

An Improved Neural Networks Prediction Model and Its Application in Supply Chain

Xiaoni Dong¹, Guangrui Wen²

¹Department of Industrial Engineering, Xi'an Siyuan University,
Xi'an Shaanxi 710038, China, birdy_dong@163.com

²College of Mechanical Engineering, Xi'an Jiaotong University,
Xi'an Shaanxi 710049, China, grwen@mail.xjtu.edu.cn

Abstract: Accurate prediction of demand is the key to reduce the cost of inventory for an enterprise in Supply Chain. Based on recurrent neural networks, a new prediction model of demand in supply chain is proposed. The learning algorithm of the prediction is also imposed to obtain better prediction of time series in future. In order to validate the prediction performance of recurrent neural networks, a simulated time series data and a practical sales data have been used. By comparing the prediction result of Multi-Layer feedback neural networks and recurrent neural networks, it can be shown that the recurrent neural networks prediction model can help in improving the prediction accuracy. [Nature and Science. 2006;4(3):23-27].

Keywords: supply chain; inventory management; recurrent neural networks; multi-step prediction

1. Introduction

A company may hold inventories of raw materials, parts, work in process, or finished products for a variety of reasons, such as the following [1]: To create buffers against the uncertainties of supply and demand; To take advantage of lower purchasing and transportation costs associated with high volumes; To take advantage of economics of scale associated with manufacturing products in batches; To build up reserves for seasonal demands or promotional sales; To accommodate product flowing from one location to another (work in process or in transit); To exploit speculative opportunities for buying and selling commodities and other products.

Recently, attention has focused on developing better forecasting models that reduce or eliminate inventories, which reflects the cost of inventory management in supply chain. ANN, as a new tool, has already been used in forecasting field due to their ability to learn complicated functions.

In 1995, Chao-Chee Ku [2] proposed dynamical recurrent neural networks (DRNN) that overcome the static modeling problem of multi-layers feed-forward neural networks in modeling. Y. X. He [3] combined data mining & knowledge discovery and neural networks technology to solve inventory problem. C. T. Xuan [4] summarized applications of neural networks in supply chain management (SCM) including optimization, prediction, decision-making support, modeling & simulation and management. This paper is a first attempt to use recurrent neural networks as forecasting technology to reduce the uncertainty of inventory management.

2. Prediction Model Based on Feed-forward Neural Networks

The neural networks models most widely used in time series applications are based in feed-forward neural networks with backpropagation learning algorithm [5]. This model usually constructs multi-layer feed-forward neural network, and then evaluate function F by time series. These models consist of a common nonlinear auto-regressing model appearing in Equation (1), which is also called NAR Model. In this case, the value at $k+1$ of this time series is often represented by $d+1$ entries time series values, thus:

$$x(k+1) = F(x(k), \dots, x(k-d)) \quad (1)$$

Where k represents the time variable, F is a nonlinear function that defines this time series.

The common prediction method based on neural network, namely the single-step neural network prediction method, set up non-linear prediction model of the neural network. The following predicting equation represents the prediction model:

$$x(k+1) = \sum_i^d f_i(x)x(k-i) + \varepsilon \quad i = 0, \dots, d \quad (2)$$

in which $f_i(x)$ is the non-linear function of input variable, $d+1$ is the number of input nodes in predicting network.

During the Neural network predicting, the following equation can be got from the predicting model:

$$\tilde{x}(k+1) = \sum_i^d f_i(x)x(k-i) \quad (3)$$

According to the predicting model, figure 1 shows the structure of the Multi-step predicting. Use the predicting value into the predicting model; the multi-step value of the future can be got step by step. In this kind of multi-

step predicting process, when the single-step predicting result is got, there has being a predicting error which can be expressed as:

$$\varepsilon_{k+1} = \frac{1}{2} [x(k+1) - \tilde{x}(k+1)]^2 \quad (4)$$

During the course of inference, using $\tilde{x}(k+1)$ and the following predicting value as the input of the network leads to the input error. And as the increasing of the predicting value, the accumulating error increases rapidly, which cannot ensure the accuracy of multi-step predicting. Figure 1 shows the traditional neural networks recursion multi-step prediction model.

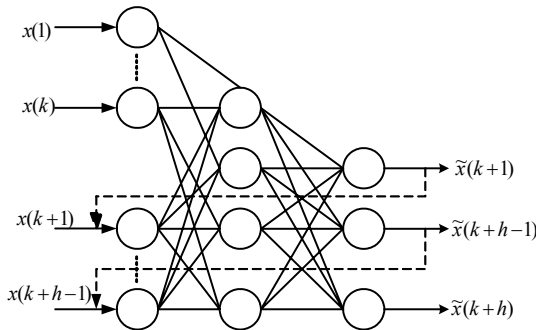


Figure 1. Traditional Neural Networks Prediction Model

3. Prediction Model Based on Recurrent Neural Networks and Its Learning Algorithm

In this section, a prediction method based on recurrent neural networks is introduced. The recurrent neural networks used to build up the prediction model consist of adding feedback connections from the output neuron to the input layer, which memorize previous prediction values^[6]. In this case, parameters of the multi-step prediction model are determined to minimize the error along interval $[k+1, k+h+1]$, where h is the length of multi-step prediction. Thus, the model is trained for long-term prediction.

3.1 Description of Recurrent Prediction Model

In order to overcome limitation of traditional feed-forward prediction model and make full use of the input vectors to the model, a new prediction model based on recurrent networks is built up^[7-8]. The recurrent network is constructed by starting from a multilayer feed-forward neural network and by adding feedback connections from the output neuron to the input layer as shown in Figure 2.

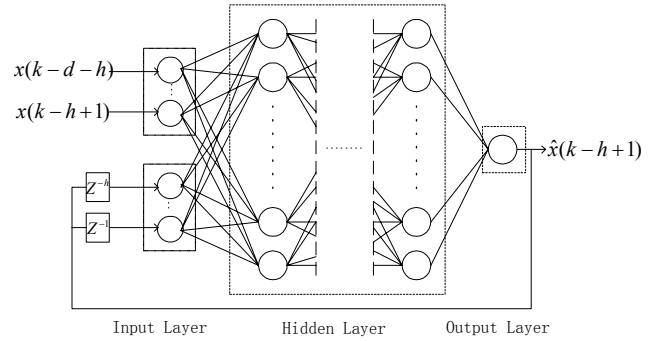


Figure 2. Recurrent network architecture

Neural network is divided into three parts: input layer, hidden layer and output layer. Where \hat{x} represents the network output, and Z^{-i} is an operator that delays by i terms the network output sequence. The input layer is composed of two groups of neurons. The first group acts as the external inputs to the network gathering the original or measured time series data. The second group is formed by the context neurons, which memorize previous output of the network. Introducing the vector $C(k)$ ($C(k) = (C_1(k), \dots, C_h(k))$) to indicate the activation of context neurons, each component is calculated as:

$$C_i(k) = Z^{-i}(\hat{x}(k+h+1)) = \hat{x}(k+h+1-i) \quad i = 1, \dots, h \quad (5)$$

3.2 Learning Algorithm[9]

Assuming that the prediction length is fixed to h , and that at instant k the goal is to predict the time series values at instants $k+1, k+2, \dots, k+h+1$, the number of input units decreases from $d+1$ to $d+1-h$, and the number of context neurons increases from 0 to h , respectively. Thus, the sequences received by the external inputs and the context neurons, at every instant k , are given by the following sequence:

1. The number of context neurons is initialized from zero, which will be included in the external inputs accordingly.

2. The future instants $k+i$ for $i = 2, \dots, h+1$ are not real, but simulated. Now the input units receive the vector $x(k+i), \dots, x(k-d-i)$, and the $(i-1)^{th}$ context neurons memorize the previous $i-1$ outputs of the network, i.e.:

$$C_1(k) = \hat{x}(k+i-1) \dots C_{i-1}(k) = \hat{x}(k+1) \quad (6)$$

3. The external inputs and the context neurons are resettled.

Below are the complete training procedures of a multi-step recurrent prediction model. At each instant k , starting with $k = d$:

Step1. The number of context neurons is initialized to zero. $d+1$ external input neurons are one set receiving the measured values of the time series, $x(k), \dots, x(k-d)$. The output of network is given by following equation:

$$\hat{x}(k+1) = \hat{F}(x(k), \dots, x(k-d), W_2) \quad (7)$$

Step 2. The number of context neurons is increased in one unit and the number of external units is decreased also in one unit. The context neuron memorizes the output of the network previously calculated, $\hat{x}(k+1)$. Thus the prediction at the simulated instant $k+2$ is given by:

$$\hat{x}(k+2) = \hat{F}(\hat{x}(k+1), x(k), \dots, x(k-d+1), W_2) \quad (8)$$

Step 3. Step 2 is repeated until h context neurons will be achieved. The output of the recurrent model at simulated instants $k+3, \dots, k+h+1$ are given by the following equations, respectively:

$$\hat{x}(k+3) = \hat{F}(\hat{x}(k+2), \hat{x}(k+1), x(k), \dots, x(k-d+2), W_2) \quad (9)$$

$$\hat{x}(k+h+1) = \hat{F}(\hat{x}(k+h), \dots, \hat{x}(k+1), x(k), \dots, x(k-d+h), W_2) \quad (10)$$

Step 4. The parameter set of the model, W_2 , is updated by following the negative gradient direction of the error function given by:

$$e(k+1) = \frac{1}{2} \sum_{i=1}^h (x(k+i+1) - \hat{x}(k+i+1))^2 \quad (11)$$

In order to avoid the long computational effort required by dynamic back-propagation rules when the prediction length is high, the updates of the parameters are realized using traditional backpropagation learning rule.

Step 5. At this moment the time variable k is increased in one unit and the procedure is returned to Step 1.

The procedure is repeated for the complete training set until it reaches the convergence.

4. Simulation Experiment

The simulated time series is given by following equation:

$$x(k+1) = \lambda x(k)(1-x(k)) \quad (12)$$

Where $\lambda = 3.97$ and $x(0) = 0.5$.

The equation describes a strong chaotic time series. As to enterprise, customer's demand is of uncertainty due to the influence in such measures as the rival, the change of customer's favor, production cycle and the price reducing, etc., the supply chain system is a complicated non-linear system. So, such unordered time array can well use simulate such a system, at the same time, can test the performance of the prediction model.

We initialize parameter k from $k=0$ to $k=100$. Table 1 and 2 show the prediction errors of two models over the training data and predicting data respectively. Figure 3 and Figure 4 display the single step prediction curves of two models respectively and Fig.5 and Fig.6 display the multi-step prediction curves of two models respectively. It is evident that multi-step recurrent prediction model's prediction accuracy is much higher than that of traditional model.

Table 1. Logistic time series: prediction errors over the training data

length h	Traditional model (3-10-1)	Recurrent model (7-15-1)
0	0.0015	0.0015
7	0.0850	0.0135

Table 2. Logistic time series: prediction errors over the predicting data

length h	Traditional model (3-10-1)	Recurrent model (7-15-1)
0	0.002	0.002
7	0.105	0.026

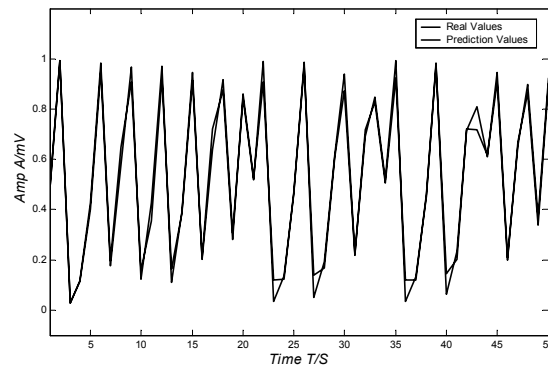


Figure 3. Traditional MFNN prediction at $h=0$

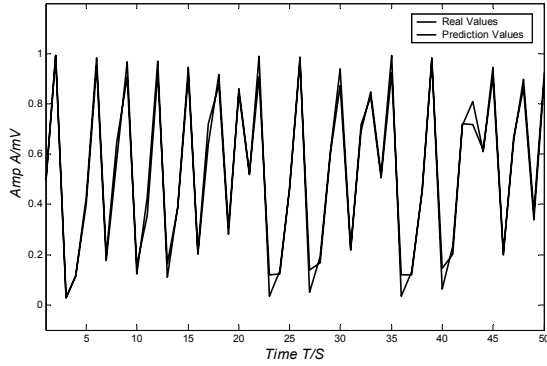


Figure 4. MSRN prediction at h=0

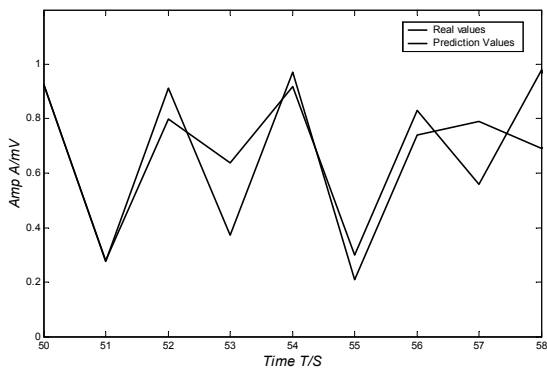


Figure 5. Traditional M-FNN model prediction at h=7

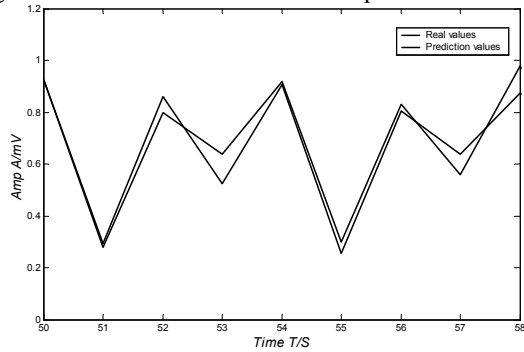


Figure 6. MSRN prediction at h=7

5. Application

This paper regards selling data of paper for one week of one papermaking enterprise as the target that is studied by the neural network^[10]. According to the experience and investigation from its marketing department, the factors affecting sales of paper include season, paper category, unit price, middleman in various regions, growth rate of macroeconomy in that year, etc..

We just consider two major influence factors, season and paper category, for building the forecasting model. Firstly, 52 data of week sales volume for No.1

warehouse in 2002 can be obtained by statistics method. Secondly, three layers recurrent neural networks model is adopted. The input layer includes four nodes, for considering the influence of different seasons. The output layer includes one node, and the hidden layer has ten nodes. First 46 data as training samples and last 6 data as testing samples are used to construct forecasting model of sales. Thirdly, in order to improve the rate of networks convergence in training processes, normalization technique is applied to 46 training data, and the batch training method (BTM) is adopted.

Table 3 lists the forecasting errors. Most errors are controlled below 0.1, and the trend of sales accords with actual situation. The results of training and forecasting are shown in Figure 7.

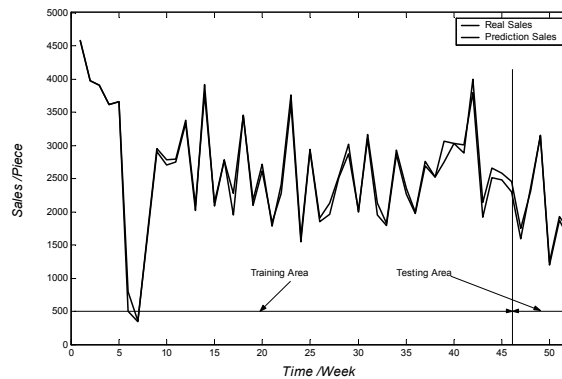


Figure 7. Training and forecasting of paper sales

In Figure 7, the data from 47th to 52th points in dash line denote the results of forecasting, and the solid line presents the real sales volume.

Table 3. Errors between Prediction Sales and Real Sales

Week	Forecasting data	Real data	error
47	1593	1743	0.0858
48	2361	2310	0.0220
49	3137	3151	0.0044
50	1201	1243	0.0337
51	1873	1924	0.0267
52	1581	1710	0.0757

6. Conclusions

This paper presents a new forecasting model for inventory management, which is based on multi-step recurrent neural networks. The recurrent neural networks have advantages of higher precision and robustness over the traditional feed-forward neural networks in multi-step prediction. The forecasting results in some paper mill certified the new method. The model will be favor of the inventory decision-making for decreasing the inventory to minimum.

Correspondence to:

Xiaoni Dong
Department of Industrial Engineering
Xi'an Siyuan University
Xian, Shaanxi 710038, China
E-mail: birdy_dong@163.com

References

1. Jeremy F. Shapiro, *Modeling the Supply Chain*, Wadsworth Group, a division of Thomson Learning, America, 2001.
2. Ku C.C., Lee K.Y., "Diagonal recurrent neural networks for dynamic Systems control", *IEEE Trans.Neuralnetworks*, Vol 6, No. 1, pp. 144-155, 1995.
3. Yanxiang He, Feng Li, Zhikai Song and Ge Zhang, "Neural Networks Technology for Inventory Management", *Computer Engineering and Application*, No.15, pp. 182-184, 2002.
4. Chaoting Xuan, Peiqing Huang and Dong Lu, Applications of Neural Network Technology in Supply Chain Management, *Industrial Engineering and Management*, No.3, pp. 41-44, 2000.
5. Svozil, Daniel, etc., Introduction to multi-layer feed-forward neural networks, *Chemometrics and Intelligent Laboratory Systems*, Vol 39, pp . 43-62, 1997.
6. M.G. Ines and I. Pedro, Multi-step learning rule for recurrent neural models: an application to time series forecasting, *Neural Processing Letters*, pp. 115-133, 2001.
7. I. Galvan and J. Zaldivar, Application of recurrent neural networks in batch reactors. Part I: NARMA modeling of the dynamic behavior of the heat transfer fluid temperature, *Chemical Engineering and Processing*, pp .505-518, 1997.
8. G.R. Wen and L.S. Qu, "Equipment Behavior Predictability Evaluation Based on Redundancy". *Journal of Xi'an Jiaotong University*, Vol 37, No.7, pp. 699-703, Jul. 2003.
9. G.R. Wen and L.S. Qu, Multi-step forecasting method based on recurrent neural networks, *Journal of Xi'an Jiaotong University*, Vol 36, No.7, pp. 722-726, Jul. 2002.
10. Shengshu Jiang , Chunjie Yang and Ping Li, Research on application of BP-ANN forecast in process supply chain, *Operations Research and Management Science*, Vol 12, No.5, pp. 57-61, Oct. 2003.