Journal of Computations & Modelling, vol.7, no.4, 2017, 39-62
ISSN: 1792-7625 (print), 1792-8850 (online)
Scienpress Ltd, 2017

The Method to improve Forecasting Accuracy by Using Neural Network –An Application to the Production Data of Udon Noodles–

Yuki Higuchi[1], Yuta Tsuchida[2] and Kazuhiro Takeyasu[3]

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 production data of Udon Noodles. 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

[1] Faculty of Business Administration, Setsunan University.
E-mail: y-higuch@kjo.setsunan.ac.jp
[2] Osaka Prefecture University. E-mail: soil.y.paddy@gmail.com
[3] College of Business Administration, Tokoha University.
E-mail: takeyasu@fj.tokoha-u.ac.jp

Article Info: *Received:* July 19, 2017. *Revised:* August 21, 2017.
Published online : October 31, 2017.

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 the method we have developed before as well as ARIMA model. We have obtained the good results.

Mathematics Subject Classification: 91B70

Keywords: forecasting, neural network, time series analysis, ARIMA model

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)[1] applied neural networks to demand forecasting and adaptive forecasting method was proposed. Baba et al. (2000) [2] combined neural networks and the temporal difference learning method to construct an intelligent decision support system for dealing stocks. Takeyasu et al. (2009)[3] 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 production data of Udon Noodles. 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 the method we have developed before (Takeyasu et al. (2012)[4]) as well as 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. Past experimental data are stated and compared in section 6, which is followed by the remarks of section 7.

2 The method for Neural Networks[5]

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.

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

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

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.

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

$$\Delta \mathrm{w}_{jk} = \eta y_j \delta_k \tag{7}$$

Therefore, weighting coefficient is updated as follows.

$$w_{jk}^{\mathrm{new}} = w_{jk}^{\mathrm{old}} + \Delta \mathrm{w}_{jk} \tag{8}$$

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.

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

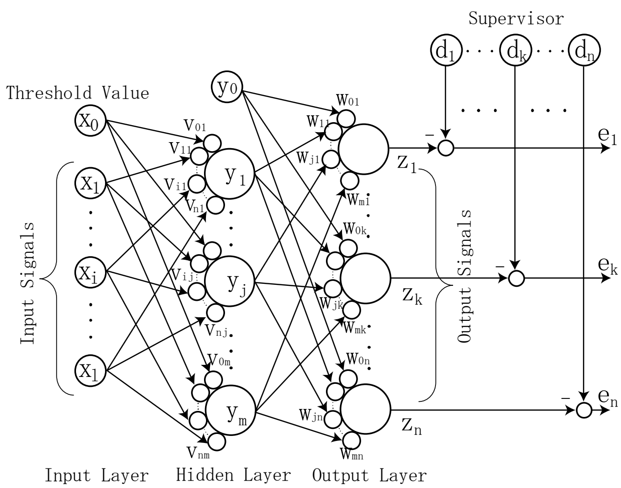

Figure 1: Multilayered neural network

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.

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

Where $F(x)$ is a neural network and X_k is a kth month's data (input signal). The number of learning patterns is $(M - N) - l$. We vary l as $l = 1,2, \dots, (M - N)/2$. 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.

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

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.

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

3.3 Forecasting Accuracy

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

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

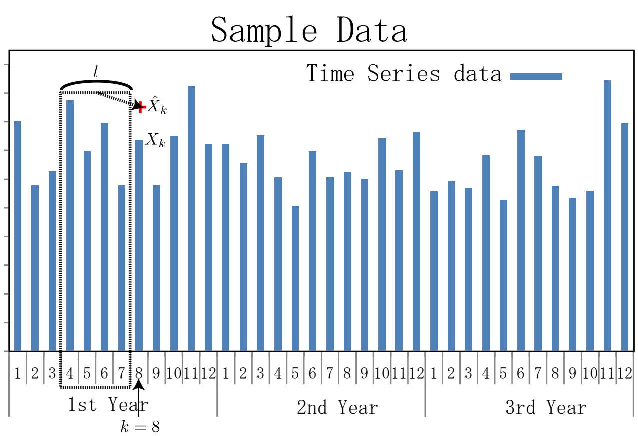

Figure 2: Choose the input data and supervisor for neural network

(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.

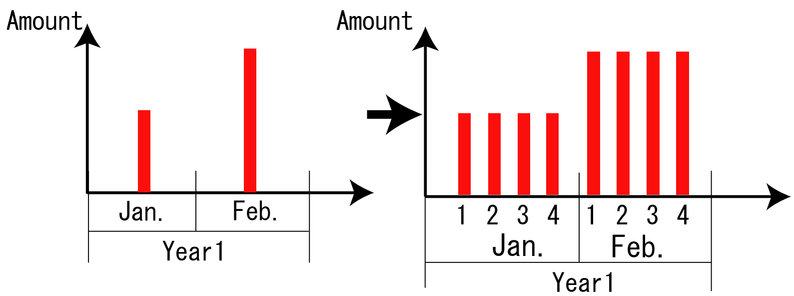

Figure 3: Change the time sampling (ex: $\tau = 4$)

5 Numerical Example

5.1 Used Data

The production data of Udon Noodles for 2 cases from January 2007 to December 2009 are analyzed. Here $M = 36$. First of all, graphical charts of these time series data are exhibited in Figure 4, 5 and 6. 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.

Latter half data $N = 15$ are the data for test and the first half 35 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.

Table 1: The maximum value and the minimum value

	1 to 36 months	Estimated
	Maximum	
	Minimum	
Raw Udon	3489	3838
Noodle	1910	1441
Boiled Udon	21080	23188
Noodle	11964	8519

5.2 Condition of Experiment

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

Table 2: The experiment of neural network

Name	Parameter	Value
The number of neurons in hidden layer	m	16
The number of output	n	1
The learning rate	η	0.035
Learning steps	s	4000

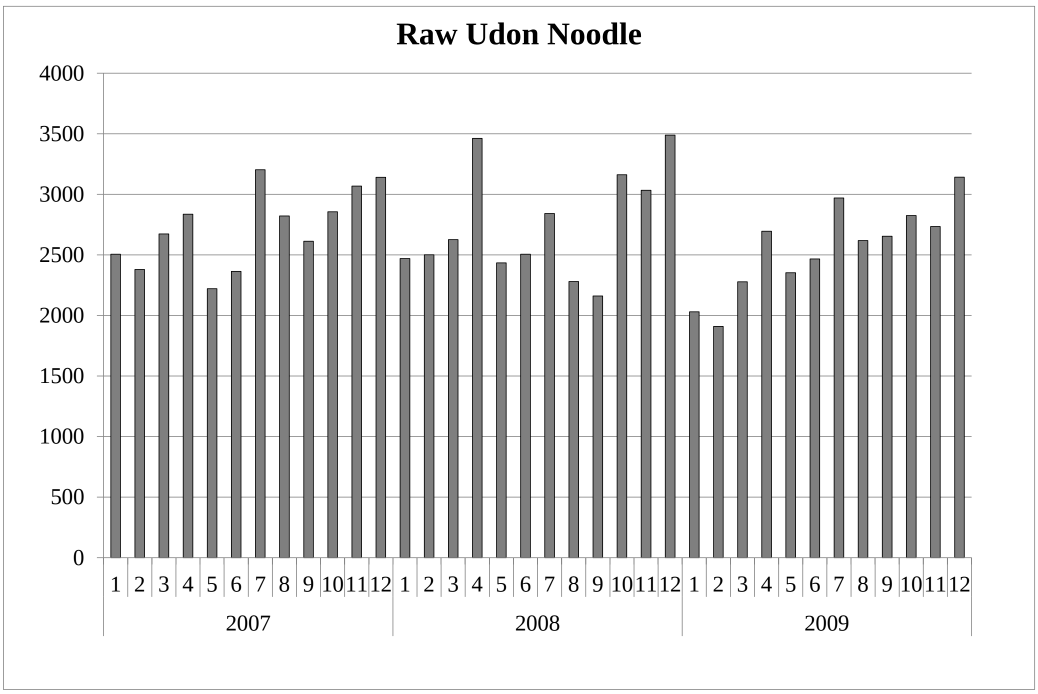

Figure 4: Data of Raw Udon Noodle

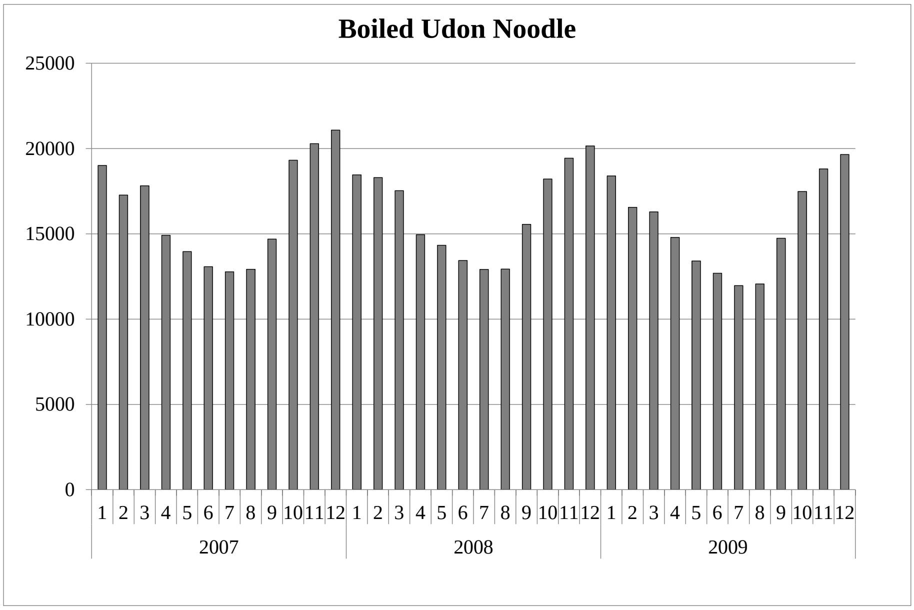

Figure 5: Data of Boiled Udon Noodle

5.3 Experimental results for $\tau = 1$ and $\tau = 4$

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 7 through 9.

Table 3: The result for Neural network [$\tau = 1$]

l	Raw Udon Noodle	Boiled Udon Noodle
1	**88.11**	91.07
2	87.59	93.28

3	87.38	**94.43**
4	87.33	93.65
5	86.59	93.91
6	85.12	93.87
7	86.43	93.24
8	86.36	92.43
9	86.42	91.14
10	85.16	92.43
11	85.79	93.28
12	86.70	93.52

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

l	Raw Udon Noodle	Boiled Udon Noodle
1	**95.51**	97.39
2	94.78	**98.09**
3	95.22	97.30
4	94.86	96.99
5	93.25	96.80
6	91.64	95.51
7	92.25	96.42
8	90.79	94.21
9	90.23	92.10
10	86.89	94.48

11	86.17	94.18
12	76.72	93.34

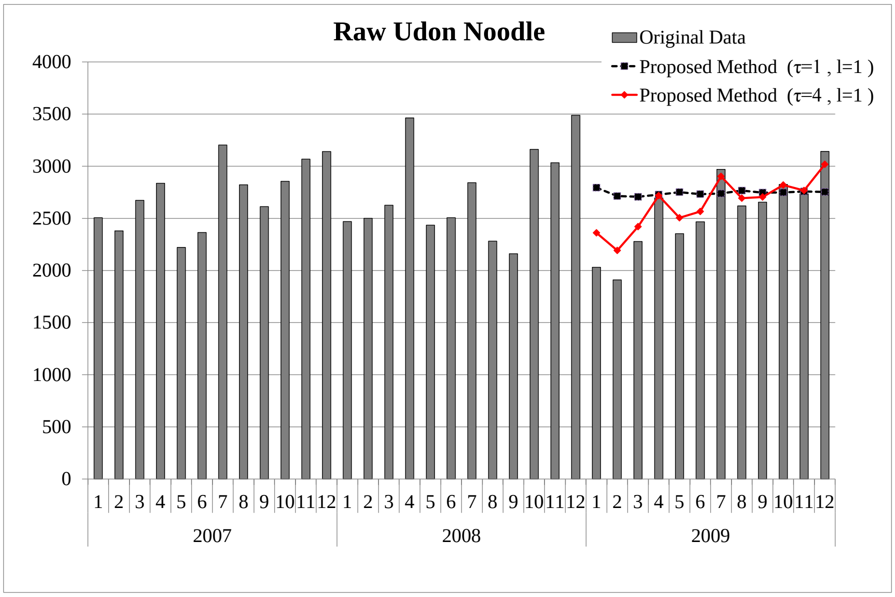

Figure 6: The result of Raw Udon Noodle ($\tau = 1, l = 1$) and ($\tau = 4, l = 1$)

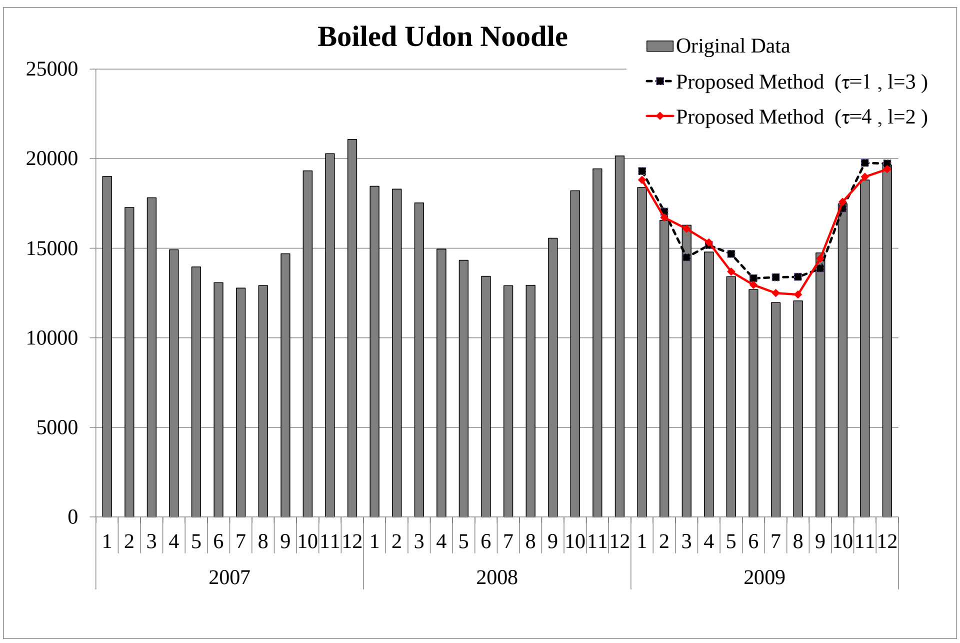

Figure 7: The result of Boiled Udon Noodle ($\tau = 1, l = 3$) and ($\tau = 4, l = 2$)

6 Past Experimental Data and Its Comparison

6.1 Outline of the Method

Focusing that the equation of exponential smoothing method(ESM) is equivalent to (1,1) order ARMA model equation, a new method of estimation of smoothing constant in exponential smoothing method is proposed before by us which satisfies minimum variance of forecasting error. Generally, smoothing constant is selected arbitrary. But in this paper, we utilize above stated theoretical solution. Firstly, we make estimation of ARMA model parameter and then estimate smoothing constants. Thus theoretical solution is derived in a simple way and it may be utilized in various fields. Furthermore, combining the trend removing method with this method, we aim to improve forecasting accuracy. An approach to this method is executed in the following method. Trend removing by the combination of linear and 2nd order non-linear function and 3rd order non-linear function is executed to the production data of Udon Noodles. Genetic Algorithm is utilized to search optimal weights for the weighting parameters of linear and non-linear function. For the comparison, monthly trend is removed after that. Theoretical solution of smoothing constant of ESM is calculated for both of the monthly trend removing data and the non-monthly trend removing data.

6.2 Theoretical Solution of Smoothing Constant in ESM

In ESM, forecasting at time $t-1$ is stated in the following equation.

$$\hat{a}_{t+1} = \hat{a}_t + \alpha(a_t - \hat{a}_t) \tag{21}$$

$$= \alpha a_t + (1-\alpha)\hat{a}_t \tag{22}$$

Here,

$\hat{a}_{t+1}$: forecasting at $t+1$

a_t : realized value at t

α : smoothing constant $(0 < \alpha < 1)$

(22) is re-stated as

$$\hat{a}_{t+1} = \sum_{l=0}^{\infty} \alpha(1-\alpha)^l \, a_{t-1} \tag{23}$$

By the way, we consider the following (1,1) order ARMA model.

$$a_t - a_{t-1} = h_t - \beta \, h_{t-1} \tag{24}$$

Generally, (p,q) order ARMA model is stated as

$$a_t + \sum_{i=1}^{p} \kappa_i \, a_{t-i} = h_t + \sum_{j=1}^{q} \lambda_j h_{t-j} \tag{25}$$

Here,

$\{a_t\}$: Sample process of Stationary Ergodic Gaussian Process $t = 1,2,\ldots,N,\ldots$

$\{h_t\}$: Gaussian White Noise with 0 mean σ_h^2 variance MA process in (25) is supposed to satisfy convertibility condition.

Finally we get

$$\lambda_1 = \frac{1-\sqrt{1-4\rho_1^2}}{2\rho_1}$$

$$\alpha = \frac{1+2\rho_1-\sqrt{1-4\rho_1^2}}{2\rho_1} \tag{26}$$

where ρ_1 is an autocorrelation function of 1st order lag. This α is a theoretical solution of smoothing constant in ESM (in detail, see [3]).

6.3 Trend Removal Method

As ESM is a one of a linear model, forecasting accuracy for the time series with non-linear trend is not necessarily good. How to remove trend for the time series with non-linear trend is a big issue in improving forecasting accuracy. In this paper, we devise to remove this non-linear trend by utilizing non-linear function. As a trend removal method, we describe linear and non-linear function, and the combination of these.

[1] Linear function

We set:

$$r = b_{11}u + b_{12} \tag{27}$$

as a linear function, where u is a variable, for example, time and r is a variable, for example, shipping amount, b_{11} and b_{12} are parameters which are estimated by using least square method.

[2] Non-linear function

We set:

$$r = b_{21}u^2 + b_{22}u + b_{23} \tag{28}$$

$$r = b_{31}u^3 + b_{32}u^2 + b_{33}u + b_{34} \tag{29}$$

as a 2nd and a 3rd order non-linear function. (b_{21}, b_{22}, b_{23}) and

$(b_{31},\ b_{32}, b_{33},\ b_{34})$ are also parameters for a 2nd and a 3rd order non-linear functions which are estimated by using least square method.

[3] The combination of linear and non-linear function

We set:

$$r = \alpha_1 \left(b_{11}u + b_{12}\right) + \alpha_2 \left(b_{21}u^2 + b_{22}u + b_{23}\right) + \alpha_3 \left(b_{31}u^3 + b_{32}u^2 + b_{33}u + b_{34}\right) \tag{30}$$

$$0 \le \alpha_1 \le 1\ , 0 \le \alpha_2 \le 1\ , 0 \le \alpha_3 \le 1 \qquad \alpha_1 + \alpha_2 + \alpha_3 = 1 \tag{31}$$

as the combination of linear and 2nd order non-linear and 3rd order non-linear function. Trend is removed by dividing the data by (30). The optimal weighting parameters $\alpha_1,\ \alpha_2,\ \alpha_3$ are determined by utilizing Genetic Algorithm.

6.4 Monthly Ratio

For example, if there is the monthly data of L years as stated bellow:

$$\{x_{\theta\phi}\}\ (\theta = 1, \dots L)\ (\phi = 1, \dots, 12)$$

where, $x \in R$ in which θ means month and ϕ means year and $x_{\theta\phi}$ is a data of θ-th year, ϕ-th month. Then, monthly ratio $\tilde{x}_\phi(\phi = 1, \dots, 12)$ is calculated as follows.

$$\tilde{x}_\phi = \frac{\frac{1}{L}\sum_{\theta=1}^{L} x_{\theta\phi}}{\frac{1}{L}\cdot\frac{1}{12}\sum_{\theta=1}^{L}\sum_{\phi=1}^{12} x_{\theta\phi}} \tag{32}$$

Monthly trend is removed by dividing the data by (32). Numerical examples both of monthly trend removal case and non-removal case are discussed later.

6.5 Forecasting Results

Forecasting results are exhibited in Figure 10 through 12. Forecasting Accuracy Ratio is exhibited in Table 5.

6.6 Remarks

In all cases, that monthly ratio was used had a better forecasting accuracy. Raw Udon Noodle had a good result in the linear function model. Boiled Udon Noodle had a good result in the combination of linear and 2nd order non-linear function model.

The minimum variance of forecasting error of GA coincides with those of the calculation of all considerable cases and it shows the theoretical solution. Although it is a rather simple problem for GA, we can confirm the effectiveness of GA approach. Further study for complex problems should be examined hereafter.

Table 5: Forecasting Accuracy Ratio

	Monthly Ratio	
	Used	Not used
Raw Udon Noodle	**89.49**	87.23
Boiled Udon Noodle	**92.04**	90.41

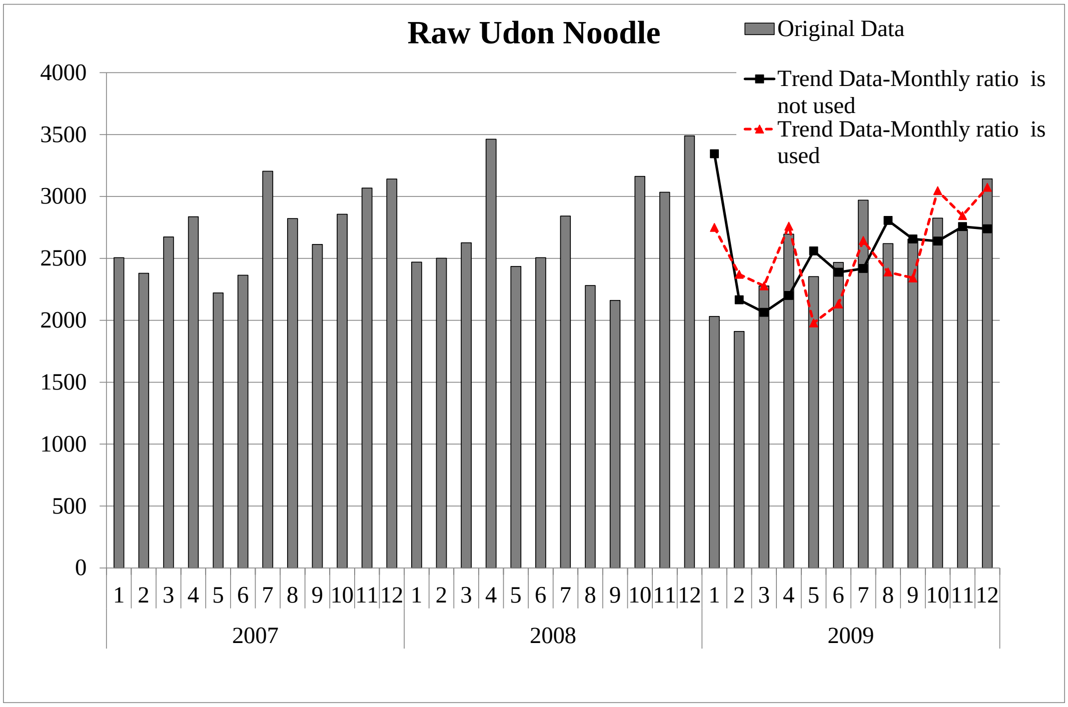

Figure 8: Forecasting Result of Raw Udon Noodle

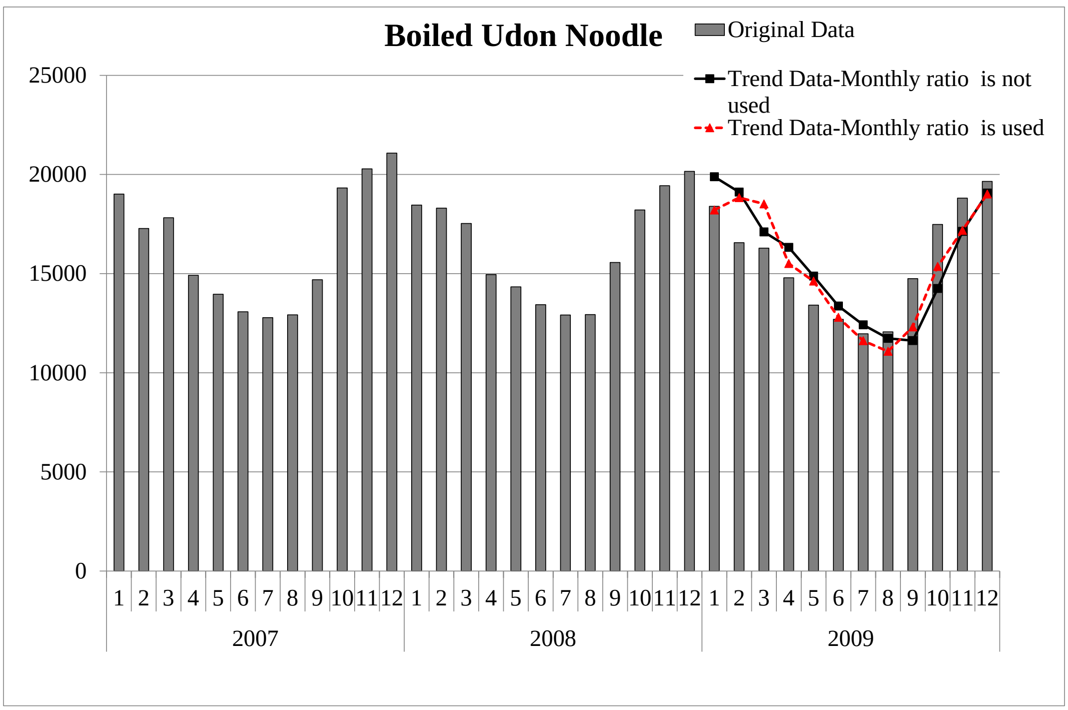

Figure 9: Forecasting Result of Boiled Udon Noodle

6.7 Forecasting Result by Using ARIMA model

The results of Forecasting Accuracy Ratio and the order of ARIMA model are exhibited in Table 6, 7.

Table 6: The result in Raw Udon Noodle: Forecasting Accuracy Ratio and its order

ARIMA	Raw Udon Noodle		
order	4,0,3	4,0,4	5,0,4
Forecasting Accuracy Ratio (One Pattern*)	**91.12**	90.67	90.84
Forecasting Accuracy Ratio (One Step**)	81.85	79.43	79.96

*Plural steps forecasting were executed at a time by using the same system parameters.

**One step forecasting was executed consecutively by using the former forecasting result.

Table 7: The result in Boiled Udon Noodle: Forecasting Accuracy Ratio and its order

ARIMA	Boiled Udon Noodle		
order	4,0,4	6,0,2	6,0,3
Forecasting Accuracy Ratio (One Pattern)	94.61	**94.71**	94.67
Forecasting Accuracy Ratio (One Step)	92.36	93.57	93.00

7 Remarks

Now, we compare these results with the past experimental results and those of

ARIMA Model in Table 8 and 9. Their comparison is shown in Fig. 10 and 11. In all cases, this newly proposed method had the best forecasting accuracy.

Table 8: The comparison result of Raw Udon Noodle

Forecasting Accuracy Ratio		
Previous Method	Monthly Ratio Not Used	87.23
	Monthly Ratio Used	89.49
Proposed Method	$\tau = 1, l = 1$	88.11
	$\tau = 4, l = 1$	**95.51**
ARIMA (One Pattern)	$p = 4, d = 0, q = 3$	91.12
	$p = 4, d = 0, q = 4$	90.67
	$p = 5, d = 0, q = 4$	90.84

Table 9: The comparison result of Boiled Udon Noodle

Forecasting Accuracy Ratio		
Previous Method	Monthly Ratio Not Used	90.41
	Monthly Ratio Used	92.04
Proposed Method	$\tau = 1, l = 3$	94.43
	$\tau = 4, l = 2$	**98.09**
ARIMA (One Pattern)	$p = 4, d = 0, q = 4$	94.61
	$p = 6, d = 0, q = 2$	94.71
	$p = 6, d = 0, q = 3$	94.67

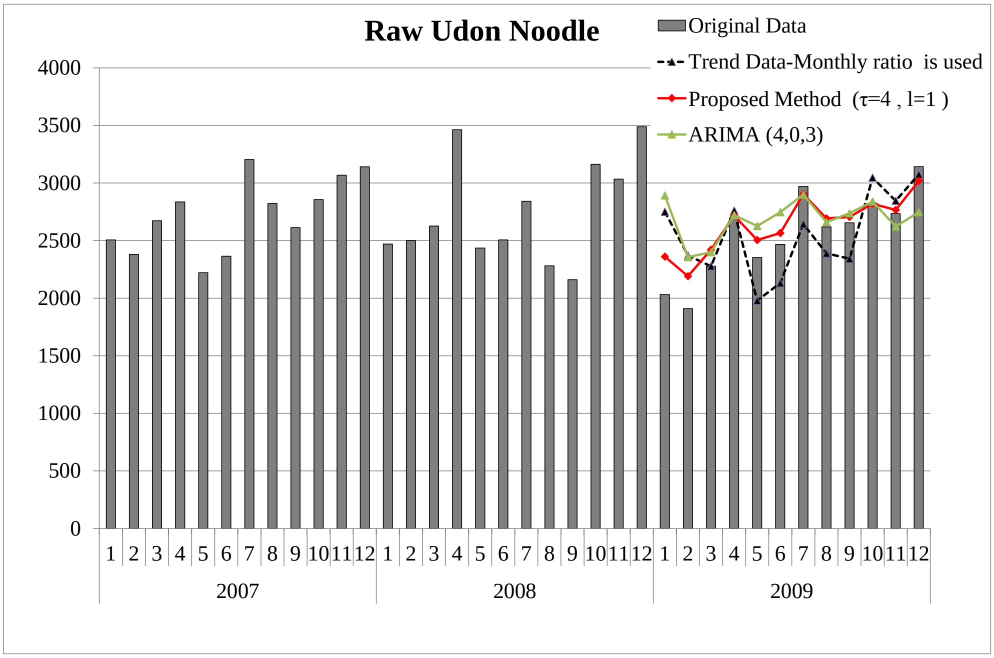

Figure 10: The result of Raw Udon Noodle: monthly ratio is used, ARIMA(4,0,3) and $(\tau = 4, l = 1)$

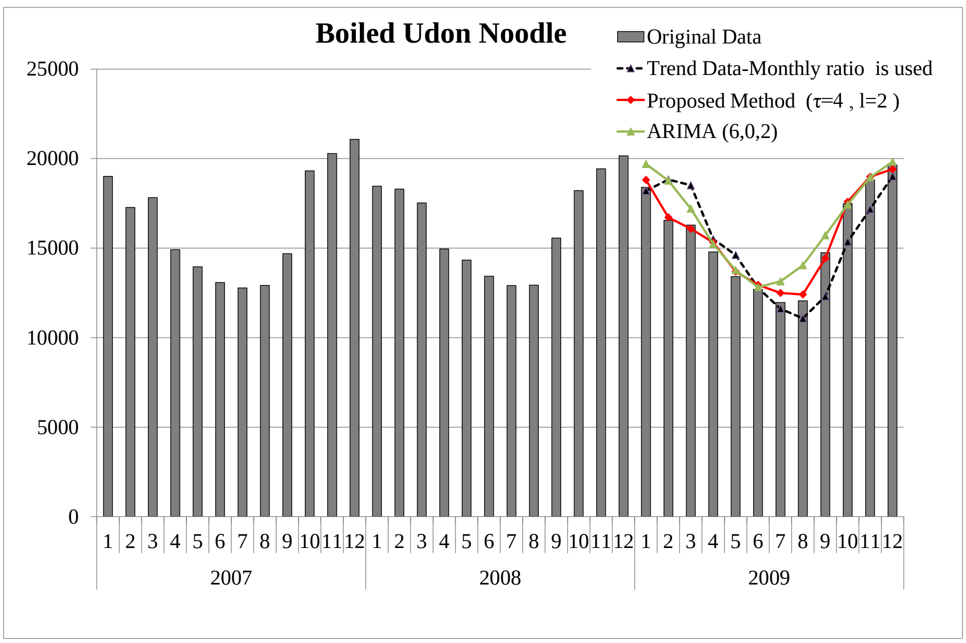

Figure 11: The result of Boiled Udon Noodle: monthly ratio is used, ARIMA(6,0,2) and $(\tau = 4, l = 2)$

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 production data of Udon Noodles. 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 was compared with the method we have developed before as well as ARIMA model. In the numerical example, all cases had the best forecasting accuracy. Thus, we have obtained the good results. Various cases should be examined hereafter.

References

[1] Aritoshi Kimura, Ikuo Arizono, Hiroshi Ohta, An Application of Layered neural networks to demand forecasting, *Japan Industrial Management Association*, **44**(5), (1993), 401-407.

[2] N. Baba, H. Suto, Utilization of artificial neural networks and the TD-learning method for constructing intelligent decision support systems, *European Journal of Operational Research*, **122**, (2000), 501-508.

[3] Kazuhiro Takeyasu, Keiko Imura, Yuki Higuchi, Estimation of Smoothing Constant of Minimum Variance And Its Application to Shipping Data With Trend Removal Method, *Industrial Engineering and Management Systems*, **8**(4), (2009), 257-263.

[4] Junpei Yamana, Yuki Higuchi, Kazuhiro Takeyasu, A Hybrid Method to Improve Forecasting Accuracy Utilizing Genetic Algorithm, The Twelfth Asia Pacific Industrial Engineering and Management Systems Conference October 14-17, 2011, Beijing, China.

[5] Edgar Sanchez-Sinencio, Clifford Lau, editors : Artificial neural networks: Paradigms, applications, and hardware implementations, IEEE press, Piscataway, N.J, (1992).