# Port Throughput Forecast Based on Nonlinear Combination Method

[1]Jianfeng Li, [1]Yan Chen, [2]Xusheng Cui
Department of Management Science and Engineering,
[1]Dalian Maritime University, [2]China Dalian International Cooperation Holdings Ltd.
Dalian city, PR. China

*Abstract:* -. Port throughput forecast plays a great role in the development of port logistics economy. This paper works over that significant subject based on a new kind of nonlinear combination method, which is composed of linear exponential smoothing model, simple moving average method and Elman network. The forecasted values through the first two methods are taken as the inputs of Elman network, which are shown as medium and long-term effect, what's more, some recent periods of port throughputs are chosen as the inputs too, which are shown as short-term effect, thus different influences are considered comprehensively. Otherwise, the actual port throughputs in specific time periods are taken as the outputs. According to the historical samples, Elman network is trained in supervised manner for the future forecast. In this way, the different singular methods are combined together well and that new kind of nonlinear method shows higher forecasting accuracy.

*Key-Words:* Nonlinear combination method; Forecast; Elman network; Port Throughput; Simple moving average method; linear exponential smoothing model

## 1 Introduction

With the formation of economic globalization, the foreign trade has been keeping a high growing rate, which promotes the development of port logistics[1]. In 2006, the freight is 2,101,800 million Yuan in China, increasing 12.8% compared with the same period of last year. For the port logistics, port throughput forecast is one of basic and significant subjects, which is useful for port transportation layout, logistic material allocation, port investment decision, and so on, thus it has been received great attention from many scholars.

Some scholars have worked over port throughput forecast in qualitative analysis. They depended mainly on their experience to predict port throughput according to some factors, such as region's economy, transferring ability and port condition. The others investigate into that affair in quantitative analysis. They built some models to predict the port throughput by applying mathematical technique according to some historical data, such as exponential smoothing model, grey model, moving average method and neural network model [2]. In the recent years, more and more scholars have paid attention to the combination theory which makes up two or more singular methods. The combination theory means that forecast will be carried on in the different points that different singular method be hold on, thereby, it can include more value forecasting information and has been proved with higher forecasting accuracy [3]. In the combination theory, the nonlinear combination method is considered as the most important subject. Many scholars have done their best to research on it, For example, Voort M and Dougtherty M have Combined Kohonen maps with ARIMA time series models to forecast traffic flow which have showed better forecasted result[4].

However, although many researchers have chosen some singular methods for forecast in the combined view, the reasons of combination in that way are often not given out clearly. Therefore, this paper makes full use of different singular method's characteristics, which also show why they are combined together, and brings forward a new kind of nonlinear combination method to predict the port throughput. It will be proven with higher forecasting accuracy.

## 2 Framework of the nonlinear combination

Port throughput is a kind of time series data, which fluctuates both in the short term and with the medium and long trend. Because linear exponential smoothing model (LESM) and simple moving average method (SMA) have a good effect on revealing the trend, they are combined in the nonlinear combination forecast method, and some recent data is also chosen to show the short term, as a result, the different influences are considered comprehensively. What's more, on account that Elman network has good effect on the time series data forecast, which is a kind of artificial neural network

(ANN) can fit any curve, it is selected as the main part of the method to reflect the mapping relationship between different influences and forecasted value. Supposing the time series port throughputs are presented as $\{p_1, p_2, p_3, \cdots, p_t\}$, where t is the period of time, the frame of nonlinear combination is shown in Fig. 1.

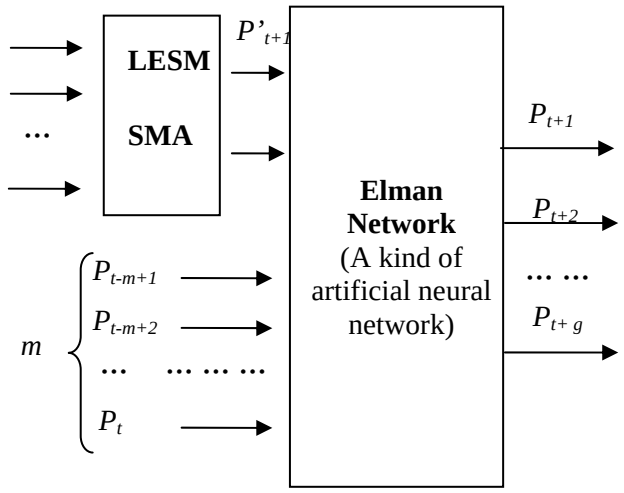


**Fig1. The framework of nonlinear combination**

Here, the forecasted values $P'_{t+1}$ are achieved in the linear exponential smoothing model and simple moving average method, and then, some port throughputs $\{p_{t-m+1}, p_{t-m+2}, p_{t-m+3}, \cdots, p_t\}$ in the short $m$ periods are also chosen to present for the short-term influence. They are as the inputs of artificial neural network, and $g$ periods of port throughputs $\{p_{t+1}, p_{t+2}, p_{t+3}, \cdots, p_g\}$ are taken as the outputs. According to the historical data, some samples can be formed, which realizes the mapping relationship from $R^{m+2}$ to $R^g$. After being trained, Elman network can be used to forecast the future port throughputs. In this way, the different singular methods are nonlinearly combined well. It embodies characteristics of the singular methods and includes the short, medium and long-term influence.

## 3 Arithmetic of nonlinear combination method

This new kind of nonlinear combination method to forecast the port throughput is mainly composed of three singular methods, which are shown below:

1) Linear exponential smoothing model: The basic formula is shown below (Formula 1), where $\alpha$ is the smooth coefficient ( $0 < \alpha < 1$ ), $S_t$ is the index-smooth value in the $t$ period, and $P_t$ is time series data, i.e. port throughput in the $t$ period.

$$S_t = \alpha P_t + (1-\alpha) S_{t-1} \quad (1)$$

The initial index-smooth value $S_1$ can be assigned with $P_1$ or the average of time series data in the front. Thus, $S_t$ will be achieved with the Formula 1, and it can be as the forecasted value of port throughput in the *t+1* period (Formula 2).

$$P'_{t+1} = S_t \quad (2)$$

2) Simple moving average method: Because time series data fluctuates around the average line, the average value can be as the forecasted value for showing the trend, The forecasted value of port throughput in the $t$+1 period can be shown in Formula 3, where $n$ is the number of the periods in which the average of time series data is calculated.

$$P'_{t+1} = \sum_{i=1}^{n} P_{t-i+1} / n \quad (3)$$

3) Elman network: As a form of recurrent neural network, it can be based on the current inputs plus a record of the previous state(s) and outputs of the network, so it has an ability of remembering the front state, thus, Elman network has more advantage on time series data forecast than other normal artificial neural networks[5]. It can be shown in Fig 2.

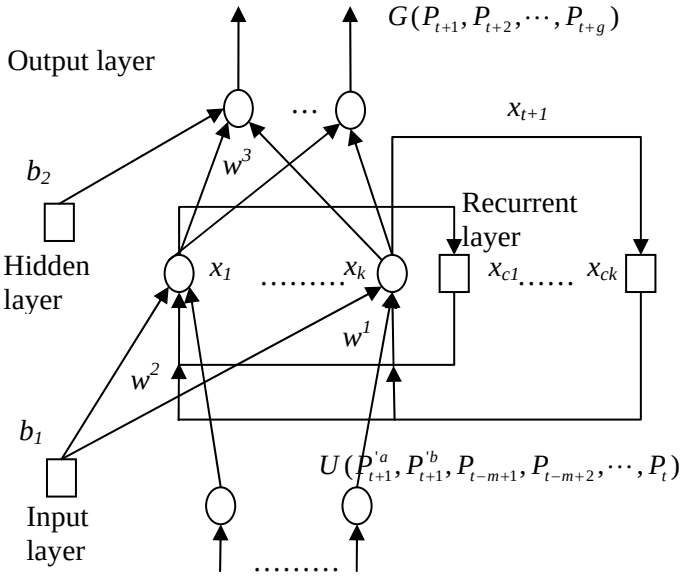


**Fig 2. The structure of Elman network**

The inputs of network are the forecasted values $(P_{t+1}^{'a}, P_{t+1}^{'b})$ and some the front port throughputs $(P_{t-m+1}, P_{t-m+2}, \cdots, P_t)$ and the outputs are the port throughputs $(P_{t+1}, P_{t+2}, \cdots, P_{t+g})$ in the $g$ periods. The $k$ is the numbers of nerve cells in the hidden (recurrent) layer; the $b_1$, $b_2$ are the biases; and the $w^1$, $w^2$, $w^3$ are the weights between the layers. The basic formulas are shown below.

$$G = g(w^3 \cdot x_{t+1} + b_2) \quad (4)$$

$$x_{t+1} = f(w^1 \cdot x_{c(t+1)} + w^2 \cdot U + b_1) \quad (5)$$

$$x_{c(t+1)} = x_t \quad (6)$$

Here, $f(\cdot)$ is the transfer function in the hidden layer and $g(\cdot)$ is the transfer function in the output layer. According to the historical data, that Elman network is trained in a supervised manner for forecasting. In this way, Elman network is as the main part of combination method and different singular methods are nonlinearly combined together well.

## 4 A Case Analysis

There are some monthly statistical data about the main port throughputs in the coast from January in 2004 to October in 2007, which come from National Bureau of Statistic of China. They are shown in Table 1.

**Table 1. The port throughputs in the main Chinese coast (billion)**

| | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|
| Jan | 0.179 | 0.229 | 0.265 | 0.311 |
| Feb | 0.181 | 0.2 | 0.236 | 0.285 |
| Mar | 0.198 | 0.229 | 0.274 | 0.302 |
| Apr | 0.199 | 0.24 | 0.282 | 0.323 |
| May | 0.202 | 0.242 | 0.283 | 0.331 |
| Jun | 0.201 | 0.239 | 0.283 | 0.328 |
| Jul | 0.2 | 0.242 | 0.277 | 0.326 |
| Aug | 0.206 | 0.246 | 0.292 | 0.328 |
| Sep | 0.214 | 0.253 | 0.295 | 0.333 |
| Oct | 0.217 | 0.259 | 0.298 | 0.325 |
| Nov | 0.218 | 0.256 | 0.294 | |
| Dec | 0.211 | 0.256 | 0.282 | |

Following the formula 1, 2, 3 of linear exponential smoothing model and simple moving average method, where $\alpha$ =0.53, $S_1 = P_1$ and n=6, the forecasted results in the different periods are achieved, which is shown in Table 2 and Table3 separately.

**Table 2. The forecasted results in LESM (billion)**

| | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|
| Jan | | 0.213 | 0.256 | 0.288 |
| Feb | 0.179 | 0.222 | 0.261 | 0.3 |
| Mar | 0.18 | 0.21 | 0.248 | 0.292 |
| Apr | 0.19 | 0.22 | 0.262 | 0.297 |
| May | 0.195 | 0.231 | 0.272 | 0.311 |
| Jun | 0.199 | 0.237 | 0.278 | 0.322 |
| Jul | 0.2 | 0.238 | 0.281 | 0.325 |
| Aug | 0.2 | 0.24 | 0.279 | 0.326 |
| Sep | 0.203 | 0.243 | 0.286 | 0.327 |
| Oct | 0.209 | 0.248 | 0.291 | 0.33 |
| Nov | 0.213 | 0.254 | 0.295 | |
| Dec | 0.216 | 0.255 | 0.294 | |

**Table 3. The forecasted results in SMA (billion)**

| | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|
| Jan | | 0.211 | 0.252 | 0.29 |
| Feb | | 0.216 | 0.256 | 0.295 |
| Mar | | 0.215 | 0.254 | 0.294 |
| Apr | | 0.217 | 0.258 | 0.295 |
| May | | 0.221 | 0.262 | 0.3 |
| Jun | | 0.225 | 0.266 | 0.306 |
| Jul | 0.193 | 0.23 | 0.271 | 0.313 |
| Aug | 0.197 | 0.232 | 0.273 | 0.316 |
| Sep | 0.201 | 0.24 | 0.282 | 0.323 |
| Oct | 0.204 | 0.244 | 0.285 | 0.328 |
| Nov | 0.207 | 0.247 | 0.288 | |
| Dec | 0.209 | 0.249 | 0.29 | |

Following the formula 4, 5, 6, where m=3, g=3, k=67 and the functions in the hidden and output layer are adopted as hyperbolic tangent sigmoid transfer function (tansig) and log sigmoid transfer function (logsig) separately, the forecasted results for Aug, Sep, Oct in 2007 can be achieved {0.327, 0.337, 0.318}. The training procedure in MATLAB is shown in Figure 3

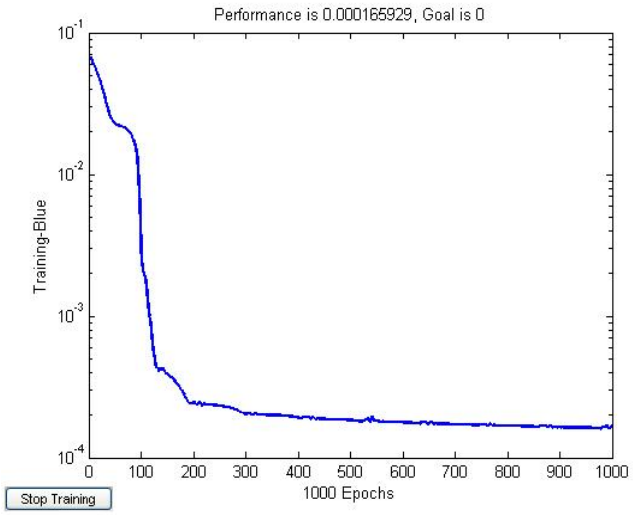


**Figure 3. The training procedure in MATLAB**

The absolute errors between forecasted results and real values are shown in Table 4. In contrast with some other methods, that new kind of nonlinear combination method has the lowest absolute mean value of errors, which reflects the higher forecasting accuracy.

**Table 4. The absolute errors between forecasted results and real values**

| | Aug,2007 | Sep,2007 | Oct,2007 | Mean value of absolute errors |
|---|---|---|---|---|
| Method in this paper | 0.001 | 0.004 | 0.007 | 0. 0040 |
| Index-smooth method | 0.002 | 0.006 | 0.005 | 0. 0043 |
| Moving average method | 0.012 | 0. 010 | 0. 003 | 0. 0083 |
| Grey forecasting model | 0.008 | 0.007 | 0.020 | 0. 0117 |

## 6. Conclusions

This paper presents a new kind of nonlinear combination method to predict the port throughput. The method possesses several characters as follows: (1) the short, medium and long-term influences are considered comprehensively; (2) the different singular methods are nonlinearly combined according to their characteristics; (3) Some original data is taken as the inputs of network with the forecasted values; (4) It has a higher forecasting accuracy.

## 7 Acknowledgements

The authors would like to thank peer reviewers for commenting this article. This work is supported by Science Foundation of Ministry of Education of China (Grant NO. 20070151022) and Research Project of the Educational Department of Liaoning Province (Leading Laboratory Project) under Grant 20060083.

*References*

[1]. John Mangan, Chandra Lalwani, Bernard Gardner. “Modelling Port/Ferry Choice in RoRo freight transportation”. Transport Management, Vol 1, 2002, 15-28.

[2]. An Dong, Cheng Zu-de. Application of Gray System Theory in the Logistics Forecasting of Shanghai Port. Logistics Sci-Tech, 2007, (3), 68-70. (In Chinese).

[3]. Elliott, G., A. Timmermann. Optimal Forecast Combinations Under General Loss Functions and Forecast Error Distributions. Journal of Econometrics. 2004, 122(1), 47-79.

[4]. Voort M.,Dougtherty M. Combining Kohonen maps with ARIMA Time Series Models to Forecast Traffic Flow. Transpotation Research. 1996, 4(5), 307-318.

[5]. Rui Zhiyuan, Ren Lina. Load Forecasting Model for Gansu Electric Power Network Based on Elman Neural Network. Modern Electric Power, 2007, 24(2), 26-29. (In Chinese)