International Journal of Operations Research

International Journal of Operations Research Vol. 10, No. 4, 161−170 (2013)

Forecasting Accuracy by Neural Network Applying to the Airlines Passengers and Cargo Data-Comparison with ARIMA Model

KazuhiroTakeyasu[1,*], Yuki Higuchi[2], Yuta Tsuchida[3] and Tatsuhiro Kuroda[3]

[1]College of Business Administration Tokoha University, Shizuoka, Japan.

[2]Faculty of Business Administration Setsunan University, Neyagawa City, Japan.

[3]Osaka Prefecture University, Osaka, Japan.

Received August 2013; Revised November 2013; Accepted December 2013

Abstract —In industry, making a correct forecasting is a very important matter. If the correct forecasting is not executed, there arise a lot of stocks and/or it also causes lack of goods. Time series analysis, neural networks and other methods are applied to this problem. In this paper, neural network is applied and Multilayer perceptron Algorithm is newly developed. The method is applied to the Airlines Passengers and Cargo Data. When there is a big change of the data, the neural networks cannot learn the past data properly, therefore we have devised a new method to cope with this. Repeating the data into plural section, smooth change is established and we could make a neural network learn more smoothly. Thus, we have obtained good results. The result is compared with ARIMA model. We have obtained the good results.

Keywords —forecasting, neural networks, time series analysis

1. INTRODUCTION

In industry, how to make a correct forecasting such as sales forecasting is a very important issue. If the correct forecasting is not executed, there arise a lot of stocks and/or it also causes lack of goods. Time series analysis, neural networks and other methods are applied to this problem. There are some related researches made on this. Reviewing past researches, Kimura *et al.* (1993) applied neural networks to demand forecasting and adaptive forecasting method was proposed. Baba and Suto (2000) combined neural networks and the temporal difference learning method to construct an intelligent decision support system for dealing stocks. Takeyasu *et al.* (2009) devised a new trend removing method and imbedded a theoretical solution of exponential smoothing constant. As a whole, it can be said that an application to sales forecasting is rather a few. In this paper, neural network is applied and Multilayer perceptron Algorithm is newly developed. The method is applied to the Airlines Passengers and Cargo Data. When there is a big change of the data, the neural networks cannot learn the past data properly, therefore we have devised a new method to cope with this. Repeating the data into plural section, smooth change is established and we could make a neural network learn more smoothly. Thus, we have obtained good results. The result is compared with ARIMA model.

The rest of the paper is organized as follows. In section 2, the method for neural networks is stated. An application method to the time series is introduced in section 3. In section 4, a new method is proposed to handle the rapidly changing data. Numerical example is stated in section 5. ARIMA model is stated and compared in section 6, which is followed by the remarks of section 7.

2. THE METHOD FOR NEURAL NETWORKS

In this section, outline of multilayered neural networks and learning method are stated. In figure l, multilayered neural network model is exhibited. It shows that it consist of input layer, hidden layer and output layer of feed forward type. Neurons are put on hidden layer and output layer. Neurons receive plural input and make one output.

Now, suppose that input layer have input signals $x_i (i = 1, 2, \ldots, l)$, hidden layer has m neurons and output layer has n neurons. Output of hidden layer $y_j (j = 1, 2, \ldots, m)$ is calculated as follows. Here $x_0 = -1$ is a threshold of hidden layer.

$$y_j = f\left(\sum_{i=0}^{l} v_{ij} x_i\right) \tag{1}$$

* Corresponding author's email: takeyasu@fj.tokoha-u.ac.jp

$$f\left(x\right)=\frac{1}{1+\exp(-x)} \tag{2}$$

When v_{ij} is a weighting parameter from input layer to hidden layer and (2) is a sigmoid function. $y_0=-1$ is a threshold of output layer and has the same value in all patterns. The value of the neuron of output layer, $z_k(k=1,2,\ldots,n)$ which is a final output of network, is expressed as follows.

$$z_k=f\left(\sum_{j=0}^{m}w_{jk}y_j\right) \tag{3}$$

When w_{jk} is a weighting parameter of Hidden layer through Output layer, Learning is executed such that v,w is updated by minimizing the square of "output – supervisor signal" Evaluation function is shown as follows.

$$E=\frac{1}{2}\sum_{k=0}^{n}(d_k-z_k)^2 \tag{4}$$

where d_k is a supervisor signal. Error signal is calculated as follows.

$$e_k=d_k-z_k \tag{5}$$

Δw_{jk} (Output layer) is calculated as follows.

$$\delta_k=e_kz_k(1-z_k) \tag{6}$$

$$\Delta w_{jk}=\eta y_j\delta_k \tag{7}$$

Therefore, weighting coefficient is updated as follows.

$$w_{jk}^{\text{new}}=w_{jk}^{\text{old}}+\Delta w_{jk} \tag{8}$$

where η is a learning rate.

Δv_{ij} (Hidden layer) is calculated as follows.

$$\gamma_j=y_j(1-y_j)\sum_{k=1}^{n}w_{jk}^{\text{new}}\delta_k \tag{9}$$

$$\Delta v_{ij}=\eta x_i\gamma_j \tag{10}$$

v_{ij} is updated as follows.

$$v_{ij}^{\text{new}}=v_{ij}^{\text{old}}+\Delta v_{ij} \tag{11}$$

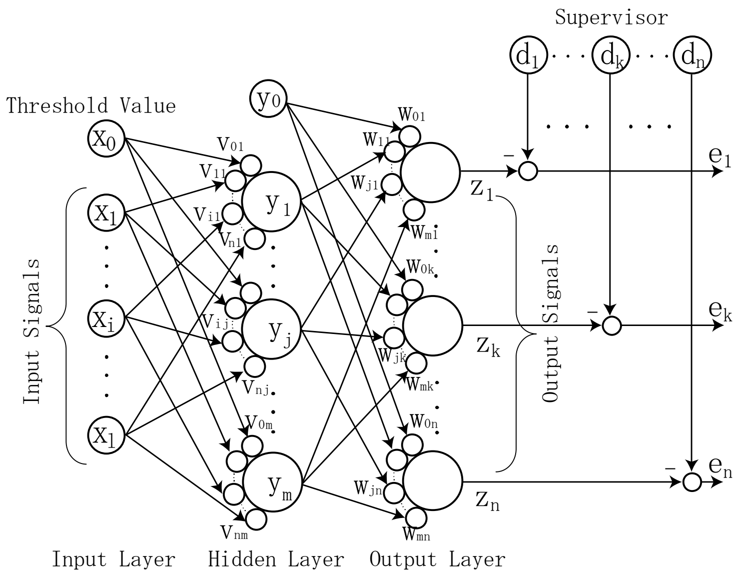

Figure 1. Multilayered neural network

3. AN APPLICATION METHOD TO THE TIME SERIES

Now we apply neural networks to the forecasting of time series. Suppose there are M months' time series data. We use them as follows: Latter half N months' data for test, the first half $(M-N)$ months' data for learning.

3.1 Normalization

Data should be normalized because output is controlled by sigmoid function. We use time series this time, therefore data is normalized in the range [0:1]. We obtained $\max, ?\ \min$ from 1 through $(M-N)$ months, which is a learning data period. We cannot grasp the test data range as the time of learning. Therefore estimated values $\widehat{\max}$, $\widehat{\min}$ are calculated as follows,

$$\widehat{\max} = \max \cdot \mu_{\max} \tag{12}$$

$$\widehat{\min} = \frac{\min}{\mu_{\min}} \tag{13}$$

where $\mu_{\max}$, $\mu_{\min}$ are margin parameters. Set a_k as time series data, then a_k is normalized as follows.

$$X_k = \frac{a_k - \widehat{\min}}{\widehat{\max} - \widehat{\min}} \tag{14}$$

3.2 Forecasting Method

Forecasting is executed as follows.

$$\hat{X}_k = F(X_{(k-l)}, X_{(k-l+1)}, \ldots, X_{(k-l+i)}, \ldots, X_{(k-1)}) \tag{15}$$

Where $F(x)$ is a neural network and X_k is a k th month's data (input signal). The number of learning patterns is $\left(M-N\right)-l$. We vary l as $l = 1, 2, \ldots, \left(M-N\right)/2$. The relation of learning data and supervisor data is shown as Figure 2. In this figure, input data is shown by the broken line when X_8 is targeted for learning under $l = 4$. Learning is executed recursively so as to minimize the square of $\hat{X}_k - X_k$, where $\hat{X}_k$ is an output.

$$\left(\hat{X}_k - X_k\right)^2 \to \varepsilon \tag{16}$$

This time, ε is not set as a stopping condition of iteration, but predetermined s steps are adopted for the stopping condition. Forecasted data $\hat{a}_k$ is reversely converted to Eq. (17) from Eq. (14) as follows.

$$\hat{a}_k = \hat{X}_k\left(\widehat{\max} - \widehat{\min}\right) + \widehat{\min} \tag{17}$$

3.3 Forecasting Accuracy

Forecasting accuracy is measured by the following "Forecasting Accuracy Ratio (FAR)".

$$\mathrm{FAR} = \left\{1 - \frac{\sum_{k=M-N}^{N} | a_k - \hat{a}_k |}{\sum_{k=M-N}^{N} a_k}\right\} \cdot 100 \tag{18}$$

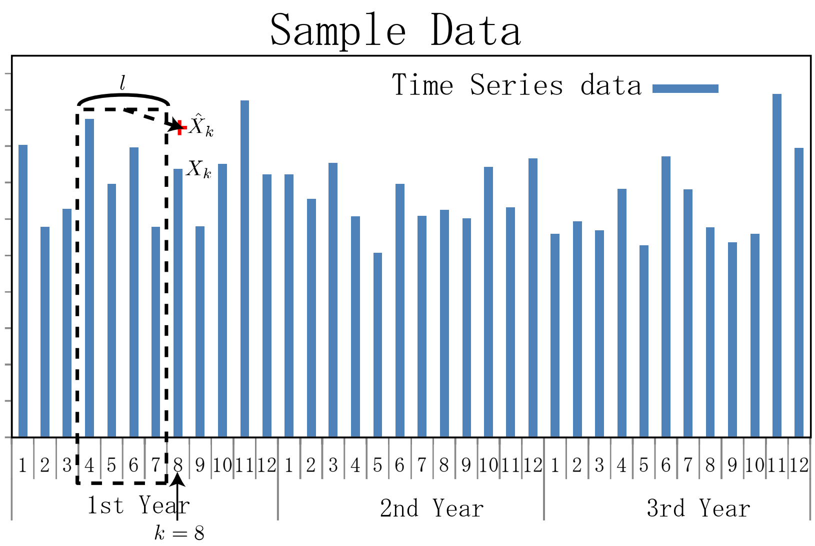

Figure 2. Choose the input data and supervisor for neural network (ex: $l = 4, k = 8$)

4. A NEWLY PROPOSED METHOD

We have found that the mere application of neural networks does not bear good results when there is a big change of the data. Therefore we have devised a new method to cope with this. Repeating the data into plural section, we aim to make a neural network learn more smoothly. The concept of the change of data sampling is exhibited in Figure 3. Data is repeated τ times and after the learning, the value is taken average by τ in order to fit for the initial condition.

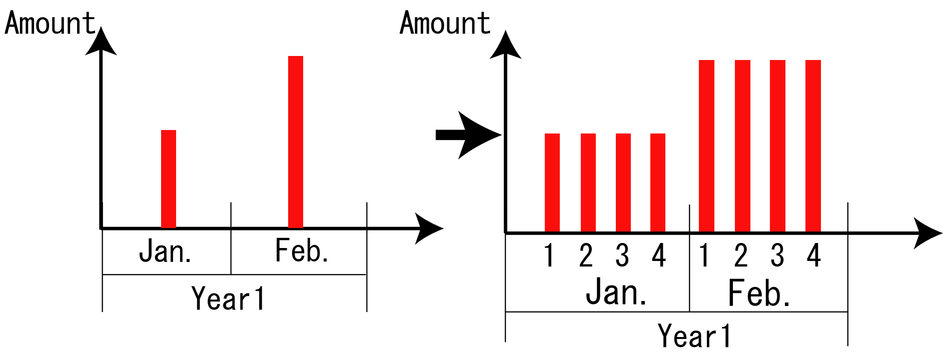

Figure 3. Change the time sampling (ex: $\tau = 4$)

5. NUMERICAL EXAMPLES

5.1 Used Data

The JAL passengers and cargo data for 4 cases from January 2009 to December 2011 were analyzed. These are the data of domestic passengers, domestic cargo, international passengers and international cargo. Here $M = 36$. Latter half data $N = 12$ are the data for test and the first half 24 data are the data for learning. $\mu_{\max}$ and $\mu_{\min}$ are set as follows

$$\mu_{\max} = 1.1 \tag{19}$$

$$\mu_{\min} = 1.5 \tag{20}$$

Each maximum, minimum and estimated maximum, minimum data are exhibited in Table 1.

Table1.The maximum value and the minimum value

	1 to 36 months	Estimated
	Maximum	
	Minimum	
Domestic Passenger	3,715,527	4,087,080
	1,871,991	1,570,351
Domestic Cargo	51,380	56,518
	27,625	21,207
International Passengers	1,039,603	1,143,563
	409,086	3,95,535
International Cargo	56,552	62,207
	17,242	13,322

5.2 Condition of Experiment

Condition of the neural network's experiment is exhibited in Table 2. Experiment is executed for 12 patterns $\left(l = 1, 2, \ldots, 12\right)$ and the Forecasting Accuracy Ratio is calculated based on the results.

Table2. The experiment of neural network

Name	Parameter	Value
The number of neurons in hidden layer	m	4
The number of output	n	1
The learning rate	η	0.035
Learning steps	s	4000

5.3 Experimental Results for τ=1 and τ=4

Now, we show the experimental results executed by the method stated in 3.2. The Forecasting Accuracy Ratio is exhibited in Table 3 and 4. Minimum score among 12 cases is written in bold for each case. In all cases, the case $\tau = 4$ is better than those of $\tau = 1$. Forecasting results for the minimum case of l are exhibited in Figures 4 through 7.

Table3. The result for Neural network [$\tau = 1$]

l	Domestic Psgrs	Domestic Cargo	International Psgrs	International Cargo
1	76.02	77.16	81.52	92.00
2	75.14	75.72	78.06	90.93
3	73.44	76.47	68.62	94.73
4	72.48	75.96	69.22	94.69
5	75.39	75.31	**84.52**	**<u>95.83</u>**
6	68.99	70.90	79.85	94.11
7	71.66	74.44	83.33	93.60
8	74.84	73.41	81.86	92.30
9	74.20	71.67	84.03	87.82
10	**76.15**	74.66	78.63	92.38
11	73.08	73.91	71.60	91.62
12	71.94	**77.29**	70.95	90.16

Table 4.The result for Neural network [$\tau = 4$]

l	Domestic Psgrs	Domestic Cargo	International Psgrs	International Cargo
1	94.26	**91.46**	94.78	96.26
2	**94.71**	88.30	92.52	94.52
3	92.09	91.42	91.35	96.21
4	91.72	91.32	**95.47**	**97.19**
5	93.83	90.26	94.78	95.09
6	92.01	79.02	94.43	93.33
7	85.55	77.97	94.14	91.67
8	87.35	69.09	92.94	94.57
9	84.93	77.73	92.92	94.59
10	85.20	75.65	89.76	90.29
11	76.44	72.11	86.43	86.73
12	67.03	70.26	71.53	82.70

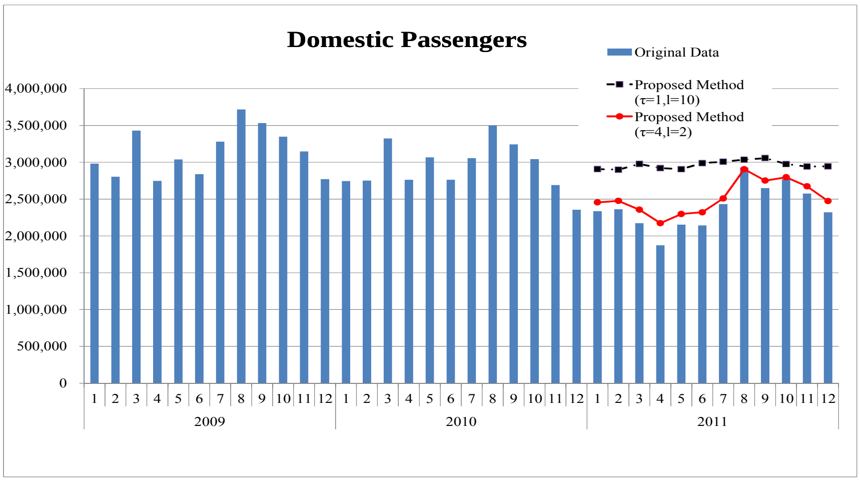

Figure 4.The result of Domestic Passengers ($\tau = 1, l = 10$) and ($\tau = 4, l = 2$)

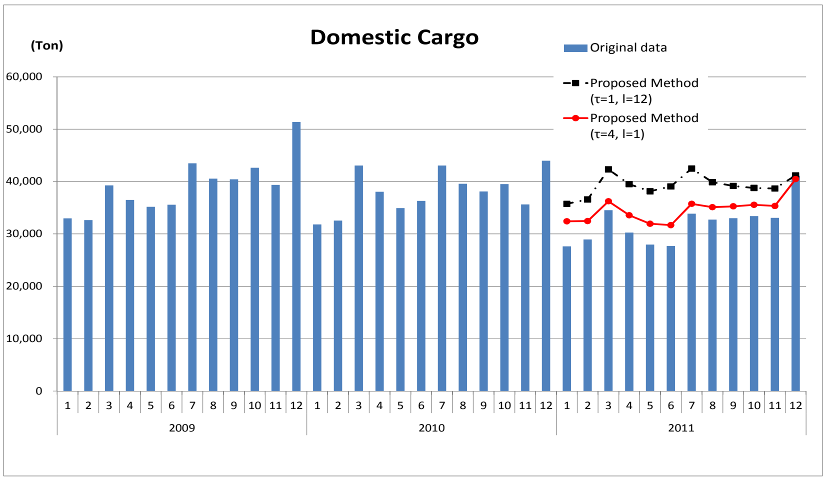

Figure 5. The result of Domestic Cargo ($\tau = 1, l = 12$) and ($\tau = 4, l = 1$)

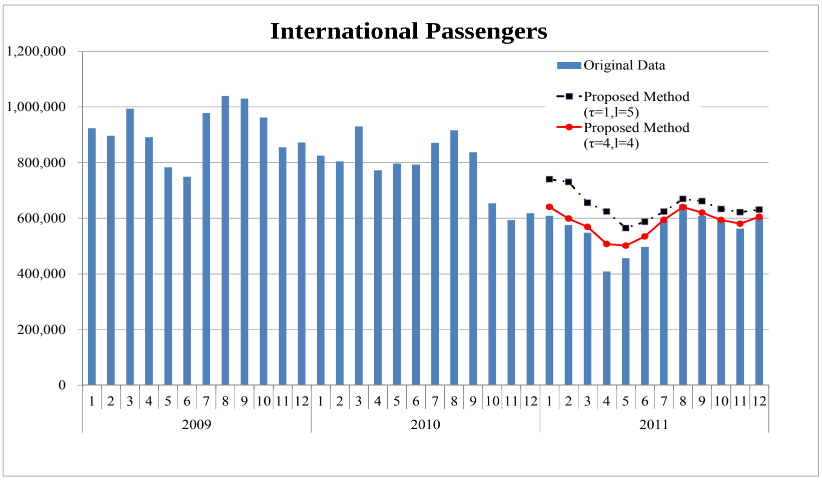

Figure 6. The result of International Passengers ($\tau = 1, l = 5$) and ($\tau = 4, l = 4$)

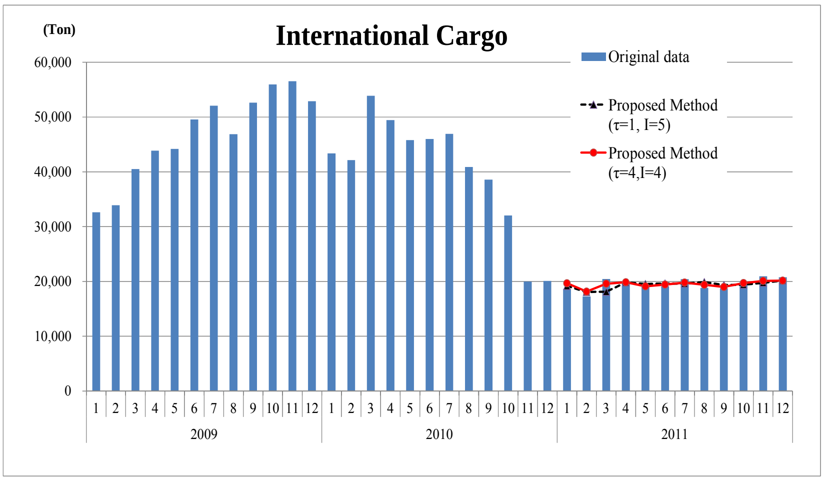

Figure 7. The result of International Cargo ($\tau = 1, l = 5$) and ($\tau = 4, l = 4$)

6. COMPARISON WITH ARIMA MODEL

6.1 ARIMA model

p - th order AR model is stated as

$$x_t + a_1 x_{t-1} + \cdots + a_p x_{t-p} = e_t \tag{21}$$

Using the delay operator z^{-1} which means

$$z^{-1} x_t = x_{t-1} \tag{22}$$

Define

$$A\left(z^{-1}\right) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_p z^{-p} \tag{23}$$

Then AR model is stated as

$$A\left(z^{-1}\right) x_t = e_t \tag{24}$$

q -th order MA model is also stated as

$$x_t = e_t + b_1 e_{t-1} + \cdots + b_q e_{t-q} \tag{25}$$

And define

$$B\left(z^{-1}\right) = 1 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_q z^{-q} \tag{26}$$

then MA model is stated as

$$x_t = B\left(z^{-1}\right) e_t \tag{27}$$

ARIMA model

$$x_t + a_1 x_{t-1} + \cdots a_p x_{t-p} = e_t + b_1 e_{t-1} + \cdots + b_q e_{t-q} \tag{28}$$

is stated as

$$A\left(z^{-1}\right) x_t = B\left(z^{-1}\right) e_t \tag{29}$$

(p, d, q) order ARIMA model of d times difference from the original data is stated as

$$A\left(z^{-1}\right)\left(1 - z^{-1}\right)^d x_t = B\left(z^{-1}\right) e_t \tag{30}$$

The order of ARIMA model is determined by calculating AIC. AIC is calculated as follows.

$$AIC = -2 \operatorname{Ln} L + 2\left(p + d + q\right) \tag{31}$$

where L is a Log of Likelihood Function.

6.2 Forecasting results

Forecasting results which are compared with former results are exhibited in Figure 8 through 11. Forecasting Accuracy Ratio is exhibited in Table 5. The order of ARIMA model is exhibited in Table 6.Comparison of the both results is exhibited in Table 7.

Table 5. The result for ARIMA: Forecasting Accuracy Ratio

	Domestic		International	
	Psgrs	Cargo	Psgrs	Cargo
ARIMA	90.84	84.65	91.14	89.87

Table 6. The order of ARIMA model

	Domestic		International	
	Psgrs	Cargo	Psgrs	Cargo
ARIMA	0,2,4	0,2,2	0,2,1	0,2,1

Table 7. Comparison of the both results

	Forecasting Accuracy Ratio	
	Previous Method (ARIMA)	Proposed Method
Domestic Passengers	90.84	**94.71** $(\tau = 4, l = 2)$
Domestic Cargo	84.65	**91.46** $(\tau = 4, l = 2)$
International Passengers	91.14	**95.47** $(\tau = 4, l = 4)$
International Cargo	89.87	**97.19** $(\tau = 4, l = 4)$

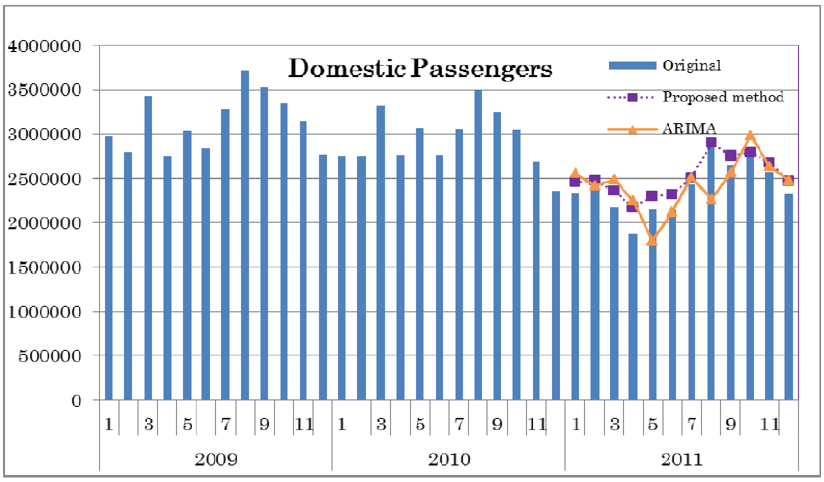

Figure 8.Comparison of the both results-The result of Domestic passengers

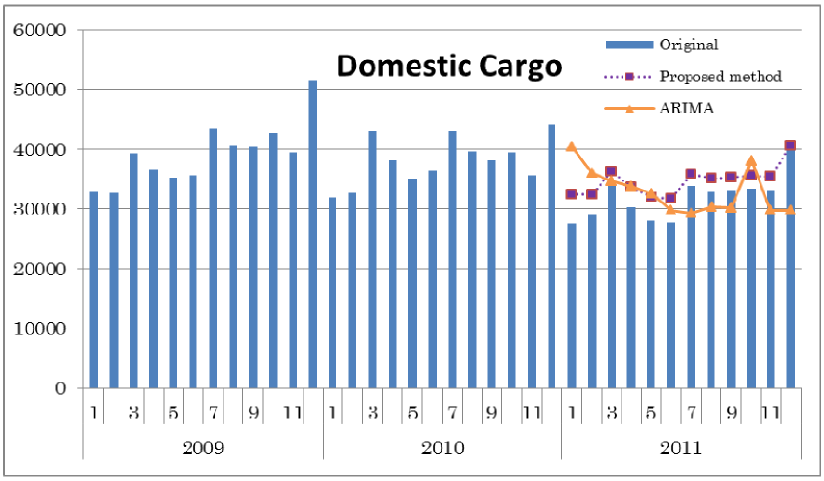

Figure 9. Comparison of the both results- The result of Domestic Cargo

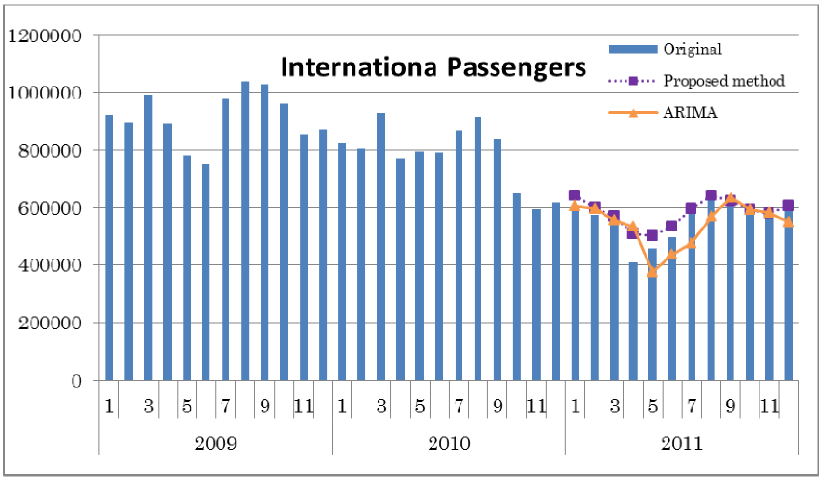

Figure 10. Comparison of the both results- The result of International Passengers

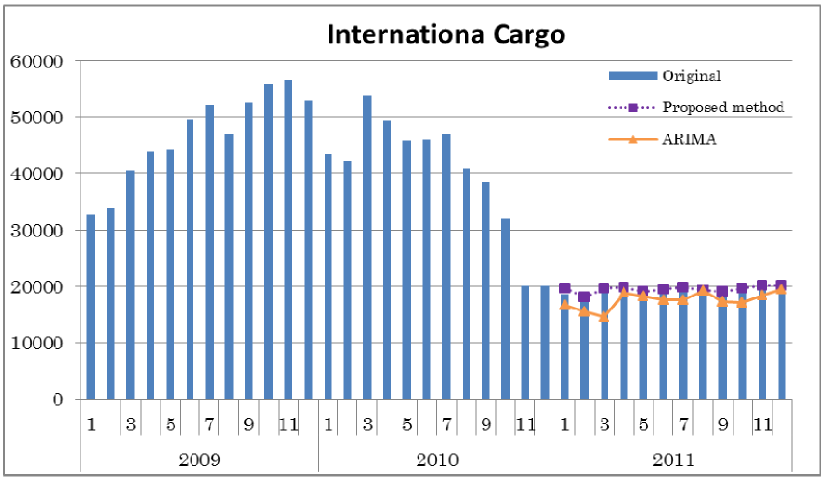

Figure 11. Comparison of the both results- The result of International Cargo

7. REMARKS

Now, we compare with both results. In Table 7, both results are stated and compared. Their comparison is shown in Figure 8, 9, 10 and 11. In all cases, this newly proposed method had a better forecasting accuracy than ARIMA model.

8. CONCLUSION

In industry, making a correct forecasting is a very important matter. If the correct forecasting is not executed, there arise a lot of stocks and/or it also causes lack of goods. Time series analysis, neural networks and other methods are applied to this problem. In this paper, neural network is applied and Multilayer perceptron Algorithm is newly developed. The method is applied to the Airlines Passengers and Cargo Data. When there is a big change of the data, the neural networks cannot learn the past data properly, therefore we have devised a new method to cope with this. Repeating the data into plural section, smooth change is established and we could make a neural network learn more smoothly. Thus, we have obtained good results. The result is compared with ARIMA model. We have obtained the good results. In the numerical example, all cases had a better forecasting accuracy than ARIMA model. Various cases should be examined hereafter.

REFERENCES

1. Baba, N. and Suto, H. (2000). Utilization of artificial neural networks and the TD-learning method for constructing intelligent decision support systems. *European Journal of Operational Research*, 122(2): 501 — 508.
2. Kimura, A., Arizono, I., and Ohta, H. (1993). An application of layered neural networks to demand forecasting. *Journal of Japan Industrial Management Association*, 44(5): 401 — 407.
3. Sanchez-Sinencio, E. (1992). *Artificial Neural Networks: Paradigms, Applications, and Hardware Implementations*. IEEE Press.
4. Takeyasu, K., Imura, K., and Higuchi, Y. (2009). Estimation of smoothing constant of minimum variance and its application to shipping data with trend removal method. *Industrial Engineering and Management Systems*, 8(4): 257-263.