World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:5, No:3, 2011

Time series forecasting using a hybrid RBF neural network and AR model based on binomial smoothing

Fengxia Zheng and Shouming Zhong

Abstract—ANNARIMA that combines both autoregressive integrated moving average (ARIMA) model and artificial neural network (ANN) model is a valuable tool for modeling and forecasting nonlinear time series, yet the over-fitting problem is more likely to occur in neural network models. This paper provides a hybrid methodology that combines both radial basis function (RBF) neural network and auto regression (AR) model based on binomial smoothing (BS) technique which is efficient in data processing, which is called BS-RBFAR. This method is examined by using the data of Canadian Lynx data. Empirical results indicate that the over-fitting problem can be eased using RBF neural network based on binomial smoothing which is called BS-RBF, and the hybrid model—BS-RBFAR can be an effective way to improve forecasting accuracy achieved by BS-RBF used separately.

Keywords—binomial smoothing (BS), hybrid, Canadian Lynx data, forecasting accuracy.

I. Introduction

TIME series analysis and forecasting is an active research area. The accuracy of time series forecasting is fundamental to many decision processes. A wide variety of research, including statistics, neural networks, wavelets, fuzzy set, principal component analysis and chaos, has been undertaken to deal with the different characteristics of time series. Recently, artificial neural networks (ANNs) have been extensively studied and used in time series forecasting [1]-[2].

In essence, a real-world problem is often complex in nature and any single model may not be able to capture different patterns equally well. In such cases, utilization of hybrid model can be of benefit since the well-known M-competition [3] in which combination of forecasts from more than one model often leads to improve forecasting performance. Hybrid methods are robust analysis tools for a large class of complex problems not amenable to traditional classical methods. By combining different methods, complex autocorrelation structures in the data can be modeled more accurately. In addition, the problem of model selection can be eased with little extra effort [1]. In neural network forecasting research, a number of combining schemes have been proposed, including ANNARIMA model [1],[4]-[6], fuzzy ANN model [7], Chaos ANN model [8]-[9], Wavelet Analysis ANN Model [10], and Principal Component Analysis ANN model [11]-[13].

One of the most important and widely used time series models is the ANNARIMA model. A hybrid ARIMA and neural network model was proposed to forecast short-term electricity price [5]. Durdu Ömer Faruk described a hybrid neural network and ARIMA model for water quality prediction [6]. These hybrid models take advantage of the unique strength of ANN and ARIMA in nonlinear and linear modeling. The benefits of such methods appear to be substantial especially when dealing with non-stationary series: non-stationary nonlinear component using neural network model and the residual is stationary linear component which can be modeled by ARIMA model. Although ANNARIMA model is a valuable tool for modeling and forecasting nonlinear time series, the over-fitting problem is more likely to occur in neural network models [14]. Since the parameters of RBF neural networks are only lambda values that need to be trained and binomial smoothing technique which is efficient in data processing [4],[15], In this paper, we propose a hybrid approach to time series forecasting using both RBF neural network and AR model based on binomial smoothing in data processing, which is called BS-RBFAR. In the hybrid technique, a complex problem is decomposed into parts solved by a combination of methods. First, the data sets are smoothed by binomial smoothing technique. Then the smoothed data sets are modeled by RBF neural network. By doing so, the non-stationary nonlinear component is removed. The residuals, which are assumed to be stationary linear, are modeled by AR model. Finally, the predicted components from both RBF neural network and AR model are aggregated to obtain the overall prediction. The result indicate that the over-fitting problem can be eased using RBF neural network based on binomial smoothing (BS-RBF) in data processing, and the hybrid model—BS-RBFAR is believed to greatly improve the prediction performance of the single BS-RBF in forecasting Canadian Lynx.

The remaining sections of this paper are organized as follows. The next section, we review the ANN modeling with binomial smoothing preprocessing approaches and AR modeling approaches to time series forecasting. The hybrid methodology is introduced in Section III. Empirical results from real data sets are reported in Section IV. Section V contains the concluding remarks.

Fengxia Zheng and Shouming Zhong are with the School of Mathematics Science, University Electronic Science and Technology of China, Chengdu 611731, PR China.

Shouming Zhong is with Key Laboratory for NeuroInformation of Ministry of Education, University of Electronic Science and Technology of China, Chengdu 611731, PR China.

Email address: zhengfengxiade@163.com.

II. Time series forecasting models

A. Binomial smoothing preprocessing

With ANNs, the nonlinear model form must be estimated from the data. Therefore, the over-fitting problem is more likely to occur in neural network models [14]. That is, the network fits the training data very well, but has poor generalization ability for data out of the sample. This paper employs smoothing technique in data processing.

Since smoothing by long least-squares polynomial (LSP) sequences leads to transmission zeros, phase reversals, and overshoots that may be objectionable in some applications [15], an alternative method, the binomial smoothing technique because a smoothing sequence is defined by the binomial coefficients can improve these problems. What's more, this binomial filter is most effectively computed. It is applicable in a wide range of situations [15].

Each three-point binomial smoothing $\{y_k\}$ of an n-point data sequence $\{x_k\}$ can be performed as follows:

$$\begin{aligned}
z_1 &= (x_1 + x_2)/2, ..., \\
z_k &= (x_k + x_{k+1})/2, ..., \\
z_{n-1} &= (x_{n-1} + x_n)/2, \\
y_2 &= (z_1 + z_2)/2, ..., \\
y_k &= (z_{k-1} + z_k)/2 \\
&= [(x_{k-1} + x_k)/2 + (x_k + x_{k+1})/2]/2 \\
&= (x_{k-1} + 2x_k + x_{k+1})/2^2, ..., \\
y_{n-1} &= (z_{n-2} + z_{n-1})/2.
\end{aligned}$$

And finally $y_1 = (x_1 + z_1)/2$ and $y_n = (z_{n-1} + x_n)/2$. This keeps the end points fixed and avoids cumulative end effects. This method is much faster than using longer binomial sequences directly or LSP smoothing.

A three-point binomial smoothing sequence is defined by the binomial coefficients as follows:

$$\{1, 2, 1\}/2^2$$

A $n+1$-point binomial smoothing sequence is defined by the binomial coefficients as follows:

$$\left\{ \begin{pmatrix} n \\ 0 \end{pmatrix}, \begin{pmatrix} n \\ 1 \end{pmatrix}, ..., \begin{pmatrix} n \\ k \end{pmatrix}, ..., \begin{pmatrix} n \\ n \end{pmatrix} \right\} /2^n$$

B. The RBF neural network

Neural networks are a class of flexible nonlinear models that can discover patterns adaptively from the data without prior assumption about the underlying relationship in a particular problem. It is often not an easy task to build an multilayer perceptron (MLP) for time series analysis and forecasting because of the large number of factors related to the model section process [14], yet the parameters of RBF neural networks are only lambda values that need to be trained, the function approximation ability and convergence rate are better than standard neural networks. In addition, RBF neural network is a class of local minimum network which can avoid local optimum [4],[16]. This paper employs RBF neural network to capture nonlinear patterns.

The RBF neural network is composed of three layers of nodes. Fig. 1 shows a typical radial basis function (RBF) neural network model used for forecasting purposes. The input nodes are the previous lagged observations, the output provides the forecast for the future value. The hidden nodes contain the radial basis functions (RBF). Suppose we have n-point data sequence $\{y_k\}$, r inputs and one output node are used in neural model, then we have $n - r$ training patterns. The first pattern is composed of $(y_1, y_2, ..., y_r)$ as the inputs and y_{r+1} as the output. The second training pattern contains $(y_2, y_3, ..., y_{r+1})$ for the inputs and y_{r+2} for the output. Finally, the last training pattern is $(y_{n-r}, y_{n-r+1}, ..., y_{n-1})$ for the inputs and the y_n for the output. The relationship between the output y_t and the inputs $(y_{t-1}, y_{t-2}, ..., y_{t-r})$ has the following mathematical representation:

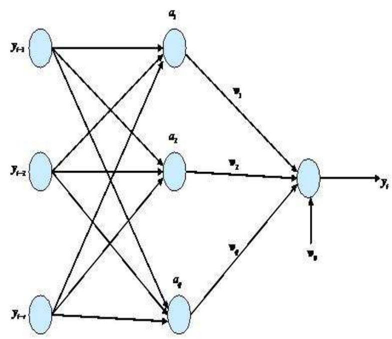

Fig. 1. RBF neural network architecture.

$$\begin{cases} y_t = \sum\limits_{i=1}^{q} w_i a_i + w_0 \\ a_i = \exp(-\sum\limits_{j=1}^{r} (y_{t-j} - \hat{y}_{t-j})^2/\delta_i^2) \end{cases} \tag{1}$$

where y_{t-j} is the ith RBF unit for input variable j, $\hat{y}_{t-j}$ is the centre of jth variable of input pattern t , δ_i is the width of ith RBF unit, w_i is the weight between ith RBF unit and output, w_0 is the biasing term at output node, r is the number of input nodes, q is the number of hidden layer nodes. Note that Eq. (1) indicates a linear transfer function is employed in the output node as desired for forecasting problems. This simple structure of the network model has been shown to be capable of approximating arbitrary function.

C. The AR model

When the nonlinear restriction of the model form is relaxed, the possible number of linear structures that can be used to describe and forecasting a time series is enormous. For a stationary time series, the auto regression (AR) model relates the future value to past and present values in a linear fashion. A process is said to be an autoregressive process of order p if

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \tag{2}$$

where $\phi_i (i = 1, 2, ..., p)$ are the autoregressive parameters, μ is the mean of the series and ε_t is a random process with a mean of zero and a constant variance of δ^2. The Box-Jenkins methodology includes three iterative steps of model identification, parameter estimation and diagnostic checking. The order of the process p can be estimated using various methods such as autocorrelation plots, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). AIC is generally used if one needs to automate the autoregressive order identification process [17]. The lower the AIC is the better the model is. Once the order is identified, estimation of the model parameters is straightforward. The parameters can be estimated using methods such as least squares or Yule-Walker equations. The last step of model building is the diagnostic checking of model adequacy. Several diagnostic statistics and plots of the residuals can be used to examine the goodness of fit of the tentatively entertained model to the historical data.

III. The hybrid methodology

Since real-world time series are rarely pure linear or nonlinear. Hence, by combining ANN model with ARIMA model, complex autocorrelation structures in the data can be modeled more accurately. Additionally, because of the possible unstable or changing patterns in the data, using the hybrid method can reduce the model uncertainty which typically occurred in statistical inference and time series forecasting [18].

It may be reasonable to consider a time series to be composed of a nonlinear component and a linear autocorrelation structure. The hybrid model Y_t can then be represented as follows:

$$Y_t = N_t + L_t \tag{3}$$

where N_t denotes the nonlinear component, L_t denotes the linear component. These two components have to be estimated from the data. First, an n-point data sequence $\{x_k\}$ are smoothed by binomial smoothing technique. Then the smoothed data sequence $\{y_k\}$ are modeled by RBF neural network. By doing so,the non-stationary nonlinear component is removed. Let e_t denote the residual at time t from the nonlinear model, then

$$e_t = x_t - \hat{N}_t \tag{4}$$

where $\hat{N}_t$ is the forecast value for time t from the estimated relationship (1).

Since the RBF model cannot capture the linear structure of the data, the residuals of nonlinear model will contain information about the linearity. By modeling residuals using AR model, linear relationships can be discovered. With the AR model for the residuals will be

$$e_t = \mu + \phi_1 e_{t-1} + \phi_2 e_{t-2} + ... + \phi_p e_{t-p} + \varepsilon_t \tag{5}$$

where ε_t is the random error. Residuals are important in diagnosis of the sufficiency of linear models. A linear model is not sufficient if there are still linear correlation structures left in the residuals. Denote the forecast from (5) as $\hat{L}_t$, the combined forecast will be

$$\hat{Y}_t = \hat{N}_t + \hat{L}_t \tag{6}$$

The results from the AR model can be used as predictions of the error terms for the RBF model. The hybrid model exploits the unique feature and strength of ANN model as well as ARIMA model in determining different patterns. Thus, it could be advantageous to model linear and nonlinear patterns separately by using different models and then combine the forecasts to improve the overall modeling and forecasting performance.

IV. Experiments

A. Data preprocessing and forecasting evaluation methods

A well-known data set—the Canadian lynx data is used in this study to demonstrate the effectiveness of the neural network based on binomial smoothing and the effectiveness of the hybrid method. The lynx series contains the number of lynx trapped per year in the Mackenzie River district of Northern Canada. The data set has 114 observations, corresponding to the period of 1821-1934. Training set has 100 observations, corresponding to the period of 1821-1920. Test set has 14 observations, corresponding to the period of 1921-1934. It has been extensively analyzed in the time series literature with a focus on the nonlinear modeling. Following other studies [1],[19], before the forecasting study, the logarithms (to the base 10) of the data are used in the analysis. Then three-point binomial smoothing was used in this paper. The data and the smoothed data series are plotted in Fig. 2.

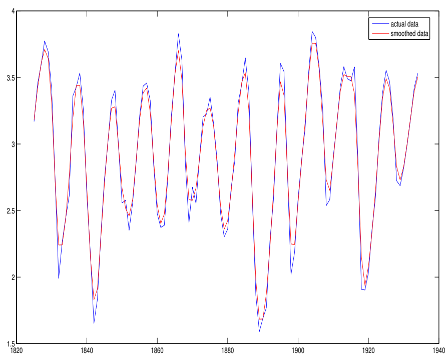

Fig. 2. The actual and the smoothed Canadian Lynx data series.

There are several methods of measuring a time series model's accuracy. We examine the forecasting accuracy by calculating two different evaluation statistics: the root mean square error (RMSE) and the mean absolute percentage error (MAPE) in this paper. They are expressed in the following:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(A_t - F_t)^2}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}|\frac{A_t - F_t}{A_t}|100\%$$

where A_t and F_t are the ith actual and predicted values respectively, and N is the total number of predictions.

World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:5, No:3, 2011

B. Results

Short-term predictions are modeled and different mathematical models are presented to forecast Canadian Lynx data. In this study, neural network models are built using the RBF neural network. Only the one-step-ahead forecasting is considered. Four inputs are used in neural model. Fig. 3-6 gives the actual values and the forecast values with individual

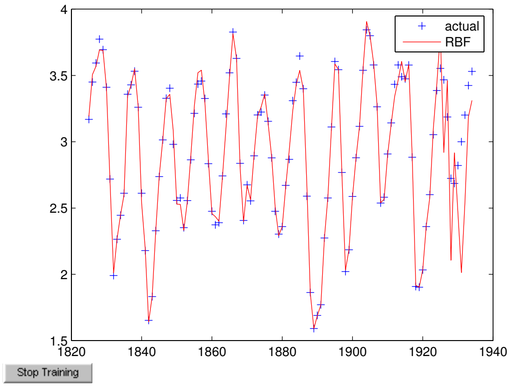

Fig. 3. RBF neural network prediction of Canadian Lynx data.

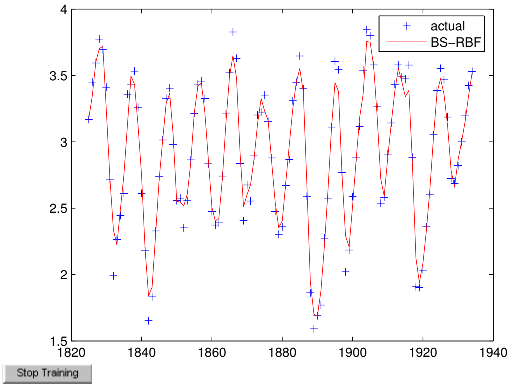

Fig. 4. BS-RBF prediction of Canadian Lynx data.

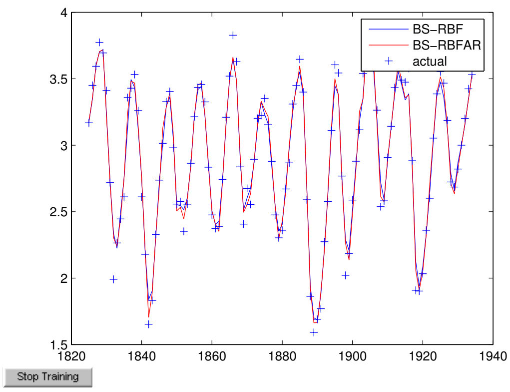

Fig. 5. BS-RBF and BS-RBFAR prediction of Canadian Lynx data.

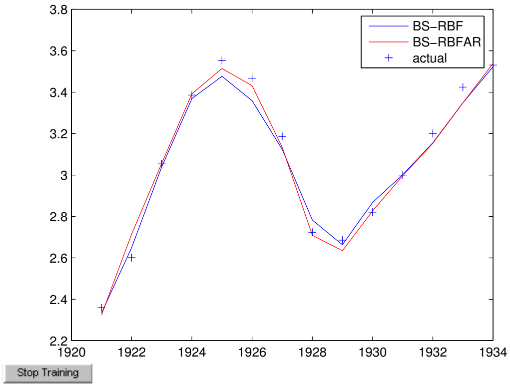

Fig. 6. BS-RBF and BS-RBFAR prediction of Canadian Lynx data(1921-1934).

TABLE I
THE FORECASTING RESULTS FOR THE CANADIAN LYNX DATA

model	RMSE	MAPE
BS-RBF	0.0530	0.0142
BS-RBFAR	0.0469	0.0118

models of RBF neural network and RBF neural network based on binomial smoothing (BS-RBF) as well as the hybrid RBF neural network and AR model based on binomial smoothing (BS-RBFAR). Table 1 gives the forecasting results for the Canadian lynx data.

Results show that while applying RBF neural network based on binomial smoothing can ease the over-fitting problem and can be used for forecasting time series. But the RBF neural network based on binomial smoothing can not capture all of the patterns in the data. The results of the hybrid model show that by combining two models together, the overall forecasting errors can be significantly reduced the MAPE and RMSE, the percentage improvements of the hybrid RBF neural network and AR model based on binomial smoothing (BS-RBFAR) over the RBF neural network based on binomial smoothing (BS-RBF).

V. CONCLUSION

Theoretical as well empirical evidences in the literature suggest that the hybrid model will have lower generalization variance or error and can reduce the model uncertainty which typically occurred in statistical inference and time series forecasting. Furthermore, by binomial smoothing preprocessing to the data, the over-fitting problem that is more strongly related to neural network models can be eased.

ACKNOWLEDGMENT

This research was supported by National Basic Research Program of China (2010CB732501) and UESTC Research Fund For Youth Science & Technology (L08010701JX0809).

REFERENCES

[1] G. Peter Zhang, Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing*, vol. 50, pp. 159 - 175, 2003.

[2] A. Chaouachi, R.M. Kamel, R. Ichikawa, H. Hayashi, and K. Nagasaka, Neural Network Ensemble-based Solar Power Generation Short-Term Forecasting, *International Journal of Information and Mathematical Sciences*, vol. 5, pp. 332-337, 2009.
[3] S. Makridakis, A. Anderson, R. Carbone, R. Fildes, M. Hibdon, R. Lewandowski, J. Newton, E. Parzen, R. Winkler , The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *J. Forecasting*, vol. 1, pp. 111-153, 1982.
[4] D.K. Wedding II, K.J. Cios, Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model. *Neurocomputing*, vol. 10, pp. 149-168, 1996.
[5] Areekul, P., Senjyu, T., Toyama, H., Yona, A., A Hybrid ARIMA and Neural Network Model for Short-Term Price Forecasting in Deregulated Market, *IEEE Transactions on Power Systems* , vol. 25, pp. 524-530, 2010.
[6] Durdu Ömer Faruk, A hybrid neural network and ARIMA model for water quality time series prediction, *Engineering Applications of Artificial Intelligence*, vol. 23, pp. 586-594, 2010.
[7] Mehdi Khashei, Seyed Reza Hejazi, Mehdi Bijari, A new hybrid artificial neural networks and fuzzy regression model for time series forecasting, *Fuzzy Sets and Systems*, vol. 159, pp. 769-786, 2008.
[8] C.J. Dong and Z.Y. Liu, Multi-layer neural network involving chaos neurons and its application to traffic-flow prediction *Journal of System Simualtion*, vol. 19, pp. 4450-4453, 2007.
[9] Y.L. Gu, C.F. Shao, and X.Y. Tan, Dynamic traffic volume forecast method based on chaotic neural network, *Traffic & Computer*, vol. 25, pp. 28-30, 2007.
[10] Ying Chen, Luh, P.B., Che Guan, Yige Zhao, Michel, L.D., Coolbeth, M.A., Friedland, P.B., Rourke, S.J., Short-Term Load Forecasting: Similar Day-Based Wavelet Neural Networks, *IEEE Transactions on Power Systems*, vol. 25, pp. 322-330, 2010.
[11] X.L. ZHANG, and G.G. HE, Forecasting approach for short-term traffic flow based on principal component analysis and combined neural network, *Systems Engineering - Theory & Practice*, vol. 27, pp. 167-171, 2007.
[12] G.J. Song. C. Hu, K.Q. Xie, and R. Peng, Process neural network modeling for real time short-term traffic flow prediction, *Journal of Traffic and Transportation Engineering*, vol. 9, pp. 73-77, 2009.
[13] Zarita Zainuddin, Ong Pauline and Cemal Ardil, A Neural Network Approach in Predicting the Blood Glucose Level for Diabetic Patients, *International Journal of Information and Mathematical Sciences*, vol. 5, pp. 72-79, 2009.
[14] C.J. Dong and Z. YLiu, A simulation study of artificial neural networks for nonlinear time-series forecasting , *Computers & Operations Research*, vol. 28, pp. 381–396, 2001.
[15] P. Marchand and L. Marmet, Binomial smoothing filter: A way to avoid some pitfalls of least-squares polynomial smoothing, *Review of Scientific Instruments*, vol. 54, pp. 1034 - 1041, 1983.
[16] Liu Hong, Cui Wenhua, Zhang Qingling, Nonlinear Combination Forecasting Model and Application Based on Radial Basis Function Neural Networks, *IITA International Conference on Control, Automation and Systems Engineering*, Zhangjiajie, 2009, pp. 387-390.
[17] H. Akaike, A New Look at the Statistical Model Identification, *IEEE Trans.Automat.Control*, vol. 19 pp. 716-723, 1974.
[18] C. Chatfield, Model uncertainty and forecast accuracy, *J. Forecasting*, vol. 15 pp. 495–508, 1996.
[19] M.J. Campbell, A.M. Walker, A survey of statistical work on the MacKenzie River series of annual Canadian lynx trappings for the years 1821-1934 and a new analysis, *J. R. Statist. Soc. Ser. A* , vol. 140 pp. 411–431, 1977.

Fengxia Zheng was born in Hubei Province,China,in 1985. She received the B.S. degree from Hubei University for Nationalities, Enshi, in 2009. She is currently pursuing the M.S. degree in applied mathematics from University of Electronic Science and Technology of China, Chengdu, China. Her research interests include neural networks and time series forecasting.

Shouming Zhong was born in 1955 in Sichuan, China. He received B.S. degree in applied mathematics from UESTC, Chengdu, China, in 1982. From 1984 to 1986, he studied at the Department of Mathematics in Sun Yatsen University, Guangzhou, China. From 2005 to 2006, he was a visiting research associate with the Department of Mathematics in University of Waterloo, Waterloo, Canada. He is currently as a full professor with School of Applied Mathematics, UESTC. His current research interests include differential equations, neural networks, biomathematics and robust control. He has authored more than 80 papers in reputed journals such as the International Journal of Systems Science, Applied Mathematics and Computation, Chaos, Solitons and Fractals, Dynamics of Continuous, Discrete and Impulsive Systems, Acta Automatica Sinica, Journal of Control Theory and Applications, Acta Electronica Sinica, Control and Decision, and Journal of Engineering Mathematics.