to obtain predictions. Forecasting the *Dst* index several months ahead of the real measurements is also seldom required. To be practically useful, the predictions should be made on some time scale which is intermediate between these two extreme cases. Both 6- and 12-hour-ahead predictions based on the model (25) were considered here. The comparisons between the 6- and 12-hour-ahead predictions, the model predicted outputs (the model free-run behaviour) and the measurements are shown in Figure 2, which clearly shows that both the short and long term predictions of the *Dst* index based on the identified model are excellent. The discrepancy between the predicted outputs and the measured values of the *Dst* index might be the result of other inputs which affect the system output but were not included in the current model.

Table 1 The selected model terms, estimated parameters and the corresponding ERR values for the magnetospheric dynamic system

Number	$B_i(t)$	θ_i	$ERR_i \times 100\%$
1	$\phi_{0,-1}(y(t-1))$	1.43819914e+000	9.85860803e+001
2	$\phi_{0,-1}(u(t-1))$	-4.09138701e-001	1.22631148e+000
3	$\phi_{0,0}(y(t-4))$	1.22124826e+000	3.88133408e-002
4	$\phi_{0,-1}(y(t-4))$	-1.85552706e-001	3.93542828e-003
5	$\phi_{0,0}(u(t-2))$	2.05974080e+000	1.25672979e-003
6	$\phi_{0,0}(y(t-2))$	-3.28263544e-001	1.18231886e-004
7	$\phi_{0,0}(y(t-3))$	5.26676127e-001	3.85101910e-004
8	$\phi_{0,-1}(y(t-3))$	-4.31071342e-002	3.98627720e-005
9	$\phi_{0,-1}(y(t-2))$	-2.13486968e-002	5.21787785e-006

Note: 1) $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$ — the 4th-order B-spline functions defined by (13);

2) $y(t)=Dst(t)$ and $u(t)=VBs(t)$.

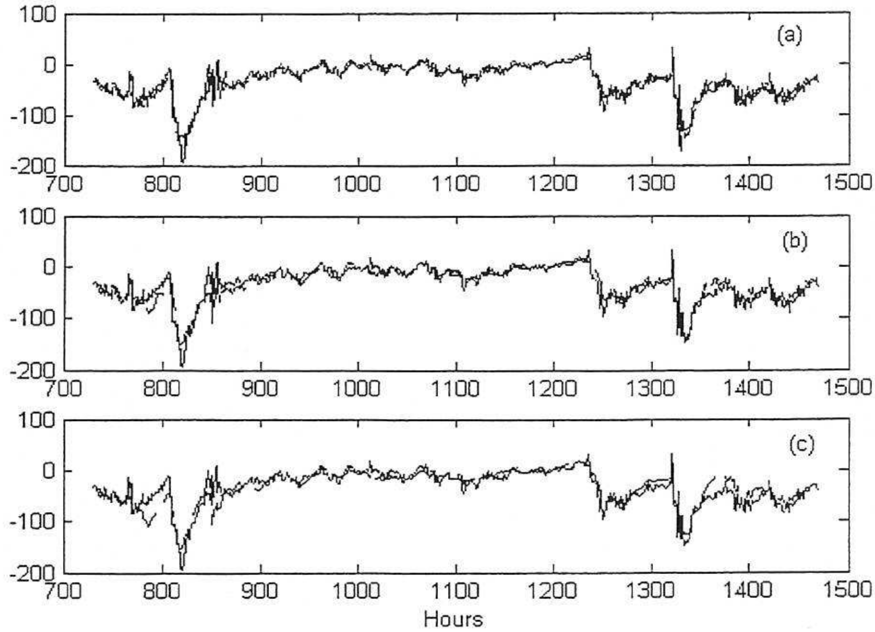

Figure 2 Comparisons of the 6- and 12-hour-ahead predictions, model predicted outputs and the measurements for the solar wind Dst index of April 1979. (a) 6-hour-ahead predictions; (b) 12-hour-ahead predictions; (c) Model predicted outputs. (Solid—measurements; Dashed—6-hour-ahead, 12- hour-ahead, or model predicted outputs).

5. Conclusions

A novel approach has been proposed for predicting the *Dst* index using multiresolution B-spline wavelet decompositions. By expanding a high-order nonlinear model into a sum of additive and interactive low-dimensional submodels, the common problem associated with the *curse-of-dimensionality* in high-order nonlinear system modelling has been greatly alleviated. Wavelets have excellent approximation properties which outperform many other approximation schemes and are well-suited for approximating general nonlinear signals, even those with sharp discontinuities. The intrinsic dynamics in nonlinear systems can be easily captured by a well fitted wavelet model with a small number of wavelet basis functions. This suggests that wavelets are a powerful tool in nonlinear system identification.

The main disadvantage of the new wavelet based modelling approach seems to be that a large number of candidate wavelet basis functions might be involved in the initial wavelet model for a high-dimensional system with several variables (large time lags for the system input and/or system output). Fortunately, this problem can be successfully resolved by employing the well-known and widely used ORS-ERR algorithm, which selects and ranks the significant model terms (regressors).

The prediction results for the *Dst* index in the example provided in section 4 clearly demonstrates the applicability and effectiveness of the new identification approach. This implies that, given some observational data on the input and output of the magnetospheric dynamic system, the modelling approach introduced in the present study can be used to directly identify a wavelet model based on the observational data, and the identified model can then be used to accurately predict the future behaviour of the *Dst* index.

Acknowledgment

The authors gratefully acknowledge that part of this work was supported by EPSRC.

References

Billings, S.A., Korenberg, M. and Chen, S.(1988), Identification of nonlinear output-affine systems using an orthogonal least-squares algorithm, *International Journal of Systems Science*, **19(8)**,1559-1568.

Billings,S.A., Chen,S. and Korenberg,M.J.(1989), Identification of MIMO non-linear systems suing a forward regression orthogonal estimator, *International Journal of Control*, **49(6)**,2157-2189.

Billings, S.A. and Coca, D.(1999), Discrete wavelet models for identification and qualitative analysis of chaotic systems, *International Journal of Bifurcation and Chaos*, **9(7)**, 1263-1284.

Billings, S.A.and Wei, H.L.(2003), The wavelet-NARMAX representation: a hybrid model structure combining the polynomial models and multiresolution wavelet decompositions (submitted for publication).

Boaghe, O.M., Balikhin, M.A., Billings, S.A., and Alleyne, H.(2001), Identification of nonlinear processes in the magnetosphere dynamics and forecasting of *Dst* index, *Journal of Geophysical Research-Space Physics*, **106(A12)**, 30047-30066.

Chen, S., Billings,S.A., and Luo,W.(1989), Orthogonal least squares methods and their application to non-linear system identification, *International Journal of Control*, **50(5)**,1873-1896.

Chen, Z.H. (1993), Fitting multivariate regression functions by interaction spline models, *Journal of the Royal Statistical Society, Series B (Methodological)*, **55(2)**, 473-491.

Chui, C. K.(1992), *An introduction to wavelets.* Boston : Academic Press.

Chui, C. K., and Wang, J. Z.(1992), On compactly supported spline wavelets and a duality principle, *Trans. of the American Mathematical Society*, **330(2)**, 903-915.

Coca, D. and Billings, S.A.(2001), Non-linear system identification using wavelet multiresolution models, *International Journal of Control*, **74(18)**,1718-1736.

Daubechies, I.(1992), *Ten lectures on wavelets*, Society for Industrial and Applied Mathematics, Philaelphia, Pennsylvania.

Goertz, C.K., Shan, L.H., and Smith, R.A.(1993), Prediction of geomagnetic activity, *Journal of Geophysical Research-Space Physics*, **98(A5)**, 7673-7684.

Gonzalez, W.D., Joselyn, J.A., Kamide, Y., Kroehl, H.W., Rostoker, G., Tsurutani, B.T., and Vasiliunas, V.M. (1994), What is a geomagnetic storm?, *Journal of Geophysical Research-Space Physics*, **99(A4)**, 5771-5792.

Hernandez, J.V., Tajima, T., and Horton, W.(1993), Neural net forecasting for geomagnetic activity, *Geophysical Research Letters*, **20(23)**, 2707-2710.

Klimas, A.J., Vassiliadis, D., Baker, D.N., and Roberts D.A.(1996), The organized nonlinear dynamics of the magnetosphere, *Journal of Geophysical Research-Space Physics*, **101(A6)**, 13089-13113.

Klimas, A.J., Vassiliadis, D., Baker, D.N.(1997), Data-derived analogues of the magnetospheric dynamics, *Journal of Geophysical Research-Space Physics*, **102(A12)**, 26993-27009.

Klimas, A.J., Vassiliadis, D., Baker, D.N.(1998),Dst index prediction using data-derived analogues of the magnetospheric dynamics, *Journal of Geophysical Research-Space Physics*,**103 (A9)**, 20435-20447.

Klimas, A.J., Vassiliadis, D., Baker, D.N., and Valdivia J.A.(1999), Data-derived analogues of the solar wind magnetosphere interaction, , *Phys. Chem.Earth (C)*, **24(1-3)**, 37-44.

Korenberg, M., Billings, S.A., Liu, Y. P. and McIlroy P.J.(1988), Orthogonal parameter estimation algorithm for non-linear stochastic systems, *International Journal of Control*, **48(1)**,193-210.

Leontaritis,I.J. and Billings, S.A.(1985), Input-output parametric models for non-linear systems, (part I: deterministic non-linear systems; part II: stochastic non-linear systems), Int. Journal of Control, **41(2)**,303-344.

Mallat,S.G.(1989),A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. On Pattern analysis and machine intelligence*, **11(7)**,674-693.

McPherron, R.L. (1999), Predicting the Ap index from past behaviour and solar wind velocity, *Phys. Chem. Earth (C)*, **24(1-3)**, 45-56.

Pearson, R.K.(1999), *Descrete-time dynamic models*, New York; Oxford: Oxford University Press.

Vassiliadis, D., Klimas, A.J., and Baker, D.N. (2000), The nonlinear dynamics of space weather, *Phys. Chem. Earth (C)*, **26(1)**, 197-207.

Vassiliadis, D., Klimas, A.J., Valdivia, J.A., and Baker, D.N. (1999), Models of Dst geomagnetic activity and of its coupling to solar wind parameters, *Phys. Chem. Earth (C)*, **24(1-3)**, 107-112.

Wei, H.L., and Billings, S.A., 2002, Identification of time-varying systems using multi-resolution wavelet models, *International Journal of Systems Science*, **33**, 1217-1228.

Wei, H.L., Billings, S.A., and Balikhin, M.A.(2003a), Wavelet-based nonparametric models for nonlinear input-output system identification (submitted for publication).

Wei, H.L., Billings, S.A., and Liu J. (2003b), Term and variable selection for nonlinear system identification, (in print, accepted for publication by the International Journal of Control).

Wu, J.G., and Lundsted, H.(1997), Neural network modelling of solar wind magnetosphere interaction, *Journal of Geophysical Research-Space Physics*, **102(A7)**,14457-14466.