ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.17, No.5, May 2006, pp.1042−1050
DOI: 10.1360/jos171042

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

参数可变系统时间序列短期预测方法*

肖 芬, 高协平+

(湘潭大学 信息工程学院,湖南 湘潭 411105)

An Approach for Short-Term Prediction on Time Series from Parameter-Varying Systems

XIAO Fen, GAO Xie-Ping+

(College of Information Engineering, Xiangtan University, Xiangtan 411105, China)

+ Corresponding author: Phn: +86-732-8293249, Fax: +86-732-8293249, E-mail: xpgao@xtu.edu.cn

Xiao F, Gao XP. An approach for short-term prediction on time series from parameter-varying systems. *Journal of Software*, 2006,17(5):1042−1050. http://www.jos.org.cn/1000-9825/17/1042.htm

Abstract: Time series prediction is a very important problem in many applications and the current prediction techniques are nearly all based on the Takens' embedding theorem. Many realistic systems are parameter-varying systems, and the embedding theorems are invalid, predicting the behavior of parameter-varying systems is more difficult. This paper proposes the novel prediction techniques for parameter-varying systems reconstruction, which are based on wavelet neural network (WNN) and multiwavelets neural network (MWNN). These techniques absorb the advantages of high resolution of wavelet and learning of neural networks. The significant improvement is that the error's functions of both networks are convex, and the problem of poor convergence and undesired local minimum can be solved remarkably. Ikeda time series generated by the parameter-varying systems is adopted to check the prediction performance of the proposed models. The numerical experiments show that the three proposed models are feasible, MWNN has the top performance, and WNN could lead the better results than NN in the prediction of the parameter-varying systems.
Key words: wavelet neural network; multiwavelets neural network; time series prediction; parameter-varying dynamical system

摘 要: 时间序列预测是一类非常重要的问题,但基本上局限于参数不可变问题的研究,而对实际问题中经常出现的更重要的参数可变系统的预测,由于构成几乎所有已有预测技术基础的Taken嵌入定理不再成立,所以这方面的研究成果极少.使用一种将(多)小波变换与反向传播神经网络相结合的新型网络结构——(多)小波神经网络,尝试对参数可变时间序列的预测.因为(多)小波神经网络的误差函数是一个凸函数,这在一定程度上可以避免经典神经网络容易陷入局部极小、收敛速度慢等问题.对著名的Ikeda参数可变系统的实验表明,多小波神经网络的预测性能较单小波神经网络要好,而单小波神经网络的性能较BP网要好.因此,该方法不失为时间可变

* Supported by the National Natural Science Foundation of China under Grant No.60375021 (国家自然科学基金); the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry (教育部留学回国人员科研启动基金); the Hu'nan Provincial Natural Science Foundation of China under Grant Nos.03JJY3096, 04JJ20010, 05JJ10011 (湖南省自然科学基金); the Scientific Research Fund of Hunan Provincial Education Department of China under Grant Nos.04A056, 05C092 (湖南省教育厅基金)

Received 2004-08-13; Accepted 2005-07-08

系统预测的一种好的推荐.
关键词: 小波神经网络;多小波神经网络;时间序列预测;参数可变系统
中图法分类号: TP183 文献标识码: A

1 Introduction

Time series is a sequence of observed data and usually ordered in time. Real-life time series can be taken from physical science, business, management, social and behavioral science, economics, etc. Time series prediction, which is based on the idea that the time series carry within them the potential for predicting their future behavior, is a very important problem in many applications. Analyzing observed data produced by a system can give both the good insight into a system and the knowledge about laws underlying the data.

Numerous studies on time series prediction have been undertaken by a lot of researchers. Unfortunately, most of them are related to the construction of structure-invariable system whose parameter values do not change all the time[1−7]. In fact, many realistic systems are naturally structure-variable, which means that the parameter values of these systems are always changing with time. The stock system, the financial expenditure system for one country, the seismic system and the climatic system are such examples since the environments of these systems change rather rapidly before they settle down at some asymptotic states. In this case, the embedding theorems are invalid, which means that the parameter-varying systems reconstruction is more difficult than parameter-invariable systems prediction. Only few related results can be found in Ref.[8], and some theoretical problems still need to be investigated.

Some recent works show that feed-forward neural network (NN), trained with backpropagation and a weight elimination algorithm, outperforms traditional nonlinear statistical approaches in time series prediction[1−5]. In spite of its numerous advantages, like robustness and ability to learn, neural network is hard to determine its structure and it often converges to local minimum. At the same time, wavelet analysis has been developed recently as a powerful analytical tool[9,10]. It has been widely applied to signal analysis and feature extraction due to some excellent properties, for example, in making local analysis. Recently, due to the similarity between wavelet decomposition and one-hidden-layer neural network, the idea of combining both wavelet and neural network has been proposed in various works. The resulting structure is called wavelet neural network (WNN). They show that the training and adaptation efficiency of the wavelet neural network is better than other neural networks[6−8]. Furthermore, using wavelet in the framework of neural network facilitates the theoretical analysis of its asymptotic properties such as universal approximation and consistency, and helps us to determine the structure of neural network. Meanwhile, multiwavelets as a generalization of wavelet appear to have better properties than traditional wavelet, which makes a new way of the theory and application study of wavelet theory. It is possible to construct a basis for multiwavelets with compact support, orthogonal, approximation order greater than 1 and symmetric properties, which are not all simultaneously possible for traditional wavelet basis[11]. And this greatly motivates us to constitute a new kind of wavelet neural networks, called as multiwavelets neural network(MWNN)[12−14], and try to use it for predicting time series from parameter-varying systems. The mean square error function of the network is convex respect to all training parameters, so the problem of poor convergence and undesired local minim are avoided remarkably. Hence one may expect that this wavelet network is a more powerful tool for predicting chaotic series compared to other prediction techniques.

Principal Components Analysis (PCA)[15] is a canonical and widely used method for dimensionality reduction of multivariate data. Applications include the exploratory analysis and visualization of large data sets, as well as the denoising and decorrelation of inputs for algorithms in statistical learning. PCA as a prevalent analysis tool has two advantages: the order of uncorrelated principal components is explicitly given in terms of their variances, and the underlying structure of series can be revealed in using the first few principal components.

The main contribution of the present paper is the use of WNN/MWNN for short-term prediction on time series generated from the parameter-varying systems. And the corresponding prediction techniques are developed. We check the prediction performance of the proposed models by measuring the accuracy of the prediction results for available data sets. Time series generated by Ikeda map is chosen for the purposes of checking prediction ability. In fact, applying the WNN/MWNN for parameter-varying system reconstruction is a new attempt. Fortunately, for the selected prediction tasks, we get promising results by using these three networks. From the simulation results, WNN could lead better results than NN, MWNN appears to yield the best prediction performance and it should to be recommended for further prediction study.

The rest of paper is organized as follows. In Section 2, after a brief introduction of parameter-varying systems, we discuss the short-term predictability for them. Section 3 provides an overview of wavelet analysis, and the detailed description of the wavelet/multiwavelets neural network architecture and prediction algorithm is also presented. Section 4 reports the computer simulations for time series prediction with parameter-varying systems. Finally, in Section 5, a conclusion is provided.

2 Parameter-Varying Dynamical Systems

In many realistic systems, the parameter values are always changing with time, we call this kind of systems as parameter-varying dynamical systems. In some such systems, we can clearly see the phenomenon of period-doubling bifurcations in time if the parameter values vary relatively slow. Such bifurcations are obviously different from the usual period-doubling ones.

The following parameter-varying dynamical system is considered:

$$X(n+1)=g(X(n),\mu(n)) \tag{1}$$

where $g:R^m\times R^p\to R^m$ is a continuous smooth function, and $\mu(n)$ is a p-dimensional parameter vector of the system at time n with its components increasing monotonically with n. Suppose we can only obtain one-dimensional observations from this system, we record the x-component value of each iterations as x_n. In this case, Takens' theorem and its extensions are at least in principle invalid since the given system has no asymptotic sets. Reconstruction of parameter-varying systems will be more difficult because of the lack of the related theory.

Remarkably, in the last two decades, there has been a lot of work trying to apply the chaotic dynamics to many natural systems. In principle, many of these natural systems such as biological, ecological and economical systems belong to parameter-varying systems. Moreover, it is found recently that every parameter-varying system can be transferred into a corresponding higher-dimensional parameter-invariable system by adding some new variables so that the nonlinear prediction algorithms, such as neural networks, wavelet, and chaos analysis, may work well for parameter-varying systems. This result encourages us to apply these prediction algorithms to reconstruct parameter-varying systems.

Consider that there is a time series $x_1,x_2,\ldots,x_n$ from a parameter-varying system, according to the analysis above, there exists a function F such that

$$x_n\approx F(x_{n-d},x_{n-(d-1)},\ldots,x_{n-1}) \tag{2}$$

The problem of predictability is how to find a good estimate of F based on the past history of x_n, on which various techniques of nonlinear deterministic prediction have been developed. In this paper, we will use the prediction algorithm based on the WNN/MWNN.

3 Method

3.1 The wavelet/multiwavelets neural network model

The multiresolution analysis (MRA) introduced by Mallat *et al*.[9,11] gives us a uniform framework to construct the wavelet/multiwavelets basis. Let $\phi:=(\phi_1,\phi_2,\ldots,\phi_r)^T$, $\phi_1,\phi_2,\ldots,\phi_r\in L^2(R)$, $r\geq 1$, satisfy

$$\phi(x)=\sum_{k\in Z}P_k\phi(2x-k),\quad x\in R,$$

for some $r\times r$ matrices P_k, such that the collection of integer translations $\{\phi_l(\cdot-k):1\leq l\leq r,k\in Z\}$ constitutes a Riesz basis of V_0. Such functions $\phi_1,\phi_2,\ldots,\phi_r$ are called (multi)-scaling functions, and they are said to generate a MRA $\ldots\subset V_{-1}\subset V_0\subset V_1\subset\ldots$ of $L^2(R)$, where $V_j:=\{f:f(2^{-j}\cdot)\in V_0\}$, $j\in Z$. Especially, if $\{\phi_l(\cdot-k):1\leq l\leq r,k\in Z\}$ forms an orthonormal basis of V_0, then (V_j) is called an orthogonal MRA.

Define $\phi^i_{M,k}(x)=2^{\frac{M}{2}}\phi_i(2^M x-k),M,k\in Z$,$\Gamma_d=\{\sigma(d)|\sigma(d)=\sigma_1\oplus\sigma_2\oplus\ldots\sigma_d\}$, where σ_i is any integer from 1 to r, and $\oplus$ is connection. For example, $2\oplus 6=26$.

Using tensor product of one-dimensional (multi)-scaling functions, we construct a set of orthogonal basis of subspace V_M as:

$$\Phi_{M,K}(x_1,x_2,\ldots,x_d)=\left\{\prod_{p=1}^{d}\phi^{\sigma_p}_{M,k_p}(x_p)\middle|\sigma(d)\in\Gamma_d,k_p\in Z\right\},K=(k_1,k_2,\ldots,k_d)\tag{3}$$

From the theory of MRA, when M is sufficiently large, one has $L^2(R^d)\approx V_M$, that is, for $\forall f\in L^2(R^d)$, $\forall\varepsilon>0$, $\exists f_M\in V_M$, such as:

$$f_M=\sum_{\sigma(d)\in\Gamma_d}\sum_{\substack{k_p\in Z\\ p=1,2,\cdots,d}}<f(y_1,\ldots,y_d),\prod_{p=1}^{d}\phi^{\sigma_p}_{M,k_p}(y_p)>\prod_{p=1}^{d}\phi^{\sigma_p}_{M,k_p}(x_p)\tag{4}$$

when M is larger enough, there is $\|f_M-f\|\leq\varepsilon$, where $\|\cdot\|$ indicates $L^2(R^d)$ norm,$\langle\cdot,\cdot\rangle$ is the inner product.

In many practical problems, since either the (multi)-scaling functions or the functions have finite support, the sum about k_p in Eq.(4) contains only a finite terms J_M, so we can rewrite the approximation of any $f\in L^2(R^d)$ as:

$$\hat{f}_{M,N}(x_1,x_2,\ldots,x_d)=\sum_{\sigma(d)\in\Gamma_d}\sum_{k_p\in J_M}\hat{C}^{\sigma(d)}_{N,K}\prod_{p=1}^{d}\phi^{\sigma_p}_{M,k_p}(x_p),K=(k_1,k_2,\ldots,k_d)\tag{5}$$

The above equation shows that a function can be approximated by an orthogonal function sets. Since Eq.(5) has a linear-in-parameter structure which can be realized by a neural network, the coefficients $\hat{C}^{\sigma(d)}_{N,K}$ can be regarded as the weights of the neural networks, called wavelet neural network (WNN) for $r=1$, and multiwavelets neural network (MWNN) for $r\geq 2$. The proposed WNN/MWNN are made of three layers: the input layer, the hidden layer and the output layer. In this paper, we just discuss separable wavelet/multiwavelets in the case of $L^2(R^d)$to deal with the high dimensional problem. A schematic diagram of the networks is presented in Fig.1. The following lemma is valuable to calculate the size of the hidden layer.

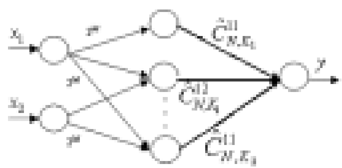

Fig.1 Wavelets/Multiwavelets neural network structure (d=2)

Lemma 1[13]. Suppose the support of $\Phi(x_1,x_2,\ldots,x_d)$ is $[0,u]^d$, and the support of the approximated function F is $[0,1]^d$, then the number of nodes in the hidden layer must be $r^d*(2^M+u-1)^d$, and the set of the threshold value should be $J_M=\{-u+1,\ldots,2^M-1\}$.

3.2 Prediction algorithm

There are many training algorithms for our network. In order to compare the prediction performance with the neural network, the gradient-based techniques for updating the parameters are used for the given network. Because wavelet networks are linear about the training parameters (connecting weights from hidden layer to output layer), the mean square error's function of the network is convex. Therefore gradient-based techniques can avoid the problem of poor convergence and undesired local minimum remarkably.

The basic steps of our algorithm for the reconstruction of parameter-varying systems by WNN/MWNN are as follows:

Step 1. Data pre-processing: The inputs to a system can be anything from external forcing to internal state variables. For the complex time series, the dimension of input information is often too high, and thus an extremely large number of hidden neurons and training examples should be used to train the networks. A preferable strategy is to attempt to reduce the dimension of the input data, and techniques like principal components are often applied to achieve a more compact, more stable model.

Given a training set $T_N = \{(\vec{x}^i, y^i): y^i = x_{i+d}, \vec{x}^i = (x_i, x_{i+1}, ..., x_{i+d-1}), i = 1, ..., N\}$, where x_i is the time series generated by the parameter-varying system. The new "constructed" time series by using PCA on initial inputs are then used as the inputs of the networks.

Step 2. Select a small integer for M (for example M=0), calculate the size of the hidden layer by Lemma 1, the threshold value of each hidden node and the connection weights between the input layer and hidden layer are decided.

Step 3. Initialize $\hat{C}_{N,K}^{\sigma(d)}$ by randomly assigning values between 0 and 1, or by the following equation:

$$\hat{C}_{N,K}^{\sigma(d)} = \frac{1}{N}\sum_{i=1}^{N} y^i \prod_{p=1}^{d} \phi_{M,k_p}^{\sigma_p}(x_p^i).$$

Step 4. Calculate the actual output of the network using the present values of $\hat{C}_{N,K}^{\sigma(d)}$.

Step 5. Adjust $\hat{C}_{N,K}^{\sigma(d)}$ to minimize the following square error:

$$E_{M,N} = \sum_{(x^i,y^i)\in T_N} \left\| o^i - y^i \right\|_R^2 = \frac{1}{N}\sum_{k=d+1}^{d+N} (x_k - \hat{f}_{M,N}(x_{k-1}, x_{k-2}, ..., x_{k-d}))^2$$

where o^i and y^i are the actual and desired outputs when the input vector is x^i. By combining gradient descent techniques and delta rule to adjust the connection weights, we have:

$$\Delta \hat{C}_{N,K}^{\sigma(d)^n} = \lambda \frac{\partial E_{M,N}(f, \hat{f}_{M,N})}{\partial \hat{C}_{N,K}^{\sigma(d)}} + \alpha \Delta \hat{C}_{N,K}^{\sigma(d)^{n-1}} \tag{6}$$

where parameters $\lambda, \alpha \in (0,1)$, λ is called the learning rate, and α is the momentum.

Step 6. Repeat steps 4~5 until the error $E_{M,N}$ is less than a given tolerant value.

4 Simulation Results

In order to investigate the capability of the proposed networks for forecasting the future state of parameter-varying system, the experiments have been performed on the same chaotic time series using three different network models, which are NN, WNN and MWNN with the same network structure. A time series generated by the Ikeda map with one parameter as a variable is chosen. It is described by the following equations[8]:

$$\begin{cases} x_{n+1} = 1 + \mu_n (x_n \cos(t) - y_n \sin(t)) \\ y_{n+1} = \mu_n (x_n \sin(t) + y_n \cos(t)) \\ \mu_{n+1} = \mu_n + 10^{-4}(1 - 0.5\sin(n)) \end{cases} \tag{7}$$

where $t = 0.8 - \dfrac{15}{1 + x_n^2 + y_n^2}$.

Iterate Eq.(7) with initial conditions x_0=0.87,y_0=−0.40, and μ_0=−0.34 until μ_n increases to 0.7 (10,400 iterations). We record the x-component value of each iteration and show the time series x_n,n=0,1,2,…,10,400 in

Fig.2.

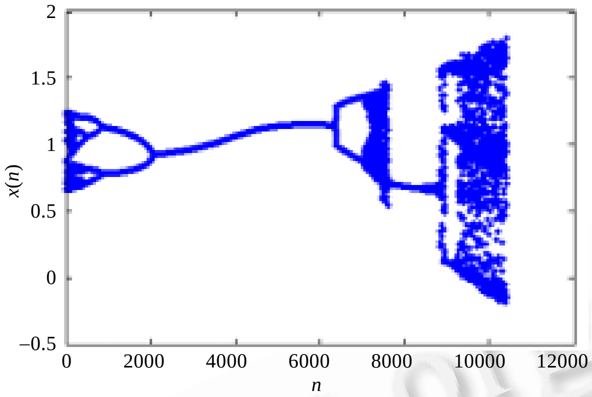

Fig.2 Time series generated by Ikeda map

The time series is divided into three subsets referred to: the chaotic subset, the training and test subsets. The first 400 points are abandoned for their irregularity. The training set consists of 400 data points $x_{400},x_{401},\dots,x_{799}$, we use the training set to fit F by using WNN/WNN and make one-step prediction on the next 200 values.

In our experiments, we reduce the input dimension to 2 by PCA. For the training set, we can calculate its covariance matrix R, the eigenvalues and the corresponding eigenvectors. The eigenvalues take the following values: λ_1=0.1002,λ_2=0.0053,λ_3=0.0018. Obviously, in this case, the first two eigenvectors contains almost all the energy. So the data could be well approximated with a two-dimensional representation.

We choose Daubechies-3 scale function in Ref.[10] as the activation function of the hidden layer in WNN, which is descried as $\phi(x)=\sqrt{2}\sum_{k=0}^{5}H_k\phi(2x-k)$, where H_0=0.3326705529500825, H_1=0.8068915093110924, H_2=0.4598775021184914, H_3=–0.1350110200102546, H_4=–0.0854412738820267, H_5=0.0352262918857095.

The activation function of the hidden layer in MWNN is assumed to be orthonormal multi-scaling functions[11] supported on [0,2] describes in Fig.3:

$$\Phi(x)=\sum_{k=0}^{2}P_k\Phi(2x-k),\quad P_0=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ -\frac{\sqrt{7}}{4} & -\frac{\sqrt{7}}{4}\end{bmatrix}\quad P_1=\begin{bmatrix}1 & 0\\ 0 & \frac{1}{2}\end{bmatrix}\quad P_2=\begin{bmatrix}\frac{1}{2} & -\frac{1}{2}\\ \frac{\sqrt{7}}{4} & -\frac{\sqrt{7}}{4}\end{bmatrix}.$$

Using Lemma 1, we can get that the numbers of the hidden neurons in WNN and MWNN are 25,16 respectively. For the convenience of comparison, we limit the support of the activation function of WNN on [0,4](Fig.4), therefore the number of the hidden neurons in WNN is 16, the same with the MWNN. So the three kinds of networks with the same structure 2-16-1 (i.e., the input vector is composed of 2 components, the hidden layer is composed of 16 neurons, and the output layer is composed of 1 neuron) can be obtained. Although the same network structure, the numbers of the free parameters are different (there are only 17 parameters in the multiwavelets neural network and wavelet neural network while neural network has 65 parameters).

In the experiments, in order to compare the capability of predicting the future state of parameter-varying systems, some error functions are defined as follows:

$$Mse=\frac{1}{l-m}\sum_{i=m}^{l-1}(y_i-\hat{y}_i)^2,\quad Error1=\frac{\sqrt{\sum_{i=l}^{k-1}(y_i-\hat{y}_i)^2}}{k-l},\quad Error2=\frac{1}{k-l}\sum_{i=l}^{k-1}\left|\frac{y_i-\hat{y}_i}{y_i}\right|,$$

$$Error3=\max_{i=l}^{k-1}\left(\left|\frac{y_i-\hat{y}_i}{y_i}\right|\right),\quad Error4=\max_{i=l}^{k-1}(|y_i-\hat{y}_i|).$$

where m=400, l=800, k=1000, y_i is the desired output and $\hat{y}_i$ denotes the actual output.

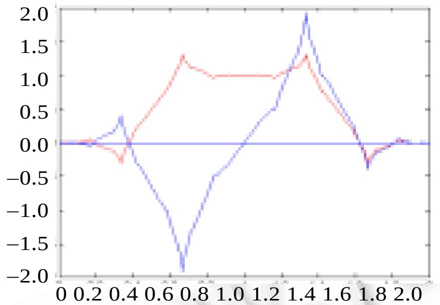

Fig.3 Multi-scaling functions

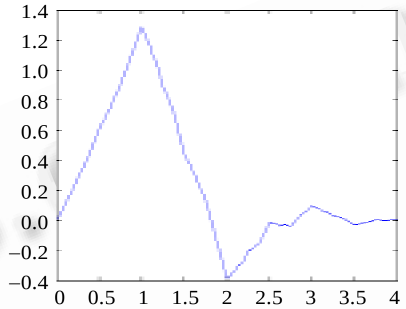

Fig.4 Daubechies scale function

Experiments are conducted on Ikeda time series to testify the prediction performance of the networks. To ensure the relevance of the comparisons, two types of training conditions of networks are recommended in our experiments to give a proper value (namely 20000) for limiting the number of training epochs and impose a desired value (namely 0.0001) as the Mse goal. Interest will first focus on the main features of the training procedures, and, afterwards, emphasis is placed on the prediction quality of the identified models. For a given training times, the achieved Mse is regarded as the training results listed in the first column of Table 1. The first column of Table 2 records the training times with the given tolerant Mse. The prediction quality of the trained networks is evaluated in terms of prediction errors for test subsets, and the mean and minimum errors (Error1, Error 2, Error3, Error 4) over 50 simulations with two different stop training criterions are shown in Table 1 and Table 2.

Table 1 Quality of networks depending on the training epochs 20000

	Network	Mse	Error 1	Error 2	Error 3	Error 4
Mean error	MWNN	7.625e−5	5.002e−4	6.755e−3	1.829e−2	1.679e−2
	WNN	8.865e−5	6.641e−4	9.583e−3	1.917e−2	1.523e−2
	NN	9.099e−5	8.357e−4	1.111e−2	2.515e−2	2.212e−2
Minimum error	MWNN	2.625e−5	1.981e−4	2.583e−3	8.906e−3	6.823e−3
	WNN	5.233e−5	2.322e−4	3.064e−3	9.007e−3	7.617e−3
	NN	2.737e−5	2.211e−4	2.660e−3	9.134e−3	9.731e−3

Table 2 Quality of networks depending on the tolerant Mse 0.0001

	Network	Training times	Error 1	Error 2	Error 3	Error 4
Mean error	MWNN	1.302e+4	5.995e−4	7.820e−3	2.230e−2	2.116e−2
	WNN	1.395e+4	6.611e−4	9.593e−3	1.802e−2	1.468e−2
	NN	1.905+4	9.233e−4	1.205e−2	2.646e−2	2.399e−2
Minimum error	MWNN	1898	2.453e−4	2.263e−3	9.956e−3	7.794e−3
	WNN	795	2.457e−4	3.282e−3	8.628e−3	6.839e−3
	NN	452	3.052e−4	3.545e−3	1.024e−2	9.290e−3

According to the definitions, we think that the measurements of average error Error 1 and Error 2 seem to be the more important indexes to check the network prediction performance than Error 3 and Error 4. From Table 1 and

Table 2, we conclude that these three kinds of networks all can be used to reconstruct the parameter-varying systems, and the performance of WNN are better than NN, while MWNN appear significantly better than WNN although the max relative error (Error 3) and the max absolute error (Error 4) between predicted and actual values trained by MWNN are slightly bigger than those obtained from WNN.

5 Conclusions

In this work, we have presented two prediction models that adopt the ideas from the wavelet analysis and neural networks, then we use them for forecasting of time series generated from the parameter-varying system. The algorithm for predicting the time series has been described in detail. The performance of the proposed networks is tested in predicting the time series generated by the Ikeda map with one parameter as a variable. The experimental results show that the MWNN has better capabilities than the WNN in forecasting the parameter-varying system. And they also show that the WNN performs better than NN on Ikeda time series. In this paper, the dimensionality reduction of the input data by PCA accounts for obtaining more precise and reasonable numerical results, and the size of the hidden layer is reduced.

As pointed out above, the application of the MWNN in conjunction with PCA gives satisfactory results for predicting time series from parameter-varying systems, and this approach would probably be preferable in practical situations.

References:

[1] Principe JC, Rathie A, Kuo JM. Prediction of chaotic time series with neural networks and the issue of dynamic modeling. Int'l Journal of Bifurcation and Chaos, 1992,2(4):989−996.

[2] Aussem A. Dynamical recurrent neural networks towards prediction and modeling of dynamical systems. Neurocomputing, 1999, 28:207−232.

[3] Dash PK, Satpathy HP, C.Liew A, Rahman S. A real-time short-term load forecasting system using functional link network. IEEE Trans. on Power System, 1997,12:675−680.

[4] Pan WM, Shen L. Regularized learning algorithm for neural network-based time series dynamic predictors. Acta Electronica Sinica, 1999,27(11):1−4 (in Chinese with English abstract).

[5] Conway AJ, Macpherson PK, Brown JC. Delayed time series predictions with neural networks. Neurocomputing, 1998,18(1): 81−89.

[6] Geva AB. Scale-Multiscale neural network architecture for time series prediction. IEEE Trans. on Neural Networks, 1998,(9): 1471−1998.

[7] Gao XP, Xiao F. Short-Term prediction of chaotic time series by wavelet networks. In: Proc. of the 5th World Congress on Intelligent Control and Automation. 2004. 1931−1935.

[8] Lyengar SS, Cho EC, V.Phoha V. Foundations of Wavelet Networks and Applications. A Chapman & Hall/CRC Press, 2002.

[9] Mallat S. A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1989,11(7):674−693.

[10] Daubechies I. Ten Lectures on Wavelets. Montpelier: Capital City Press, 1992.

[11] Chui CK, Lian J. A study of orthonormal multiwavelets. Applied Numerical Mathematics, 1996,20(3):273−298.

[12] Jiao LC, Pan J, Fang YW. Multiwavelet neural network and its approximation properties. IEEE Trans. on Neural Networks, 2001, 12(5):1060−1066.

[13] Li XL, Gao XP. Multiwavelet neural network: A novel model. In: Proc. of the IEEE Int'l Conf. on Systems, Man and Cybernetics. 2003. 2629−2632.

[14] Gao XP. A comparative research on wavelet neural networks. In: Wang L, ed. Proc. of the 9th Int'l Conf. on Neural Information Processing. Singapore: Orchid Country Club. 2002. 1699−1703.

[15] Smith LI. A tutorial on principal components analysis. 2002. http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

附中文参考文献:

[4] 潘维民,沈理.基于神经网络的时间序列动态预测器的调整学习算法.电子学报,1999,27(11):1–4.

XIAO Fen was born in 1981. She is a Ph.D. candidate at the Xiangtan University. Her current research areas are wavelet analysis and its application, neural computing, and signal processing.

GAO Xie-Ping was born in 1965. He is a professor and doctoral supervisor at the Xiangtan University and a CCF senior member. His research areas are wavelet analysis and its application, neural computing, and Data mining, etc.