TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.4, April 2014, pp. 3224 ~ 3229
DOI: http://dx.doi.org/10.11591/telkomnika.v12i4.4928

The Combined Forecasting Model of Discrete Verhulst-BP Neural Network Based on Linear Time-Varying

Shang Hongchao[1], Long Xia*[2], He Tingjie[3]
[1]School of Computer Science, Sichuan University of Science & Engineering
Sichuan University of Science & Engineering, 643000 Zigong, China, Phone: 18080466578
[2]School of Computer Science, Sichuan University of Science & Engineering
Lecturer, School of Computer Science, Sichuan University of Science & Engineering, 643000 Zigong, China, Phone: 13890058580
[3]Sichuan Institute of Technology, Institute of automation and electronic information
Institute of automation and electronic information, 643000 Zigong, China, Phone: 18227735574
*Corresponding author, e-mail: 35262719@qq.com[1], longxia10-28@163.com[2], 731129344@ qq.com[3]

Abstract

Firstly, this paper, aiming at the problem of errors produced by the transformation of differential equation directly into difference equation from traditional gray Verhulst model,through generating reciprocal for the original data sequence, constructs the discrete Vrhulst model based on linear time-varying (LTDVM model);And then we, taking the LTDVM predicted value as an input value and the original data as a mentor training value, put forward the combined forecasting model of discrete Verhulst-BP neural network based on linear time-varying. Meanwhile, in order to improve the training speed and agility and effectively avoid the saturation region of S-type function, this article normalized in advance the input data and mentor training values to better ensure the usefulness, self-learning ability and fault tolerance of the model. At last, we will study the cases to demonstrate that the model has high modeling and forecasting accuracy.

Keywords: *discrete Verhulst model, linear time-varying, BP neural network, combination forecasting*

1. Introduction

Forecast reefers to, on the basis of mastering existing information and in accordance with certain means and rules, the measure and calculation for the future things to know in advance the development process and results of tings. In actual forecast, generally based on historical data variables, we use statistical methods or system identification method to establish mathematical models for prediction. Although there are many existing predictive models, the time series model which is based on the simple regression analysis and the theory of probability and statistics, only has better predict results for the data of linear variation, thus hard to accurately describe the time series trends. Therefore, the gray system theory and neural network models the nonlinear prediction models such as the gray system theory and neural network models emerged. Gray model based on limited information can fit the overall trend of the time series and improve the prediction accuracy, but the gray system theory can not approach to a nonlinear function in spite of training learning. While neural network model is able to solve this problem very well, especially the BP neural network possessing a strong nonlinear mapping ability. Through the efforts of a large number of scholars, BP neural network has complete theoretical system and clear algorithmic process, with a strong analog and recognition. However, with the deepening of BP neural network application research, some of its problems come to expose, such as long learning time, slow convergence and poor generalization ability, and these problems have a serious impact on the prediction accuracy of BP neural network [1]. Thus, researchers began to focus on BP neural network improvement and combination forecast. Li Huanrong [2] (2000) proposed a method which can not only reduce the amount of sample input, but also can improve the convergence speed is an approach to optimize the traditional neural network. Cao Jianhua, etc. [3] (2008) according to the error determining combined weights of gray model and neural network model, build a combination forecasting model. Dai Yu (2010) [4] constructed the combined weights forecasting model of

Received September 24, 2013; Revised November 21, 2013; Accepted December 10, 2013

RBF neural network method, gray GM(1, 1) method and ARIMA method, but with error and other limited information to determine the weight and establish the combined forecasting model, the effective complementary advantages of two or more models can not be formed. Therefore, there has been more than two prediction models research for the effective integration, such as Li Weiguo [5] (2007) using gray system theory to extract the trend item of time series and applying the sample periodogram to fit periodic terms, finally create a combination forecasting model. Shi Biao [6] (2009) by the means of using PSO algorithm to train BP neural network, optimized neural network parameters and improved the generalization ability of neural network. Liu Rentao etc. [7] (2008) through using real-coded acceleration genetic algorithm to optimize GM(1, 1) parameters and taking the improved GM(1, 1) prediction value as input values and the original data as output values to combine with BP neural network, finally obtained a higher prediction accuracy. Tong Xinan etc. [8] (2011) based on the Verhulst model and BP neural network model, made a research on the combination of these two models and thought that the combination model has a good stability, while the modeling prediction results for the shape of "s" type of oscillation sequence were not ideal.

This paper firstly according to the arising problems from the transformation of differential equations directly into difference equations of traditional gray Verhulst model, and through generating countdown to the original data sequence, establish a gray model with no bias to "s" type sequence simulation--discrete Verhulst model based on linear time-varying(LTDVM model);Then this paper, taking LTDVM predicted value as an input value and the original data as a mentor training value, proposes based on linear time-varying discrete Verhulst-BP neural network combination forecast model. Meanwhile, in order to improve the training speed and agility and effectively avoid the saturation region of S-type function, this article priors to normalize the input data and mentor training value to better make sure that this model has higher practicability, self-learning ability and fault tolerance.

2. Based on Linear Time-varying Discrete Verhulst-BP Neural Network Combination Forecast Model

2.1. Discrete Verhulst Model Based on Linear Time-varying (LTDVM model)

Definition 1 take the observation value of a behavioral characteristic sequence of the system as: $X^{(0)}=\{x^{(0)}(1),x^{(0)}(2),\cdots,x^{(0)}(n)\}$, $Y^{(0)}$ is the countdown sequence of $X^{(0)}$, that is, $y^{(0)}(k)=\frac{1}{x^{(0)}(k)}(k=1,2,\cdots,n)$, $Y^{(1)}$ is a cumulative sequence of $Y^{(0)}$, that is:

$$y^{(1)}(k+1)=(\beta_1+\beta_2 k)y^{(1)}(k)+\beta_3 k+\beta_4,\ k=1,2,\cdots,n-1 \tag{1}$$

Is the discrete Verhulst model based on linear time-varying (LTDVM model).

The solving process of this model is as follows:

1) Using the least squares method to find the model parameters $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_4$

$$\hat{\beta}_1=\frac{B_1}{A},\ \hat{\beta}_2=\frac{B_2}{A},\ \hat{\beta}_3=\frac{B_3}{A},\ \hat{\beta}_4=\frac{B_4}{A}$$

2) Applying the recursion formula $\hat{y}^{(1)}(k+1)=\left(\hat{\beta}_1+\hat{\beta}_2 k\right)y^{(1)}(k)+\hat{\beta}_3 k+\hat{\beta}_4$ to find the sequence $\hat{y}^{(1)}(k+1)$

3) Using $\hat{x}^{(0)}(k+1)=\frac{1}{\hat{y}^{(0)}(k+1)}$ to find the simulation prediction value of the original sequence.

2.2. Improved BP Neural Network

The essence of prediction is the exploration of law from the seemingly chaotic historical data.

The neural network model has these features: self-learning information processing, knowledge reasoning and self-adaptation to non-deterministic rule system. Through the training to sample data to achieve some kind of mapping from input to output, so by the mapping the inherent law of sample data can be discovered. After a long period of development, artificial neural network research has achieved fruitful research results and currently the most widely used model is BP neural network model.

In the application of BP neural network to predict, the primary task is to establish BP neural network model and during the model process, determining the network layers and the number of neurons in each layer is the key.

(1) Network layer

In the BP neural network model, hidden layers determine the speed of model training, but in practice, increasing the hidden layer needs more training time, so generally only the structure containing the input with one hidden layer, hidden layer and output layer can be selected.

(2) Determination of hidden nodes

The selection of hidden layer nodes is also very important. If the number of neurons in the hidden layer is too small, the network performance is poor or can not identify the completion of training. If the selection of the number of nodes is too large, the number of iterations may increase, the training time may be prolonged, the network fault tolerance may decreased, and the generalization capacity may diminish. All these issues can lead to deterioration of the model prediction.

In order to select a reasonable number of hidden nodes, there is an empirical formula:

$$i = \sqrt{m+n} + a \tag{2}$$

Thereinto, i stands for the hidden layer nodes, m for the number of input nodes, n for the number of output nodes, and the range of a is 1-10.

(3) Data preprocessing

Because the input layer and hidden layer in the BP neural network applied the tansig function which is "'s" type of transfer function whose range is [-1,1] or [0,1].In order to improve the training speed and agility and effectively avoid the saturation region of S-type function, the range of input data is generally required between[-1,1] or [0,1].This article priors to normalize the input data and mentor training value to make its range be [0,1], and then brings the processed data into the BP neural network to train, and finally anti-normalizes the estimated results to get the required data.

Normalization formula [10]:

$$T = 0.1 + \frac{x - x_{\min}}{x_{\max} - x_{\min}} * 0.8 \tag{3}$$

There into, T represents the normalized target data and x is the original data.

Anti-normalization formula:

$$x = x_{\min} + \frac{(T - 0.1)(x_{\max} - x_{\min})}{0.8} \tag{4}$$

2.3. Based on Linear Time-varying Discrete Verhulst-BP Neural Network Combination Forecast Model

The gray system and neural network integration into a gray neural network model can complement each other. By the gray prediction method, building model requires a small amount of computation and in the case of small samples this way can achieve higher accuracy; the use of BP neural network contributes to building a model with high precision and error control. Therefore, integration of the two together can give full play to the advantages of both. At the same time, because the gray prediction model was constructed based on a linear time-varying discrete Verhulst Model(LTDVM model), this model has no bias to the data of the "s" type and can simulate and predict the oscillatory data very well; Moreover, based on the data normalized BP neural network makes further improvement, effectively improving the training speed and agility as well as avoiding the saturation of S-type function to make the combination forecasting model with higher simulation and prediction accuracy.

3. Case Analyses

Select one of application examples in the literature [9] and analyze and compare the sample of one cost of Torpedo in 1995-2003. Table 1 shows the raw data. According to statistics, the cumulative cost of Torpedo approximates curve "S" type, suitable for the establishment of new discrete Verhulst Model (LTDVM model).

Table 1. A Type of Torpedo Development Cost unit: million

years	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
cost	496	779	1187	1025	488	255	157	110	87	79

Table 2. A Type of Torpedo Development Cost unit: million

years	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
cost	496	1275	2462	3487	3975	4230	4387	4497	4584	4663

This type of Torpedo cumulative development cost is shown in Table 2. It can be seen from the table that the growth of Torpedo cumulative development cost slows down and the data series presents "S" shape. Therefore, according to the discrete Verhulst model (LTDVM model) based on linear time-varying, the data shown in Table 2 can be simulated and predicted. $y^{(0)}(k)=\frac{1}{x^{(0)}(k)}(k=1,2,\cdots,n)$, thereinto, $y^{(1)}(k)=\sum_{i=1}^{k}y^{(0)}(i)$, $k=1,2,\cdots n$; according to definition 4 and by the least squares to find the model parameters $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_4$ and get the prediction expression: $y^{(1)}(k+1)=(0.3012-0.0076k)y^{(1)}(k)+0.0002k+0.002, k=1,2,\cdots,n-1$. The results are shown in Table 3.

Table 3. LTDVM Model Simulation Results to a Type of Torpedo Cumulative Development Cost Unit: million

years	data	Traditional Verhulst model[8]		Discrete Verhulst model based on linear time-varying (LTDVM model)	
		Analog value	Relative error	Analog value	Relative error
1995	496				
1996	1275	1119.1	0.123	1274.8908	0.0086
1997	2462	2116.0	0.1405	2465.0106	0.1223
1998	3487	3177.5	0.0888	3473.1565	0.3970
1999	3975	3913.7	0.0154	3983.2205	0.2068
2000	4230	4286.2	0.0133	4241.6250	0.2748
2001	4387	4444.8	0.0132	4387.4620	0.0105
2002	4497	4507.4	0.0023	4487.3902	0.2137
2003	4584	4531.3	0.0115	4575.9492	0.1756
2004	4663	4540.3	0.0263	4671.3441	0.1789
Average relative error			0.0482		0.1765

Note: the relative error in the table taking the absolute value.

On this basis, use the improved BP network model to simulate and predict. The specific approach is the establishment of a three-layer BP neural network, taking the fit part of the discrete Verhulst model based on linear time-varying as the input of neural network and the original data as the mentor training BP neural network.

The first step: network parameter setting

The network structure is selected as 9-8-1, the transfer function of input layer and hidden layer neurons is chosen as *tansig*, the transfer function of the output layer neurons is *purelin* and the training function takes the adaptive modification learning rate algorithm *traingda.*This kind of approach in the training project automatically modifies the learning rate, let it change always in a suitable range to ensure system stability and speed of network training (learning rate is too large, which will reduce the stability of the network; on the contrary, the training time will be longer). Set the maximum number of epochs of training steps as 5000steps, the expected error goal as 0.00001, the initial learning rate is generally selected between 0.01 and 0.1 and this paper selects lr=0.05, learning rate increment lr_inc as 1.05.

The second step: data normalization

Because the input layer and hidden layer in the BP neural network applied the *tansig* function which is "'s" type of transfer function whose range is [-1,1] or [0,1].In order to improve the training speed and agility and effectively avoid the saturation region of S-type function, the range of input data is generally required between[-1,1] or [0,1]. This article priors to normalize the input data and mentor training value to make its range be [0, 1], and then brings the processed data into the BP neural network to train, and finally anti-normalizes the estimated results to get the required data.

Normalization formula [10]:

$$T = 0.1 + \frac{x - x_{\min}}{x_{\max} - x_{\min}} * 0.8 \tag{3}$$

Thereinto, T represents the normalized target data and x is the original data.

Anti-normalization formula:

$$x = x_{\min} + \frac{(T - 0.1)(x_{\max} - x_{\min})}{0.8} \tag{4}$$

The third step: using the combination model to simulate and predict

Bring the normalized input data P and the mentor training value T into the BP network to train and then get a predictive network.Using the trained neural network to predict the input data for simulation, to obtain the desired predictive value, and calculate the relative error (see Table 4). Compared with 0.0918 of the model in the literature [8], the accuracy of this combination model is much higher.

Table 4. The Comparison Table of Calculation Results of Model Unit: million

years	data	Gray Verhulst-BP network combination model[8]		Based on linear time-varying discrete Verhulst-BP neural network combination prediction model	
		Analog value	Relative error	Analog value	Relative error
1995	496	496.3691	0.0007	496	0
1996	779	777.5959	0.0018	779.3	0.0004
1997	1187	1190.0080	0.0025	1187.2	0.0002
1998	1025	1018.1120	0.0067	1025	0
1999	488	505.1010	0.035	488.2	0.0004
2000	255	225.9330	0.114	255.1	0.0004
2001	157	173.0380	0.1022	157.7	0.0045
2002	110	128.7950	0.1709	100.8	0.0836
2003	87	81.1100	0.0677	83.5	0.0402
2004	79	46.0700	0.4168	91.9	0.1633
Average relative error			0.0918		0.0293

Note: the relative error in the table taking the absolute value.

4. Conclusion

Gray Verhulst model and BP neural network model has its own shortcomings and the establishment of a combination of the prediction model effectively play their respectively advantages. Moreover, the respective improvement of gray prediction model and BP neural network makes this combination model simulate and predict the data of s-type with higher accuracy.

Acknowledgements

This work was supported by the basic application research Program of Sichuan Province (No. 2008JY0112), the Higher Education Personnel Training Quality and Teaching Reform Subject of Sichuan Province (No. P09264), the Department of education key scientific research Program of Sichuan Province (No. 2006A077), and the Foundation of Sichuan Province for Retuned Scholars (Sichuan people social letter No. 32 in 2010).

References

[1] Liu Tianshu. The improvement research and application of BP neural network. Harbin: Northeast Agricultural University. 2011.
[2] Li Huanrong, Wang Shuming. An improved BP neural network prediction method and its application. *Systems Engineering.* 2000; 18(5): 75-78.
[3] Cao Jianhua, Liu Yuan, Dai Yue. Network traffic prediction based on grey neural network integr ated model. *Computer Engineering and Applications*. 2008; 44(5): 155-157.
[4] Dai Yu. Optimal Combination Forecasting Model and Its Application. *Economic Mathematics.* 2010; 27(1): 92-98.
[5] Li Weiguo, Zhang Aiqing. The modeling method of combination forecast based on gray system. *Statistics and Decision*. 2007; (21): 11-12.
[6] Shi Biao, Li Yuxia, Yu Xinhua, Yan Wang. Short-term load forecasting of improved PSO-BP neural network model. *Computer Applications.* 2009; 29(4): 1036-1039.
[7] Liu Rentao, Fu Qiang, Feng Yan, Gai Zhaomei, Li Guoliang, Li Weiye. Gray BP neural network prediction model based on RAGA and its impact on the Sanjiang Plain groundwater dynamic forecasting. *System Engineering Theory and Practice*. 2008; 28(5): 171-176.
[8] Tong Xinan, Wei Wei. Gray Verhulst-BP Network Combined Model in Forecasting Research. *Computer Engineering and Applications.* 2011; 47(23): 245-248.
[9] Liu Sifeng, Dang Yaoguo, Fang Zhigeng etc. Gray system theory and its application. Beijing: Science Press. 2010: 176-179.
[10] Wang Yingying. Coal logistics demand prediction research based on gray neural network model. Beijing: Beijing Jiaotong University. 2012.
[11] Xia Long, Yong Wei, Ping Liao, Yan Liu. Linear Time-varying Parameters Discrete Gray Model Based on Oscillating Sequence. *Journal of Systems Science and Information*. 2012; 10(4): 313-318.
[12] Xia Long, Yong Wei, Zhao Long. The Combined Forecasting Model of Gray Model Based On Linear Time-variant and ARIMA Model. *Academic Research Publishing Agency*. 2013; 16(3).
[13] Guangyou Yang, Zhijian Ye, etc. The Implementation of S-curve Acceleration and Deceleration Using FPGA. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2013; 11(1): 279-286.
[14] TingZhong Wang, Gang Long Fan. The Research of Building Fuzzy C-Means Clustering Model Based on Particle Swarm Optimization. *TELKOMNIKA Indonesian Journal of Electrical Engineering.* 2013; 11(12): 7589-7598.