Proceedings of the
46th IEEE Conference on Decision and Control
New Orleans, LA, USA, Dec. 12-14, 2007

NARMAX Identification for Space Weather Prediction Using Polynomial Radial Basis Functions

Harish J. Palanthandalam-Madapusi, Biju Edamana, Dennis S. Bernstein,
Ward Manchester and Aaron J. Ridley

Abstract— **Solar storms can damage transformers, electrical networks, and satellites. In this paper, we use system identification methods to construct nonlinear time-series models that are used to predict solar wind conditions with a 27-day prediction horizon. To identify nonlinear time-series models, we use a set of basis functions to represent the nonlinear mapping. For these basis functions, we propose an alternative class of radial basis functions, which have fewer parameters that needs to be tuned by the user. Finally, we compare the predictions obtained using identified models with predictions obtained with existing models.**

I. Introduction

The Sun's atmosphere flows supersonically away from the Sun past the Earth and the other planets in the form of the solar wind. The Sun's magnetic field, which is commonly referred to as the interplanetary magnetic field (IMF), is embedded in the solar wind [4]. Although the solar wind typically flows at about 400 km/s, large expulsions of plasma from the Sun, called coronal mass ejections (CMEs) or solar storms, can attain speeds of over 900 km/s.

High-energy CMEs interact with the Earth's magnetic field causing geomagnetic storms, which entail ionospheric currents and aurora in both the northern and southern polar regions. Most of the time, the aurora and the ionospheric currents are weak and have minimal effect on technology. During a solar storm, however, the aurora and ionospheric currents increase, heating and expanding the upper atmosphere, which causes increased drag on satellites. In addition, large ionospheric current fluctuations can induce currents in power lines that can overwhelm and destroy transformers and electrical networks [6]. Additionally, as discussed in [10], solar storms can adversely affect animals, humans, and aircraft. It is therefore important to be able to predict when and where large ionospheric current fluctuations are likely to occur.

When the Sun's atmosphere is calm, there is little aurora and small currents. However, large ejections of magnetic and plasma energy cause large disturbances in the ionosphere. To predict these large ejections of magnetic and plasma energy, satellites monitor the solar surface and image the photospheric magnetic-field patterns on the Sun.

Existing first-principles models that predict these magnetic events [5, 9] are extremely expensive to run in real time, or must be run at low spatial and temporal resolution. Therefore, researchers must wait until the ejections reach most of the way to the Earth, where additional measurements are made. The solar wind and IMF conditions are measured by the Advanced Composition Explorer (ACE) satellite, which orbits the Lagrangian gravitational null point between the Sun and the Earth. The solar wind measurements are made by ACE 30 to 60 minutes before the solar wind encounters the Earth's magnetic field.

Empirical models are potentially useful for predicting specific quantities at specific locations. Empirical models developed for understanding and predicting magnetic-field fluctuations include Hammerstein-Wiener models [8], neural network models [12], time-series models [2], statistical models [1], and power-law models [11]. The models developed in [1, 2, 8, 12] use current ACE measurements as the inputs to the model and predict magnetic-field fluctuations on the Earth 30 to 90 minutes into the future. On the other hand, the WSA model [11] uses photospheric magnetic-field patterns from the Sun to predict solar wind conditions at ACE 3 to 5 days into the future.

The goal of this paper is to use system identification methods to construct models that can be used to predict solar wind conditions at the first Lagrangian point with a 27-day prediction horizon. The periodicity of the Sun's rotation with respect to the Earth is 27 days, that is, the Sun's rotation and the Earth's revolution around the Sun is such that the same part of the Sun faces the Earth every 27 days. As a result, the current solar wind conditions are highly correlated with the solar wind conditions and image measurements of the solar surface from 27 days earlier. Thus, the output of the models developed in this paper is the solar wind velocity measured by the ACE satellite, while the inputs to the models are photospheric magnetic field measurements and ACE measurements from 27 days earlier. By predicting future solar-wind conditions using current ACE data and photospheric magnetic-field data, our objective is to obtain long-term predictions and thus advance warnings of future disturbances. These warnings can be used to take steps to minimize damage to sensitive infrastructure [6].

Since the solar wind velocity measured by ACE is both a model input and a model output, a natural choice of models is time-series models. Moreover, since the relationship between current solar-wind velocity and past solar-wind conditions and photospheric magnetic-fields is believed to be nonlinear, we identify nonlinear time-series models.

This research was supported in part by the National Science Foundation Information Technology Research initiative, through Grant ATM-0325332.

The authors are with the University of Michigan, Ann Arbor, MI 48109, USA. {hpalanth,bij,dsbaero,chipm,ridley}@umich.edu

1-4244-1498-9/07/$25.00 ©2007 IEEE.

To identify nonlinear time-series models, we use a set of basis functions to represent the nonlinear mapping, and thus recast the problem as a linear least-squares problem. To represent nonlinear mapping of several dimensions, radial basis functions is a convenient choice. However, the user must choose a class of radial basis functions and tune the basis function parameters. Moreover, there are no guidelines for choosing the best class of radial basis functions, while the relative performance often depend on the particular application. Furthermore, although an optimization-based technique was developed in [7] to tune the basis-function parameters, the method is computationally expensive and is not guaranteed to find the global minimizer.

The methodological contribution of this paper is alternative class of radial basis functions for use in NARMAX system identification. In particular, the standard approach to nonlinear time series identification is to use a collection of Gaussian radial basis functions centered at specified grid points. These functions are characterized by a spread parameter, whose optimization entails a nonconvex optimization problem. In the present paper we consider a collection of polynomial basis function centered at each grid point. The coefficients of these basis functions are then set by least squares techniques, thus simplifying the approximation process. In addition, we compare the performance of standard polynomials, Legendre polynomials, and Bernstein polynomials. Bernstein polynomials are of interested for their ability to attain uniform convergence of continuous functions on a compact domain. After comparing the performance of these classes of polynomials on representative examples, we apply this technique to estimation of nonlinear time series models for solar wind prediction.

II. NARMAX Model Identification

Consider the nonlinear time-series model

$$\begin{aligned} y_k &= f(y_{k-\beta}, y_{k-\beta-1}, \ldots, y_{k-\beta-n}, \\ &\quad u_{k-\beta}, u_{k-\beta-1}, \ldots, u_{k-\beta-n}) + v_k, \end{aligned} \tag{II.1}$$

where $u_k \in \mathbb{R}^m$, $y_k \in \mathbb{R}^l$, $f : \mathbb{R}^{n(m+l)} \to \mathbb{R}^l$, $v_k \in \mathbb{R}^l$ is zero-mean white noise, and β is the number of time-steps corresponding to 27 days. Since a typical sample interval is 15 minutes, usually $\beta = 2592$. We consider the problem of identifying the unknown model f when measurements of outputs y_k and inputs u_k are available, and v_k is unknown.

Next, defining $x_k \in \mathbb{R}^{n(m+l)}$ by

$$x_k \triangleq \begin{bmatrix} y_{k-\beta}^{\mathrm{T}} & \cdots & y_{k-\beta-n}^{\mathrm{T}} & u_{k-\beta}^{\mathrm{T}} & \cdots & u_{k-\beta-n}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}},$$

(II.1) can be written as

$$y_k = f(x_k) + v_k. \tag{II.2}$$

Thus the problem of identifying the unknown model in (II.1) is equivalent to multivariable function approximation. Therefore, we first discuss the problem of approximating an unknown function.

III. Static Nonlinear Function Approximation

Consider a static nonlinear system

$$y_k = f(x_k) + v_k, \tag{III.1}$$

where $x_k \in \mathbb{R}^p$, $y_k \in \mathbb{R}^l$, $f : \mathbb{R}^p \to \mathbb{R}^l$, and $v_k \in \mathbb{R}^l$ is zero-mean white noise. We assume that we have measurements of y_k and x_k, while f and v_k are unknown. The problem is to estimate the unknown function f.

Writing f in terms of it's scalar components, we have

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_l(x) \end{bmatrix}, \tag{III.2}$$

where $f_i : \mathbb{R}^p \to \mathbb{R}$. Next, we assume that each component f_i of the function f can be written as a basis function expansion

$$f(x) = \begin{bmatrix} \sum_{i=1}^s b_{1i} g_i(x) \\ \vdots \\ \sum_{i=1}^s b_{li} g_i(x) \end{bmatrix}, \tag{III.3}$$

where $g_j : \mathbb{R}^p \to \mathbb{R}$ for $j = 1, \ldots, s$ are the individual basis functions and b_{ij} are the unknown coefficients of the basis function expansion. Next, defining $g : \mathbb{R}^p \to \mathbb{R}^s$ as

$$g(x) \triangleq \begin{bmatrix} g_1(x) \\ \vdots \\ g_s(x) \end{bmatrix}, \tag{III.4}$$

we can write (III.3) as

$$f(x) = Bg(x), \tag{III.5}$$

where $B \triangleq [b_{ij}] \in \mathbb{R}^{l \times s}$ is the matrix of unknown coefficients. Thus the system (III.1) can be written as

$$y_k = Bg(x_k) + v_k. \tag{III.6}$$

Since (III.6) is linear in the unknown parameters B, we can use least-squares technique to estimate B. However, the basis functions g, which are chosen by the user, affect the accuracy of the resulting approximation.

IV. Radial Basis Functions

A convenient choice of basis functions for g are radial basis functions, which can handle arguments of arbitrary dimensions [3]. A radial basis function is a function of the distance $r_j \triangleq \|x - c_j\|_2$ from a specified center point c_j. For example, Laplacian, logistic, Gaussian, and thin-plate spline radial basis functions have the form

$$g_j(x) = e^{-\alpha_j \|x - c_j\|_2}, \tag{IV.1}$$

$$g_j(x) = \frac{1}{1 + e^{-\alpha_j \|x - c_j\|_2}}, \tag{IV.2}$$

$$g_j(x) = e^{-\alpha_j \|x - c_j\|_2^2}, \tag{IV.3}$$

$$g_j(x) = \|x - c_j\|_2^2 \log(\alpha_j \|x - c_j\|_2), \tag{IV.4}$$

respectively. The parameters α_j and c_j determine the spread and center, respectively, of g_j.

Although the coefficients b_i of the basis function expansion are identified by using standard least-squares, the parameters α_j and c_j must be chosen by the user. In practice, the user can use an optimization process to tune these parameters. However, optimization with respect to α_j and c_j is in general a nonconvex optimization problem.

For example a gradient-based optimization algorithm is used in [7] to optimize the cost function with respect to the parameters α_j and c_j. However, the gradient-based optimization algorithm finds a local minimizer and not the global minimizer. Furthermore, the optimization process is computationally expensive.

V. Polynomial Radial Basis Functions

In this section, we propose an alternative type of radial basis function to avoid the need to optimize with respect to the parameters α_j. First, we rewrite g_j as a polynomial expansion of r_j of the form

$$g_j(x) = \sum_{i=0}^{q} a_{ji} r_j^i, \tag{V.1}$$

where the coefficients a_i are to be determined. Thus

$$\begin{aligned} g(x) &= \begin{bmatrix} \sum_{i=0}^{q} a_{1i} r_1^i \\ \vdots \\ \sum_{i=0}^{q} a_{si} r_s^i \end{bmatrix} \\ &= A \begin{bmatrix} 1 & r_1 & \cdots & r_1^q & 1 & r_2 & \cdots & r_s^q \end{bmatrix}^{\mathrm{T}}, \end{aligned} \tag{V.2}$$

where $A \in \mathbb{R}^{s \times s(q+1)}$ is the matrix of unknown coefficients defined as

$$A \triangleq \begin{bmatrix} a_{10} & \cdots & a_{1q} & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_{20} & \cdots & a_{2q} & \cdots & 0 & \cdots & 0 \\ \vdots & & & & & & \ddots & & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & a_{s0} & \cdots & a_{st} \end{bmatrix}$$

Thus (III.6) becomes

$$y_k = C\phi_k + v_k, \tag{V.3}$$

where $C \triangleq BA \in \mathbb{R}^{l \times s(q+1)}$, and

$$\phi_k \triangleq \begin{bmatrix} 1 & r_1 & \cdots & r_1^q & 1 & r_2 & \cdots & r_s^q \end{bmatrix}^{\mathrm{T}}. \tag{V.4}$$

Let N be the number of data points. Then, define the output matrix $Y \in \mathbb{R}^{l \times N}$ and the regression matrix $\Phi \in s(q+1) \times N$ by

$$Y \triangleq \begin{bmatrix} y_0 & \cdots & y_N \end{bmatrix}, \quad \Phi \triangleq \begin{bmatrix} \phi_0 & \cdots & \phi_N \end{bmatrix}. \tag{V.5}$$

Thus the least-squares estimate $\hat{C}$ of C is given by

$$\hat{C} = Y\Phi^{\dagger}, \tag{V.6}$$

where † represents the Moore-Penrose generalized inverse.

Now, by using least-squares to estimate the unknown matrix C, we eliminate the need to choose α_j.

VI. Bernstein and Legendre Polynomial Radial Basis Functions

In (V.6), as the order q of the polynomials increases, Φ becomes numerically ill-conditioned. Thus we use Legendre polynomials and Bernstein polynomials instead of the standard polynomials in (V.4). Although all three types of polynomials are linear combinations of each other, they differ in properties such as orthogonality and numerical conditioning.

Legendre polynomials $L_i : [-1,\, 1] \to [-1,\, 1]$ are given by

$$\begin{aligned} L_0(r) &= 1, \\ L_1(r) &= r, \\ L_2(r) &= (3r^2 - 1)/2, \\ L_3(r) &= (5r^3 - 3r)/2, \\ L_4(r) &= (35r^4 - 30r^3 + 3)/8, \\ L_5(r) &= (63r^5 - 70r^3 + 15r)/8, \\ L_6(r) &= (231r^6 - 315r^4 + 105r^2 - 5)/16, \\ &\vdots \end{aligned}$$

while Bernstein polynomials $B_{i,q} : [0,\, 1] \to [0,\, 1]$ are given by

$$B_{i,q}(r) = \frac{q!}{(q-i)!\; i!} r^i (1-r)^{q-i}. \tag{VI.1}$$

Furthermore, Legendre polynomials are orthonormal on $[-1,\, 1]$, while Bernstein polynomials are orthonormal on $[0,\, 1]$. Thus, Legendre polynomials and Bernstein polynomials lead to better numerical conditioning of Φ if used instead of the standard polynomials in (V.1) as

$$g_j(r) = \sum_{i=0}^{q} L_i r_j^i$$

or

$$g_j(x) = \sum_{i=0}^{q} B_{i,q} r_j^i.$$

VII. Examples

Consider the static nonlinear system (III.1) with $l = p = 1$ and $f(x_k) = \sqrt{|x_k|}$. To estimate the static nonlinear function using radial basis functions, we use 2000 data points with $u_k \in [-100,\, 100]$. The noise v_k is zero-mean and white with standard deviation 0.1. We compare Gaussian radial basis functions with regular, Legendre, and Bernstein polynomial radial basis functions. For the radial basis functions, 31 c_i are chosen to be distributed uniformly in $[-100,\, 100]$. Figure 1 shows the fit error as a function of the order q of three types of polynomial radial basis functions as compared to Gaussian radial basis functions with $\alpha_i = 0.01$. Although the α_i for gaussian radial basis function are tuned manually to give the least fit error, the Legendre and Bernstein polynomial radial basis functions have smaller fit errors. Furthermore, it is seen that for large orders q, regular polynomial radial basis functions have large fit errors, while Bernstein polynomial

radial basis functions have the least fit errors. The large fit errors for regular polynomial radial basis functions for high orders seem to be due to poor numerical conditioning of Φ. Furthermore, note that a 9^{th} order polynomial radial basis function has 10 unknown coefficients for each c_i, while the gaussian radial basis functions have 1 function for each c_i. However, The unknown coefficients in polynomial radial basis functions are estimated using least-squares, while the α_i in Gaussian radial basis functions are tuned manually.

Next, consider the static nonlinear system (III.1) with $l = 1, p = 2$ and $f(x_k) = |x_{2,k}|\text{atan}(x_{1,k})$. To estimate the static nonlinear function using radial basis functions, we use 961 data points with $u_k \in [-3,\, 3] \times [-3,\, 3]$, and zero-mean, white v_k with standard deviation 0.1. We compare Gaussian radial basis functions with Bernstein polynomial radial basis functions. Figure 2 shows the actual function, while Figure 3 shows the approximation using Gaussian radial basis functions with $\alpha_i = 0.7$. The centers c_i for the Gaussian radial basis functions are chosen to form a 11×11 uniform grid in $[-3,\, 3] \times [-3,\, 3]$. The fit error for the Gaussian RBF approximation is 0.5801, while the fit error for an approximation with 10^{th} order Bernstein polynomial radial basis function at each c_i of the 11×11 uniform grid is 0.0661. Note that, for the Gaussian radial basis function approximation, the user has to tune the parameter α_i. On the other hand, in the Bernstein polynomial RBF approximation there are more unknown coefficients than the Gaussian RBF approximation, but these unknown coefficients are identified using least squares and no manual tuning is required.

Next, we repeat the above example with fewer centers c_i for Bernstein polynomial RBF approximation. We choose c_i for the Bernstein polynomial RBF to be 11 uniformly distributed points along the x_2 axis. Figure 4 shows the approximation of the function using a 100^{th} order Bernstein polynomial radial basis function for each of the 11 c_i. The fit errors for 10^{th} and 100^{th} Bernstein polynomial RBF approximation are 3.2298 and 0.8416, respectively.

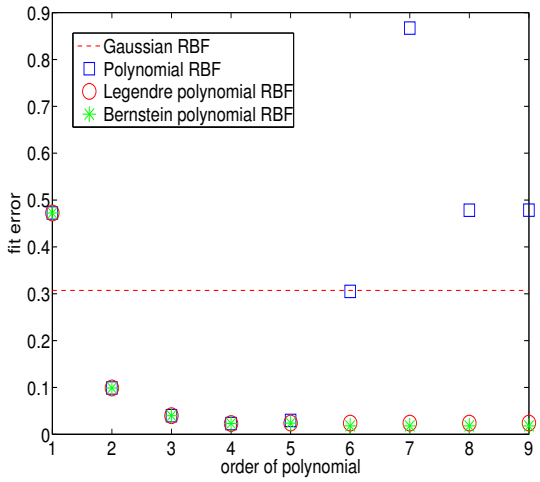

Fig. 1. Polynomial Radial Basis Functions Example. The fit error for three types of polynomial radial basis functions as a function of the order of polynomials q.

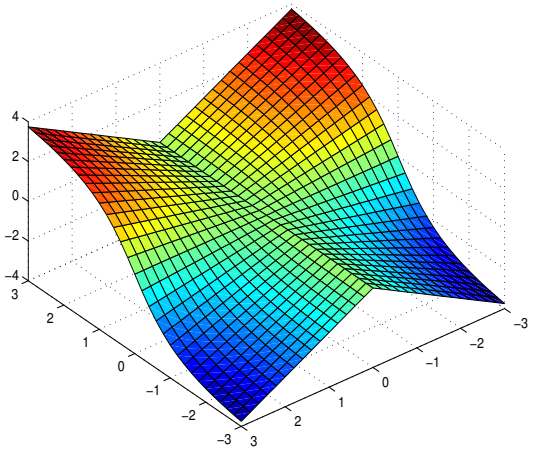

Fig. 2. Plot of the actual function $f(x) = |x_2|\text{atan}(x_1)$.

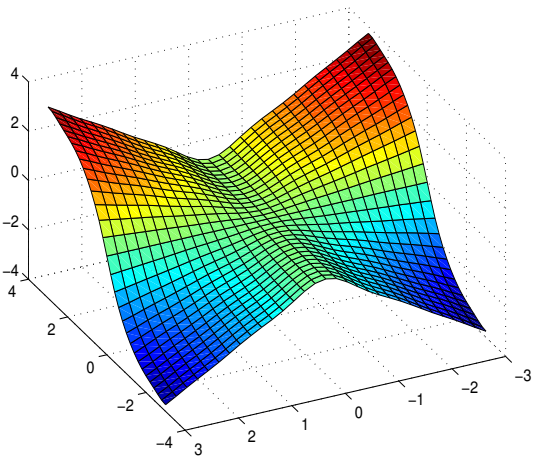

Fig. 3. Estimate of the function $f(x) = |x_2|\text{atan}(x_1)$ using Gaussian radial basis functions with 121 c_i and $\alpha_i = 0.7$.

VIII. Solar-wind Velocity Prediction

In this section, we use Bernstein polynomial radial basis functions to identify NARMAX models of the form (II.1). The output of the model is the solar-wind velocity measured by the ACE satellite.

First, we use ACE measurements of velocity, temperature, and density from 27 days earlier and the WSA model [11] predictions as the inputs to the model. The goal is to investigate whether the use of past ACE data can improve WSA predictions. Since one of the model input is the WSA model output, the prediction horizon is 4-5 days. The dynamic order n for the NARMAX model is chosen to

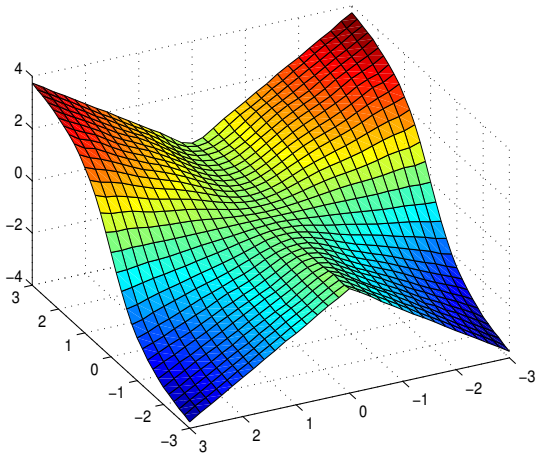

Fig. 4. Estimate of the function $f(x) = |x_2|\text{atan}(x_1)$ using a 100^{th} order Bernstein polynomial radial basis functions for each of the 11 c_i.

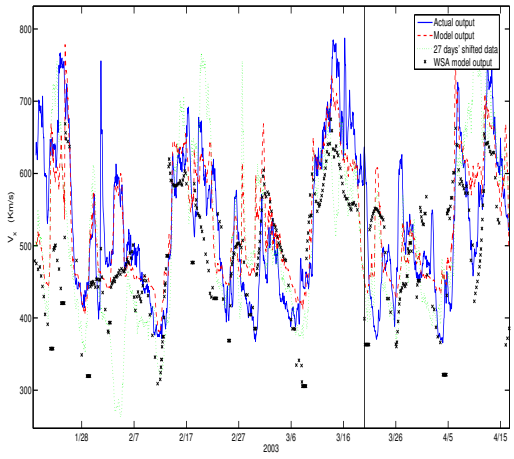

Fig. 5. Comparison of WSA model prediction of solar-wind velocity and prediction of NARMAX model for March-April 2003. Data to the left of the vertical line are used to fit the NARMAX model, while the data to the right of the vertical line is prediction. The skill coefficient of the NARMAX model prediction compared to the WSA model prediction is 2.0073.

Basis Functions	Number of centers	Order of polynomials	Fit error
Gaussain RBF	121	N/A	0.5801
Bernstein Polynomial RBF	121	10	0.0661
Bernstein Polynomial RBF	11	10	3.2298
Bernstein Polynomial RBF	11	100	0.8416

TABLE I

FIT ERRORS FOR APPROXIMATION OF THE FUNCTION $f(x) = |x_2|\text{ATAN}(x_1)$ USING GAUSSIAN RBF AND BERNSTEIN POLYNOMIAL RBF.

be 4 by trial and error. In Figure 5, the data to the left of the black vertical line are used to fit the model, while the data to the right of the vertical line are used for model validation, that is, prediction. For the prediction region, the skill coefficient of the NARMAX model as compared to the WSA model is 2.0073, where the skill coefficient of the NARMAX identified model as compared to the WSA model is defined as

$$\text{Skill coefficient} \stackrel{\triangle}{=} \frac{\text{prediction error using WSA model}}{\text{prediction error using the identified NARMAX model}}.$$

A skill coefficient that is greater than 1 indicates that the prediction using the identified NARMAX model is better than the WSA model prediction.

Next, we use ACE measurements of velocity, temperature, and density from 27 days earlier in conjunction with measurements of the Sun's photospheric magnetic field as the inputs to the model. Since the inputs are measurements made 27 days earlier, the prediction horizon of the identified model is 27 days. We use the identified models to make 27 days advance predictions of solar-wind velocity and compare these 27-day predictions to 4-day predictions of WSA model. Figure 6 shows the 27-day predictions of the NARMAX model and 4-day predictions of the WSA model for March-April 2003. The skill coefficient of the 27-day predictions as compared to the 4-day WSA predictions is 1.7181. Figure 7 shows squared errors for prediction as a function of time. The blue regions indicate time periods during which the NARMAX prediction is better than the WSA prediction, while the red regions indicate periods during which the WSA prediction is better. Despite the fact that NARMAX model predictions are made 27 days in advance versus 4 days in advance for the WSA model, the NARMAX predictions are better than the WSA model predictions 70.236% of the time.

Finally, Figure 8 shows the NARMAX model predictions for April-June 2006.

IX. CONCLUSIONS

Solar storms can damage transformers, electrical networks, and satellites. In this paper, we used system identification methods to construct nonlinear time-series models that predict solar wind conditions with a 27-day prediction horizon. To identify nonlinear time-series models, we used a set of basis functions to represent the nonlinear mapping. Furthermore, we introduced polynomial radial basis functions, which have the advantage that there are fewer parameters that need to be tuned by the user. We then compared three types of polynomial radial basis functions through numerical examples and found that Bernstein polynomial radial basis functions are advantageous. Finally, we used Bernstein polynomial radial basis functions to identify NARMAX models and predict solar wind conditions 27 days into the future and compared them to 4-day predictions obtained from existing

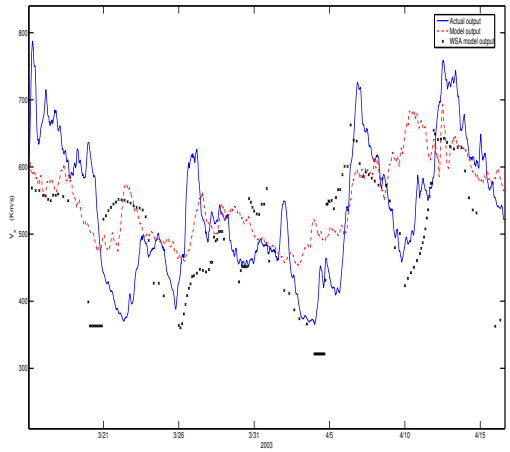

Fig. 6. Comparison of NARMAX model predictions and WSA model prediction of solar-wind velocity and for March-April 2003. Despite the fact that the NARMAX model predictions are 27-days in advance as compared to 4-days advance predictions by WSA model, the skill coefficient of the NARMAX model prediction compared to the WSA model prediction is 1.7181.

models.

REFERENCES

[1] D. Baker. Statistical Analyses in the Study of Solar Wind-Magnetosphere Coupling. *Solar Wind-Magnetosphere Coupling*, pages 17–38, 1986.

[2] D. N. Baker, R. S. Weigel, E. J. Rigler, R. L. McPherron, D. Vassiliadis, C. N. Arge, G. L. Siscoe, and H. E. Spence. Sun-to-magnetosphere modeling: CISM forecast model development using linked empirical methods. *J. of Atmosph. and Solar-Terrestrial Phys.*, 66:1491–1497, 2004.

[3] M. D. Buhmann. *Radial Basis Functions*. Cambridge University Press, July 2003.

[4] D. H. Hathaway. The solar wind. Solar Physics - Science Directorate, http://science.msfc.nasa.gov/ssl/pad/solar/sun_wind.htm.

[5] J.G. Lyon, J.A. Fedder, and C.M. Mobarry. The Lyon-Fedder-Mobarry (LFM) global MHD magnetospheric simulation code. *J. Atmos. Sol-Terr. Phys.*, 66:1333, 2004.

[6] S. F. Odenwald. *The 23rd Cycle*. Columbia University Press, February 2001.

[7] H. Palanthandalam-Madapusi, J.B. Hoagg, and D. S. Bernstein. Basis-function optimization for subspace-based nonlinear identification of systems with measured-input nonlinearities. In *Proc. Amer. Contr. Conf.*, pages 4788–4793, Boston, MA, July 2004.

[8] H. J. Palanthandalam-Madapusi, D. S. Bernstein, and A. J. Ridley. Space weather forecasting: Identifying periodically switching block-structured models to predict magnetic-field fluctuations. *IEEE Control Systems Magazine*.

[9] K.G. Powell, P.L. Roe, T.J. Linde, T.I. Gombosi, and D.L. De Zeeuw. A solution-adaptive upwind scheme for ideal magnetohydrodynamics. *J. Comp. Phys.*, 154:284, 1999.

[10] I. Sandahl. Does space weather really matter on the ground? In Ling-Hsiao Lyu, editor, *Space Weather Study Using Multipoint Techniques*, volume 12 of *Proceedings of COSPAR*, pages 349–357. September 2000.

[11] Y.-M. Wang and N. R. Sheely. Empirical relationship between the magnetic field and the mass and energy flux in the source regions of the solar wind. *Astrophysical Journal*, 1995.

[12] R. S. Weigel, A. J. Klimas, and D. Vassiliadis. Solar wind coupling to and predictability of ground magnetic fields and their time derivatives. *J. of Geophysical Research*, 108(A7).

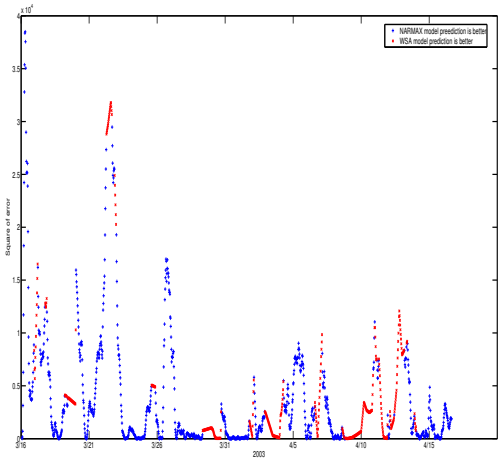

Fig. 7. Squared errors for predictions as a function of time. The blue regions indicate time periods during which the NARMAX prediction is better than the WSA prediction, while the red regions indicate periods during which the WSA prediction is better. Despite the fact that NARMAX model predictions are made 27 days in advance versus 4 days in advance for the WSA model, the NARMAX predictions are better than the WSA model predictions 70.236% of the time.

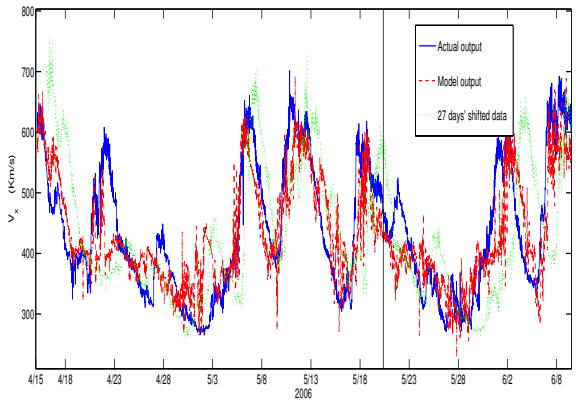

Fig. 8. Solar-wind Velocity Fit and Prediction for 2006. Data to the left of the vertical line are used to identify the model, while data to the right of the vertical line are used for validation, that is, prediction. Since the inputs are measurements that are made 27 days earlier, the prediction horizon is 27 days. The green dashed line indicates the solar wind velocity from 27 days earlier.