Journal of Computations & Modelling, vol.4, no.3, 2014, 115-137
ISSN: 1792-7625 (print), 1792-8850 (online)
Scienpress Ltd, 2014

The Method to improve Forecasting Accuracy by Using Neural Network

–An Application to the Shipping Data of Consumer Goods

Yuta Tsuchida[1], Yuki Higuchi[2],
Kazuhiro Takeyasu[3] and Michifumi Yoshioka[4]

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 original shipping data of consumer goods. 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

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

Article Info: *Received* : May 26, 2014. *Revised* : August 28, 2014.
Published online : September 10, 2014.

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]). We have obtained the good results for half cases in the total.

Mathematics Subject Classification : 92B20; 82C32; 91B70; 93B40; 97N40
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) [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 original shipping data of consumer goods. 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 for half cases in the total. The result is compared with the method we have developed before (Takeyasu et al. (2012) [4]).

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 $x_i(i = 1,2, \dots, l)$, hidden layer has m neurons and output layer has n neurons. Output of hidden layer $y_j(j = 1,2, \dots, m)$ 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, $z_k(k = 1,2, \dots, n)$ 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.

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

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

v_{ij} is updated as follows.

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

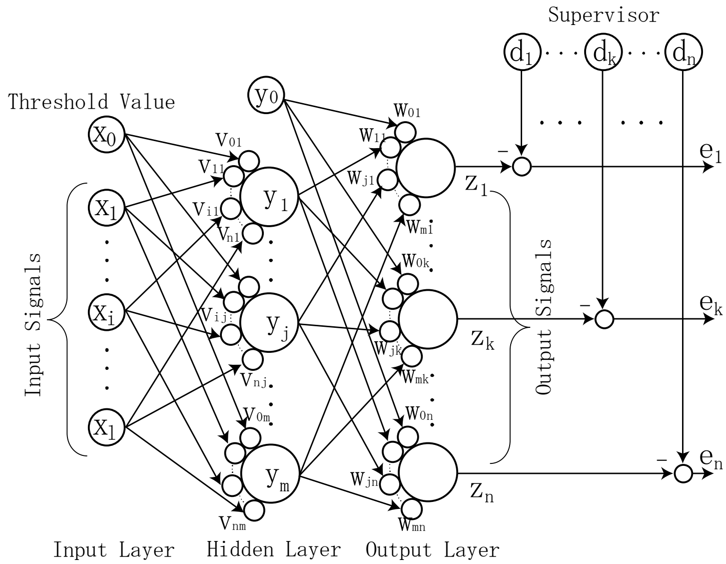

Figure 1: Multilayered neural network

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.

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

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.

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

3.3 Forecasting Accuracy

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

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

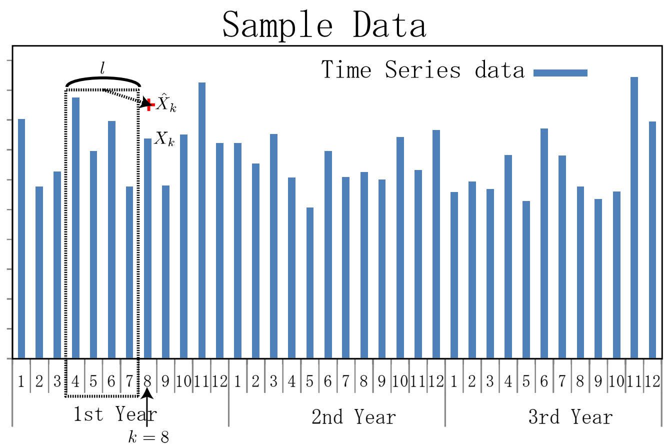

Figure 2: Choose the input data and supervisor for neural network (ex: l=4, k=8)

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 sections, 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.

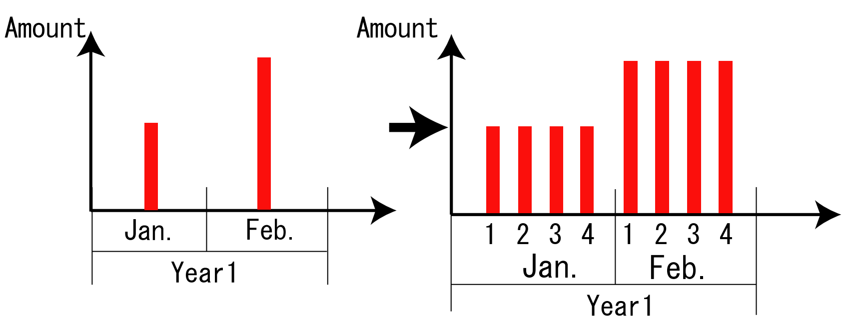

Figure 3: Change the time sampling (ex:τ=4)

5 Numerical Example

5.1 Used Data

The original shipping data of consumer goods for ten cases were analyzed. There are three years' monthly shipping data for each case. Here $M = 36$. First of all, graphical charts of these time series data are exhibited in Figure 4 through 13. 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 for each data set. Each maximum, minimum and estimated maximum, minimum data are exhibited in Table 1.

Table 1: The maximum value and the minimum value

	μ	1 to 36 months	Estimated
	$\mu_{\max}$	Maximum	
	$\mu_{\min}$	Minimum	
Food A	1.5	2180	3270
	1.5	1489	993
Food B	1.1	2561	2561
	4.0	742	266
Food C	1.1	2773	2773
	4.0	508	305
Food D	1.5	2049	2873
	1.5	536	357
Food E	1.1	1633	1633
	4.0	399	230
Food F	1.5	1415	2123
	1.5	733	501
Food G	1.5	1681	2432
	1.5	617	411
Food H	1.5	1420	2084

	1.5	519	346
Food I	1.5	1215	1823
	1.5	693	478
Food J	1.5	2041	2178
	1.5	592	395

5.2 Condition of Experiment

Condition of the neural network's experiment is exhibited in Table 2. Experiment is executed for 12 patterns $(l = 1,2,\dots,12)$ 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

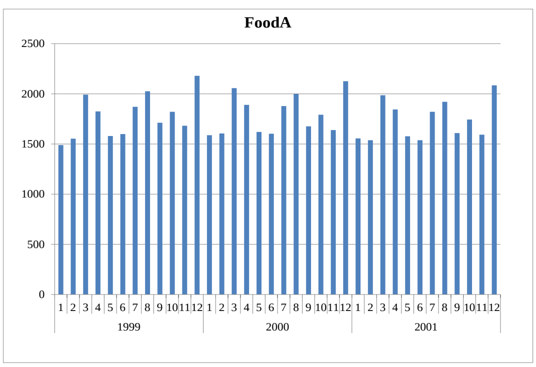

Figure 4:Shipping data of Food A

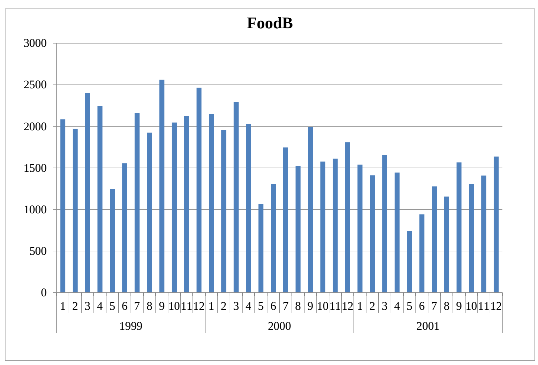

Figure 5:Shipping data of Food B

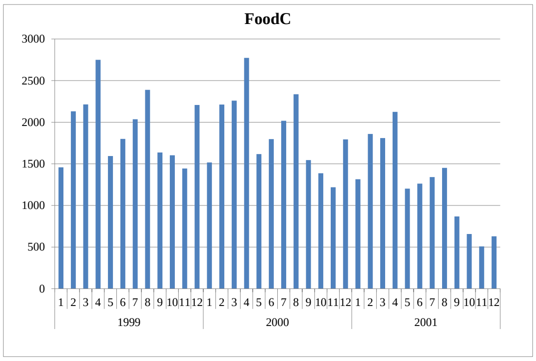

Figure 6:Shipping data of Food C

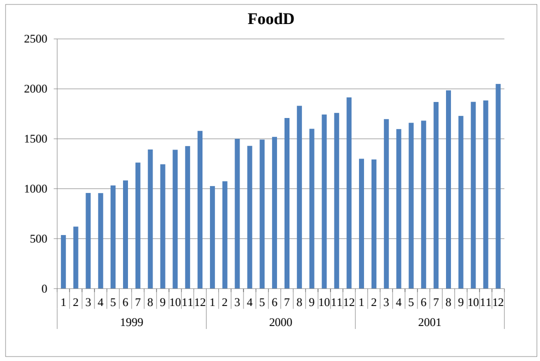

Figure 7:Shipping data of Food D

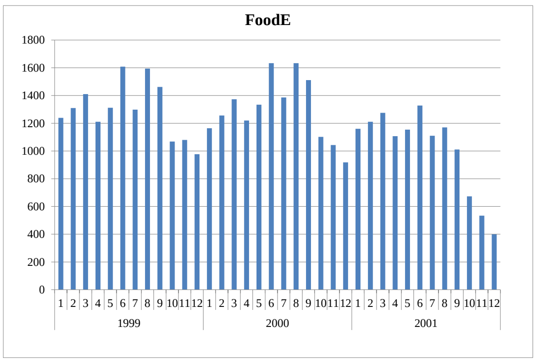

Figure 8:Shipping data of Food E

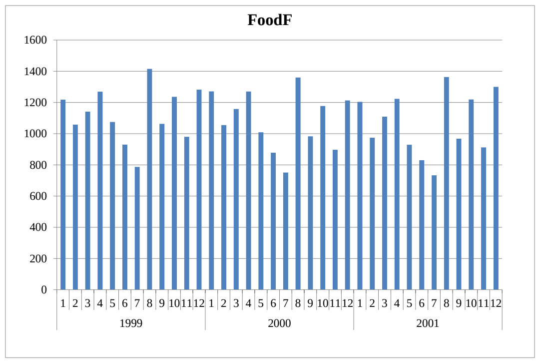

Figure 9:Shipping data of Food F

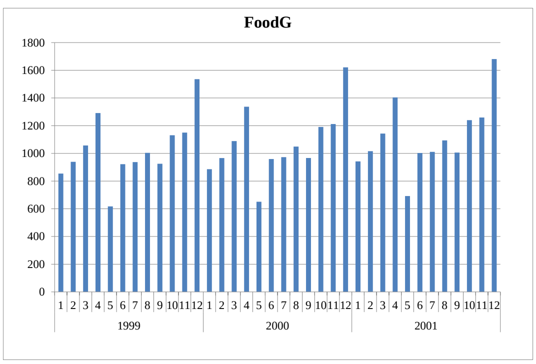

Figure 10:Shipping data of Food G

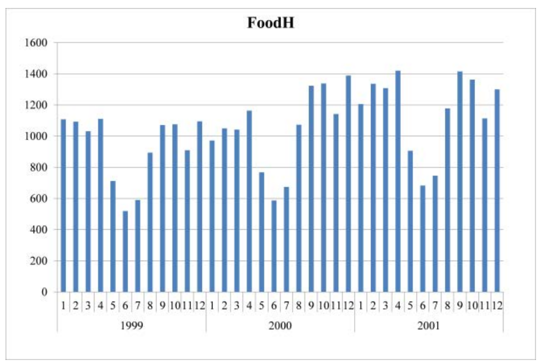

Figure 11:Shipping data of Food H

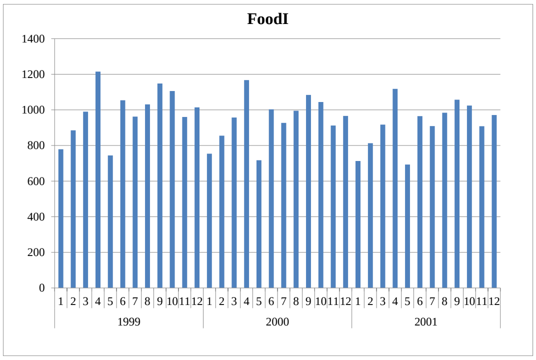

Figure 12:Shipping data of Food I

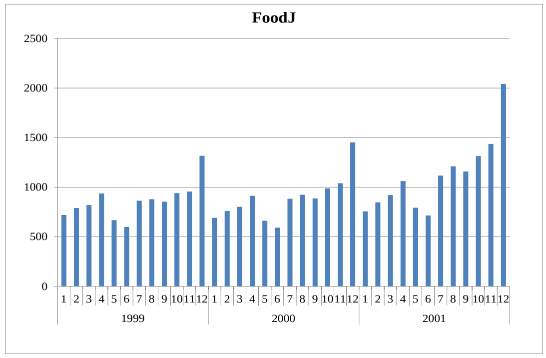

Figure 13:Shipping data of Food J

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 through 6. 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 14 through 23.

Table 3: The result for Neural network

(Food A～Food E) [τ=1]

l	A	B	C	D	E
1	90.04	71.40	47.89	89.49	78.53
2	90.55	69.40	46.95	89.20	**78.91**
3	90.61	**74.29**	37.89	90.50	77.91
4	91.27	70.05	54.98	90.51	72.84
5	90.97	73.95	**58.12**	90.45	69.56
6	91.32	71.63	50.80	89.95	71.82
7	91.45	70.00	44.35	90.72	71.25
8	92.02	61.86	46.55	90.94	72.62

Table 4: The result for Neural network

(Food F～Food J) [τ=1]

l	A	B	C	D	E
1	84.52	84.04	83.08	88.73	73.98
2	84.41	84.01	83.08	88.49	73.91
3	84.57	84.35	82.82	88.64	74.08
4	85.04	85.01	80.95	89.15	74.31
5	85.44	85.62	80.95	88.80	75.46
6	86.63	85.61	81.46	89.64	76.62
7	86.39	85.92	81.03	89.84	76.37
8	86.10	87.00	79.96	89.90	76.70

9	91.37	64.62	42.28	91.14	71.62
10	91.98	61.03	42.51	90.68	72.67
11	92.09	55.18	43.99	**91.26**	71.60
12	**93.74**	73.04	39.82	88.24	70.87

9	86.58	87.01	80.97	90.46	76.53
10	86.47	86.43	81.98	90.77	76.18
11	86.37	87.31	83.52	90.84	76.65
12	**90.46**	**90.82**	**85.76**	**93.16**	**77.06**

Table 5: The result for Neural network (Food A～Food E) [τ=4]

l	A	B	C	D	E
1	95.31	93.05	85.88	96.06	96.86
2	95.27	92.84	87.43	96.08	**96.91**
3	94.75	95.59	79.37	95.97	96.52
4	94.46	92.65	**89.56**	95.51	89.10
5	94.70	95.54	87.43	95.65	89.04
6	94.52	**95.61**	85.66	95.57	87.16
7	93.90	95.04	84.75	**96.78**	86.50
8	94.47	89.29	71.38	96.16	78.04
9	94.64	84.09	51.54	96.21	78.71
10	94.85	68.11	56.18	94.84	79.51
11	**95.49**	68.33	56.41	92.91	78.66
12	95.43	71.95	42.56	84.23	71.44

Table 3: The result for Neural network (Food F～Food J) [τ=4]

l	A	B	C	D	E
1	92.86	94.11	95.58	94.43	91.92
2	93.88	94.70	95.64	94.80	92.16
3	93.61	94.74	95.38	94.72	92.51
4	93.32	94.54	**96.06**	94.64	**93.06**
5	93.80	94.08	94.72	94.50	92.61
6	94.07	94.26	94.77	95.01	83.91
7	93.58	92.47	94.28	94.72	78.76
8	93.44	92.98	94.63	94.19	81.08
9	92.48	93.03	89.74	93.18	78.09
10	92.88	91.44	95.61	92.36	80.37
11	91.93	92.33	95.27	90.80	74.57
12	**94.53**	**96.47**	88.26	**96.09**	79.22

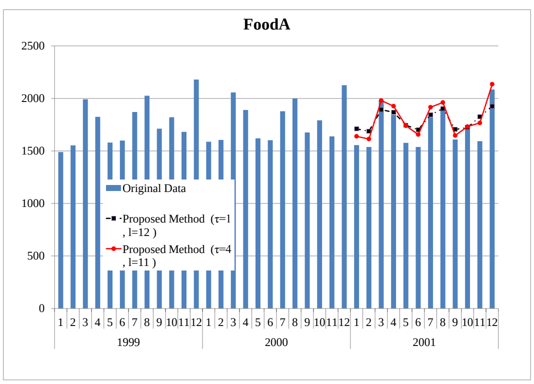

Figure 14:The result of Food A (τ=1, *l*=12) and (τ=4, *l*=11)

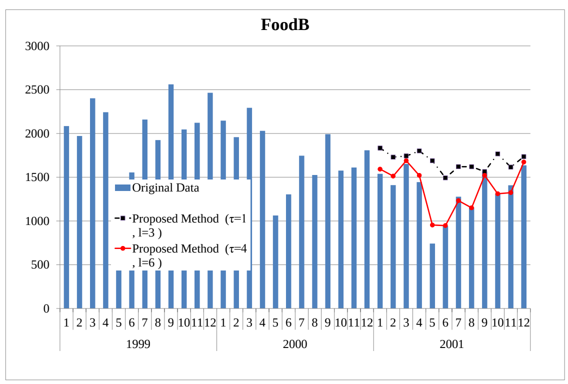

Figure 15:The result of Food B (τ=1, *l*=3) and (τ=4, *l*=6)

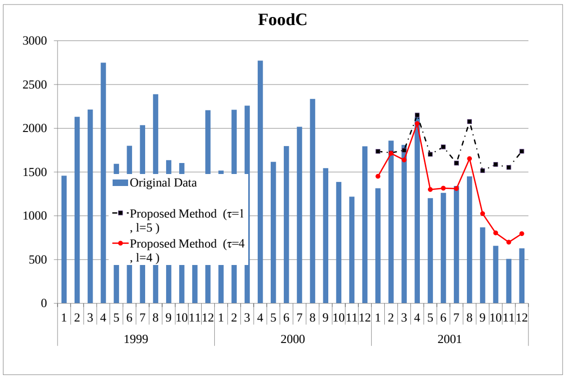

Figure 16:The result of Food C (τ=1, *l*=5) and (τ=4, *l*=4)

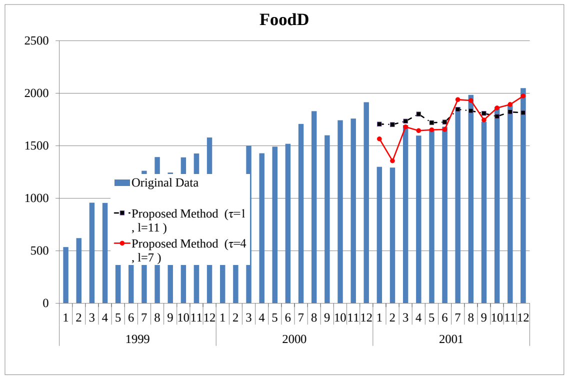

Figure 17:The result of Food D (τ=1, *l*=11) and (τ=4, *l*=7)

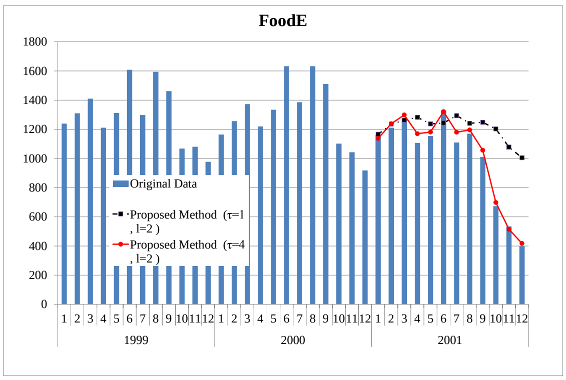

Figure 18:The result of Food E (τ=1, *l*=2) and (τ=4, *l*=2)

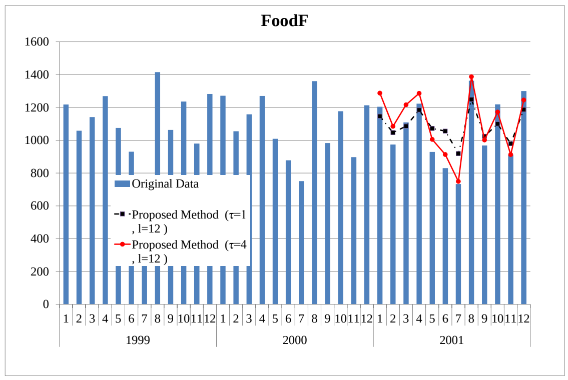

Figure 19:The result of Food F (τ=1, *l*=12) and (τ=4, *l*=12)

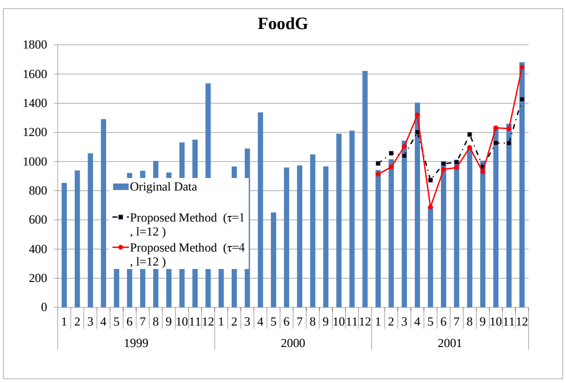

Figure 20:The result of Food G (τ=1, *l*=12) and (τ=4, *l*=12)

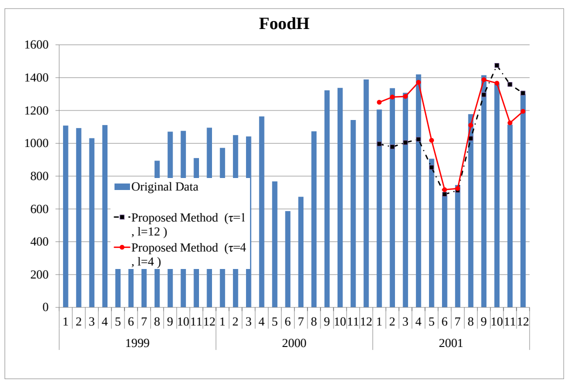

Figure 21:The result of Food H (τ=1, *l*=12) and (τ=4, *l*=4)

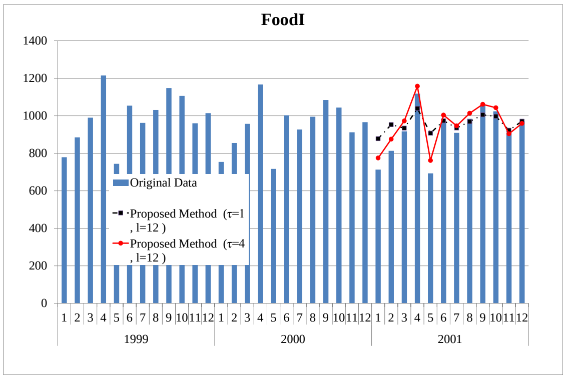

Figure 22:The result of Food I (τ=1, *l*=12) and (τ=4, *l*=12)

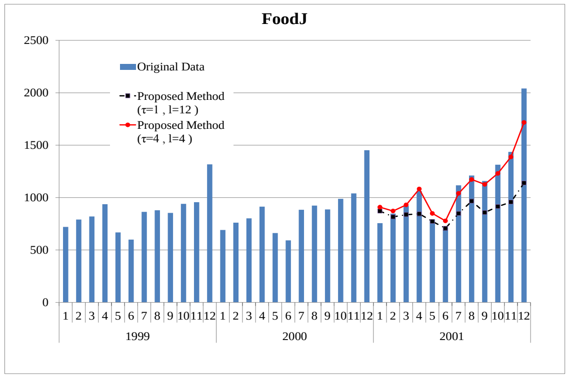

Figure 23:The result of Food A (τ=1, *l*=12) and (τ=4, *l*=4)

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 arbitrarily. 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 original shipping data of consumer goods. 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.

$$\hat{a}_{t+1} = \hat{a}_t + \alpha(a_t - \hat{a}_t) \tag{19}$$

$$= \alpha a_t + (1-\alpha)\hat{a}_t \tag{20}$$

Here,

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

a_t : realized value at t

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

(20) is re-stated as

$$\hat{a}_{t+1} = \sum_{l=0}^{\infty} \alpha(1-\alpha)^l a_{t-1} \tag{21}$$

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

$$a_t - a_{t-1} = h_t - \beta h_{t-1} \tag{22}$$

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

$$a_t + \sum_{i=1}^{p} \kappa_i a_{t-i} = h_t + \sum_{j=1}^{q} \lambda_j h_{t-j} \tag{23}$$

Here,

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

$\{h_t\}$: Gaussian White Noise with 0 mean σ_h^2 variance MA process in (23) is supposed to satisfy convertibility condition.

Finally we get

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

$$\alpha = \frac{1+2\rho_1-\sqrt{1-4\rho_1^2}}{2\rho_1} \tag{24}$$

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

$$r = b_{11}u + b_{12} \tag{25}$$

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:

$$r = b_{21}u^2 + b_{22}u + b_{23} \tag{26}$$

$$r = b_{31}u^3 + b_{32}u^2 + b_{33}u + b_{34} \tag{27}$$

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

$$\begin{aligned} r = {} & \alpha_1 (b_{11}u + b_{12}) \\ & + \alpha_2 (b_{21}u^2 + b_{22}u + b_{23}) \\ & + \alpha_3 (b_{31}u^3 + b_{32}u^2 + b_{33}u + b_{34}) \end{aligned} \tag{28}$$

$$0 \le \alpha_1 \le 1 \;, 0 \le \alpha_2 \le 1 \;, 0 \le \alpha_3 \le 1$$
$$\alpha_1 + \alpha_2 + \alpha_3 = 1 \tag{29}$$

as the combination of linear and 2nd order non-linear and 3rd order non-linear function. Trend is removed by dividing the data by (28). The optimal weighting parameters α_1, α_2, α_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.

$$\tilde{x}_\phi = \frac{\frac{1}{L}\sum_{\theta=1}^{L} x_{\theta\phi}}{\frac{1}{L} \cdot \frac{1}{12}\sum_{\theta=1}^{L}\sum_{\phi=1}^{12} x_{\theta\phi}} \tag{30}$$

Monthly trend is removed by dividing the data by (30). 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 18 through 27. Forecasting Accuracy Ratio is exhibited in Table 7.

6.6 Remarks

In all cases, that the cases monthly ratio are used have better result than those of the cases monthly ratio are not used concerning the Forecasting Accuracy Ratio (Table 7, Figure 24 through 33). This means that monthly trend removal is effective for these data on the whole.

Table 7: Forecasting Accuracy Ratio

	Monthly Ratio	
	Used	Not used
Food A	**98.20**	86.38
Food B	**92.37**	83.84
Food C	**92.36**	81.42
Food D	**98.20**	86.38
Food E	**92.11**	86.10
Food F	**94.49**	78.48
Food G	**95.45**	78.34
Food H	**90.80**	80.01
Food I	**96.78**	85.38
Food J	**93.80**	79.98

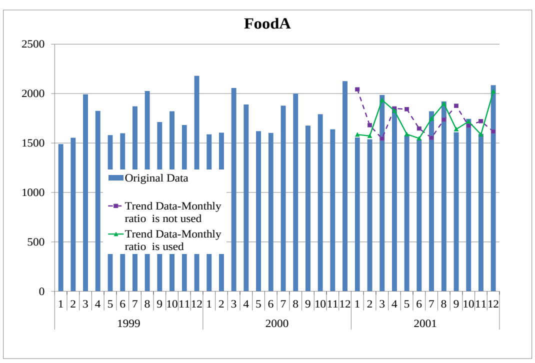

Figure 24: Forecasting result of Food A

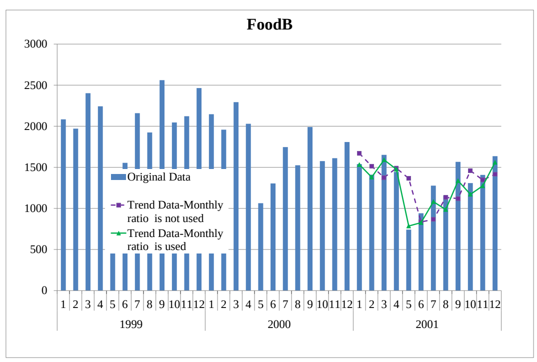

Figure 25: Forecasting result of Food B

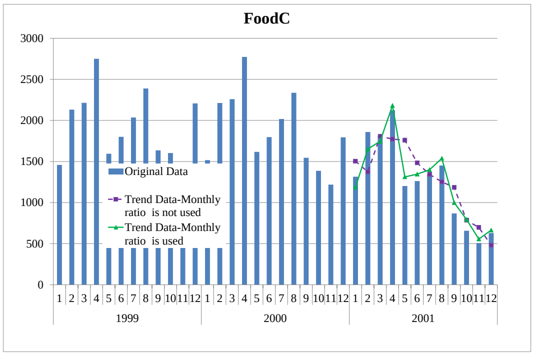

Figure 26: Forecasting result of Food C

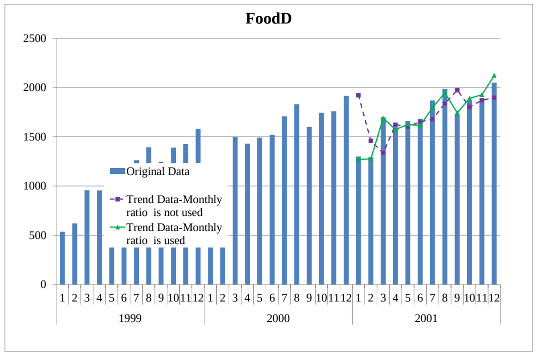

Figure 27: Forecasting result of Food D

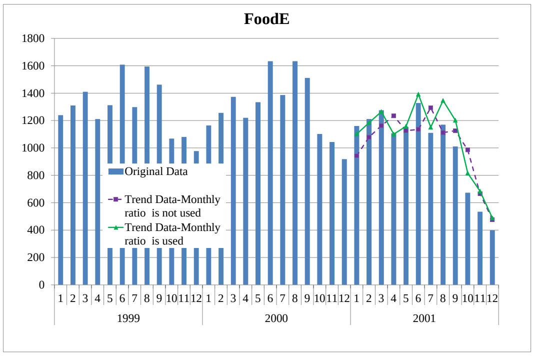

Figure 28: Forecasting result of Food E

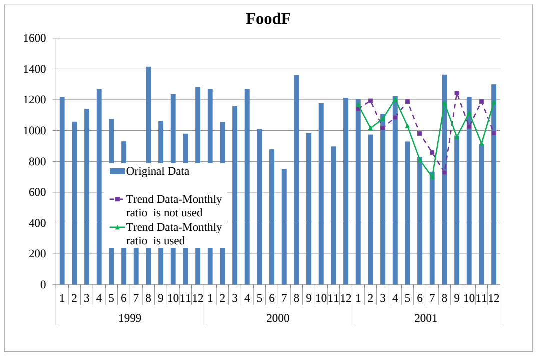

Figure 29: Forecasting result of Food F

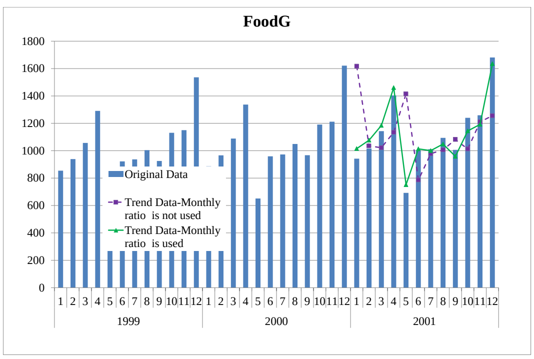

Figure 30: Forecasting result of Food G

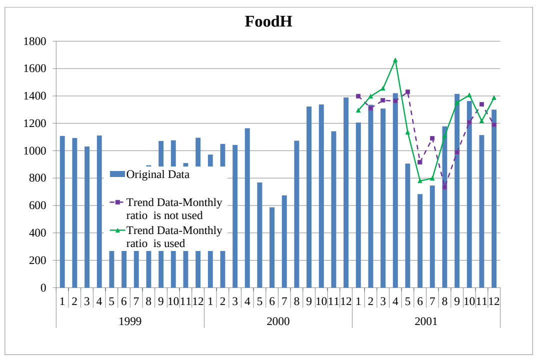

Figure 31: Forecasting result of Food H

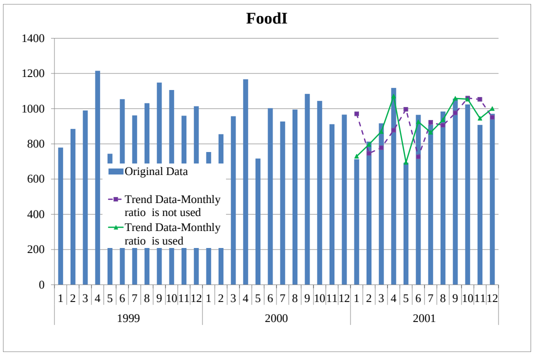

Figure 32: Forecasting result of Food I

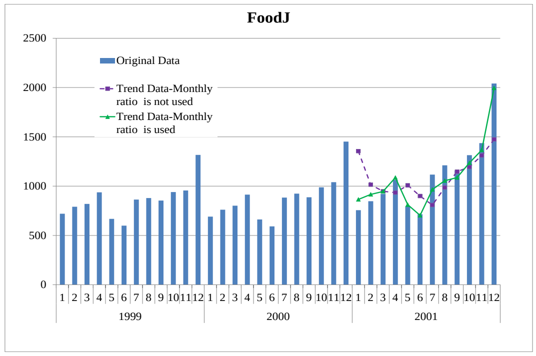

Figure 33: Forecasting result of Food J

7 Remarks

Now, we compare with both results. In Table 8, both results are stated and compared. Their comparison is shown in Figure 34 through 43. In five cases out of ten, this newly proposed method had a better forecasting accuracy than the previously proposed method.

Table 8: Comparison of the both results

	Forecasting Accuracy Ratio	
	Conventional Method	Proposed Method
Food A	98.20	**95.49** ($\tau = 4, l = 11$)
Food B	**92.37**	95.61 ($\tau = 4, l = 6$)
Food C	**92.36**	89.56 ($\tau = 4, l = 4$)
Food D	**98.20**	96.78 ($\tau = 4, l = 7$)
Food E	92.11	**96.91** ($\tau = 4, l = 2$)
Food F	94.49	**94.53** ($\tau = 4, l = 12$)
Food G	95.45	**96.47** ($\tau = 4, l = 12$)
Food H	90.80	**96.06** ($\tau = 4, l = 4$)
Food I	**96.78**	96.09 ($\tau = 4, l = 12$)
Food J	**93.80**	93.06 ($\tau = 4, l = 4$)

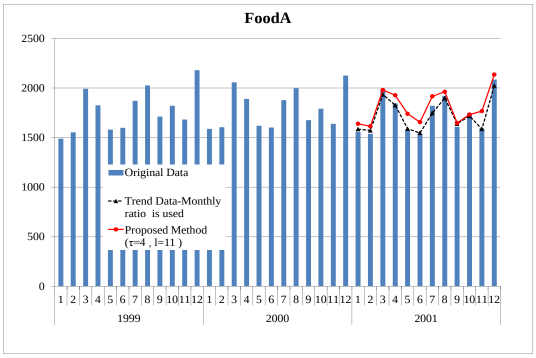

Figure 34: The result of Food A: monthly ratio is used and (τ=4, *l*=11)

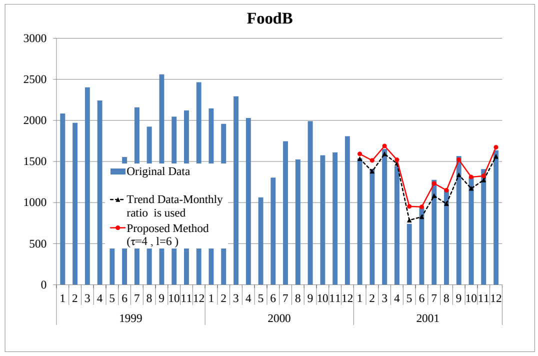

Figure 35: The result of Food B: monthly ratio is used and (τ=4, *l*=6)

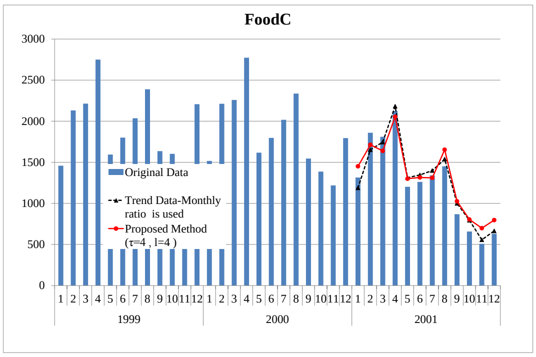

Figure 36: The result of Food C: monthly ratio is used and (τ=4, *l*=4)

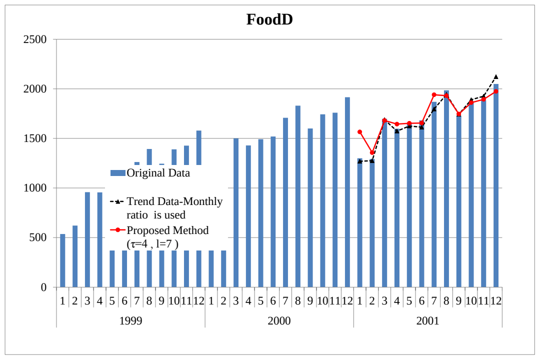

Figure 37: The result of Food D: monthly ratio is used and (τ=4, *l*=7)

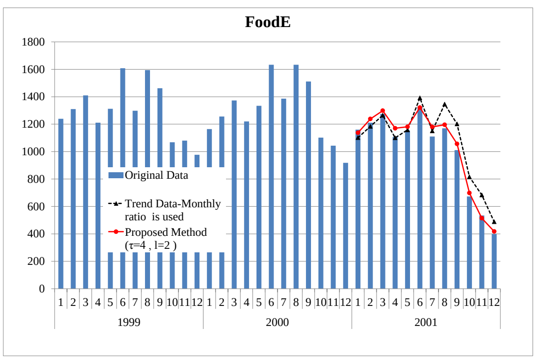

Figure 38: The result of Food E: monthly ratio is used and (τ=4, *l*=2)

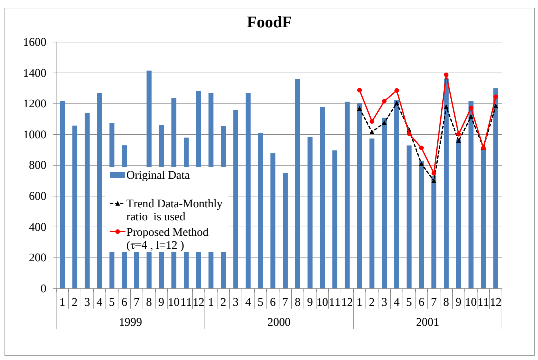

Figure 39: The result of Food F: monthly ratio is used and (τ=4, *l*=12)

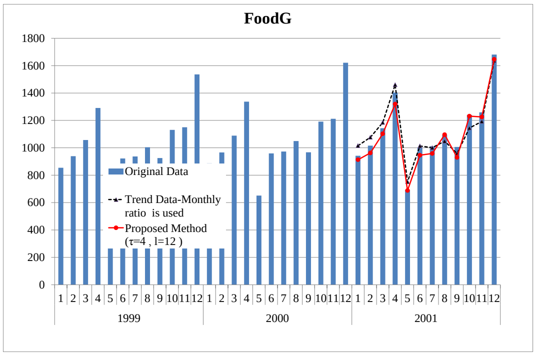

Figure 40: The result of Food G: monthly ratio is used and (τ=4, *l*=12)

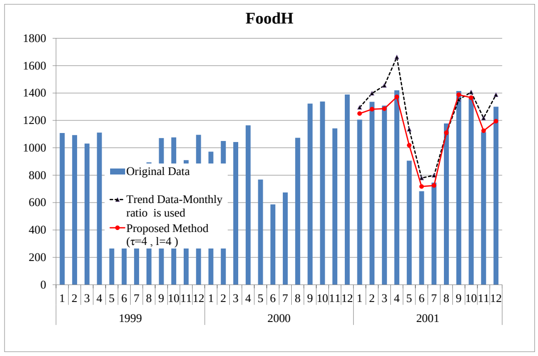

Figure 41: The result of Food H: monthly ratio is used and (τ=4, *l*=4)

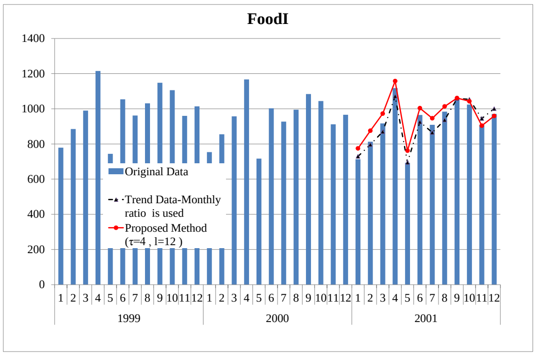

Figure 42: The result of Food I: monthly ratio is used and (τ=4, *l*=12)

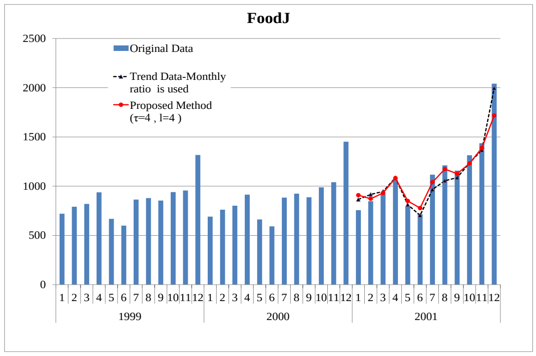

Figure 43: The result of Food J: monthly ratio is used and (τ=4, *l*=4)

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 original shipping data of consumer goods. 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. We have obtained the good results. In the numerical example, five cases out of ten had a better forecasting accuracy than the method proposed before. Various cases should be examined hereafter.

References

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

[2] N. Baba and 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 and 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] Tatsuhiro Kuroda, Yuki Higuchi and Kazuhiro Takeyasu, A Hybrid Method to Improve Forecasting Accuracy Utilizing Genetic Algorithm –An Application to the Airlines Passengers and Cargo Data, *International Journal of Information Technology and Network Application* (IJITNA), **2**(4), (2012),1-12.