# A Hybrid Method to Improve Forecasting Accuracy Utilizing Genetic Algorithm: An Application to the Data of Processed Cooked Rice

**Hiromasa Takeyasu\***
Faculty of Life and Culture, KagawaJuniorCollege, Kagawa, Japan

**Yuki Higuchi**
Department ofEconomics, Osaka Prefecture University, Osaka, Japan

**Kazuhiro Takeyasu**
College of Business Administration, Fuji-Tokoha University, Shizuoka, Japan

## ABSTRACT

In industries, shipping is an important issue in improving the forecasting accuracy of sales. This paper introduces a hybrid method and plural methods are compared. Focusing the equation of exponential smoothing method (ESM) that is equivalent to (1, 1) order autoregressive-moving-average (ARMA) model equation, a new method of estimating the smoothing constant in ESM had been proposed previously by us which satisfies minimum variance of forecasting error. Generally, the smoothing constant is selected arbitrarily. However, this paper utilizes the above stated theoretical solution. Firstly, we make estimation of ARMA model parameter and then estimate the smoothing constant. 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. This method is executed in the following method. Trend removing by the combination of linear and 2nd order nonlinear function and 3rd order nonlinear function is executed to the original production data of two kinds of bread. Genetic algorithm is utilized to search the optimal weight for the weighting parameters of linear and nonlinear function. For comparison, the 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. Then forecasting is executed on these data. The new method shows that it is useful for the time series that has various trend characteristics and has rather strong seasonal trend. The effectiveness of this method should be examined in various cases.

Keywords: Minimum Variance, Exponential Smoothing Method, Forecasting, Trend, Bread

\* Corresponding Author, E-mail: takeyasuh@kjc.ac.jp

## 1. INTRODUCTION

Many methods for time series analysis have been presented, such as autoregressive (AR) model, autoregressive moving-average (ARMA) model, and exponential smoothing method (ESM) (Box *et al.*, 1994; Brown, 1963; Kobayashi, 1992; Tokumaru *et al.*, 1982). Among these, ESM is said to be a practical simple method.

For this method, various improving methods such as adding compensating item for time lag, coping with the

time series with trend (Winters, 1960), utilizing Kalman filter (Maeda, 1984), Bayes forecasting (West and Harrison, 1989), adaptive ESM (Ekern, 1982), exponentially weighted moving averages with irregular updating periods (Johnston, 1993), making averages of forecasts using plural method (Makridakis and Winkler, 1983) are presented. For example, Maeda (1984) calculated smoothing constant in relationship with signal-to-noise ratio under the assumption that the observation noise was added to the system. But he had to calculate under supposed noise because he could not grasp observation noise. It can be said that it does not pursue optimum solution from the very data themselves which should be derived by those estimations. Ishii *et al.* (1981) pointed out that the optimal smoothing constant was the solution of infinite order equation, but he did not show analytical solution. There are some papers that utilize the neural network. Miki and Yamakawa (1996) generated time series data by nonlinear system. But it is a stationary one and does not have significant trend. Ogasahara and Inoue (2009) utilized hybrid neural network for prediction. It searches neural neighbors which have a similar typed data and utilize them as a forecast. But it needs huge data. Therefore, it cannot be used in the case we are trying to make forecast in which we handle the monthly data. Takaho *et al.* (2002) used pre data handling such as low/high pass filter. It may be a kind of trend removing. Based on these facts, we proposed a new method of estimating the smoothing constant in ESM in advance (Takeyasu and Nagao, 2008). Focusing on the equation of ESM that is equivalent to (1,1) order ARMA model equation, a new method of estimating the smoothing constant in ESM was derived.

In this paper, utilizing above stated method, a revised forecasting method is proposed. In making forecast such as production data, a trend removing method is devised. Trend removing by the combination of linear and 2nd order nonlinear function and 3rd order nonlinear function is executed to the original production data of two kinds of bread. Genetic algorithm (GA) is utilized to search the optimal weight for the weighting parameters of linear and nonlinear function. For comparison, the monthly trend is removed after that. Theoretical solution of smoothing constant of ESM is calculated for both the monthly trend removing data and the non-monthly trend removing data. Then, forecasting is executed on these data. This is a revised forecasting method. Trend removal and GA are not used in the paper of Takeyasu and Nagao (2008). The variance of forecasting error of this newly proposed method is assumed to be less than those of the previously proposed method. The rest of the paper is organized as follows. In Section 2, ESM is described by ARMA model, and the estimation method of smoothing constant is derived using ARMA model identification. The combination of linear and nonlinear function is introduced for

trend removing in Section 3. The monthly ratio is referred in Section 4. Forecasting is executed in Section 5, and estimation accuracy is examined.

## 2. DESCRIPTION OF ESM USING ARMA MODEL

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

$$\hat{x}_{t+1} = \hat{x}_t + \alpha \left( x_t - \hat{x}_t \right) \tag{1}$$

$$= \alpha x_t + (1-\alpha)\hat{x}_t \tag{2}$$

Here,

$\hat{x}_{t+1}$ : forecasting at $t+1$
$x_t$ : realized value at $t$
$\alpha$ : smoothing constant $(0 < \alpha < 1)$

(2) is re-stated as

$$\hat{x}_{t+1} = \sum_{l=0}^{\infty} \alpha (1-\alpha)^l x_{t-l} \tag{3}$$

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

$$x_t - x_{t-1} = e_t - \beta e_{t-1} \tag{4}$$

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

$$x_t + \sum_{i=1}^{p} a_i x_{t-i} = e_t + \sum_{j=1}^{q} b_j e_{t-j} \tag{5}$$

Here,

$\{x_t\}$ : Sample process of stationary ergodic Gaussian process $x(t)$ $t = 1, 2, \cdots, N, \cdots$

$\{e_t\}$ : Gaussian white noise with 0 mean $\sigma_e^2$ variance.

MA process in (5) is supposed to satisfy the convertibility condition. Utilizing the relation that

$$E\left[ e_t \mid e_{t-1}, \ e_{t-2}, \ \cdots \right] = 0$$

we get the following equation from (4).

$$\hat{x}_t = x_{t-1} - \beta e_{t-1} \tag{6}$$

Operating this scheme on $t+1$, we finally get

$$\hat{x}_{t+1} = \hat{x}_t + (1-\beta)e_t$$
$$= \hat{x}_t + (1-\beta)(x_t - \hat{x}_t) \tag{7}$$

If we set $1 - \beta = \alpha$, the above equation is the same with (1), i.e., equation of ESM is equivalent to (1,1) order ARMA model, or is said to be (0,1,1) order ARIMA model because 1st order AR parameter is $-1$. Compar-

ing with (4) and (5), we obtain

$$\begin{cases} a_1 = -1 \\ b_1 = -\beta \end{cases}$$

From (1) and (7),

$$\alpha = 1 - \beta$$

Therefore, we get

$$\begin{cases} a_1 = -1 \\ b_1 = -\beta = \alpha - 1 \end{cases} \tag{8}$$

From above, we can get the estimation of smoothing constant after we identify the parameter of MA part of ARMA model. But, generally MA part of ARMA model become nonlinear equations which are described below.

Let (5) be

$$\tilde{x}_t = x_t + \sum_{i=1}^{p} a_i x_{t-i} \tag{9}$$

$$\tilde{x}_t = e_t + \sum_{j=1}^{q} b_j e_{t-j} \tag{10}$$

We express the autocorrelation function of $\tilde{x}_t$ as $\tilde{r}$ and from (9) and (10), we get the following nonlinear equations which are well known.

$$\begin{cases} \tilde{r}_k = \sigma_e^2 \sum_{j=0}^{q-k} b_j b_{k+j} & (k \le q) \\ \quad 0 & (k \ge q+1) \\ \tilde{r}_0 = \sigma_e^2 \sum_{j=0}^{q} b_j^2 \end{cases} \tag{11}$$

For these equations, recursive algorithm has been developed. In this paper, parameter to be estimated is only $b_1$, so it can be solved in the following way.
From (4), (5), (8), and (11), we get

$$\begin{cases} q = 1 \\ a_1 = -1 \\ b_1 = -\beta = \alpha - 1 \\ \tilde{r}_0 = \left(1 + b_1^2\right)\sigma_e^2 \\ \tilde{r}_1 = b_1 \sigma_e^2 \end{cases} \tag{12}$$

If we set

$$\rho_k = \frac{\tilde{r}_k}{\tilde{r}_0} \tag{13}$$

the following equation is derived.

$$\rho_1 = \frac{b_1}{1 + b_1^2} \tag{14}$$

We can get $b_1$ as follows.

$$b_1 = \frac{1 \pm \sqrt{1 - 4\rho_1^2}}{2\rho_1} \tag{15}$$

In order to have real roots, $\rho_1$ must satisfy

$$|\rho_1| \le \frac{1}{2} \tag{16}$$

From the invertibility condition, $b_1$ must satisfy

$$|b_1| < 1$$

From (14), using the next relation,

$$(1 - b_1)^2 \ge 0$$
$$(1 + b_1)^2 \ge 0$$

(16) always holds.

As

$$\alpha = b_1 + 1$$

$b_1$ is within the range of

$$-1 < b_1 < 0$$

Finally we get

$$\left. \begin{array}{l} b_1 = \dfrac{1 - \sqrt{1 - 4\rho_1^2}}{2\rho_1} \\[3mm] \alpha = \dfrac{1 + 2\rho_1 - \sqrt{1 - 4\rho_1^2}}{2\rho_1} \end{array} \right\} \tag{17}$$

which satisfies the above condition. Thus we can obtain a theoretical solution by a simple way. Focusing on the idea that the equation of ESM is equivalent to (1,1) order ARMA model equation, we can estimate the smoothing constant after estimating ARMA model parameter. It can be estimated only by calculating 0th and 1st order autocorrelation function.

## 3. TREND REMOVAL METHOD

As trend removal method, we describe the combination of linear and nonlinear function.

### 3.1 Linear Function

We set

$$y = a_1 x + b_1 \tag{18}$$

as a linear function.

### 3.2 Nonlinear Function

We set

$$y = a_2 x^2 + b_2 x + c_2 \tag{19}$$

$$y = a_3 x^3 + b_3 x^2 + c_3 x + d_3 \tag{20}$$

as a 2nd and a 3rd order nonlinear function. $(a_2, b_2, c_2)$ and $(a_3, b_3, c_3, d_3)$ are also parameters for a 2nd and a 3rd order nonlinear functions which are estimated by using least square method.

### 3.3 The Combination of Linear and Nonlinear Function

We set

$$y = \alpha_1(a_1 x + b_1) + \alpha_2(a_2 x^2 + b_2 x + c_2) \\ + \alpha_3(a_3 x^3 + b_3 x^2 + c_3 x + d_3) \tag{21}$$

$$0 \le \alpha_1 \le 1,\ 0 \le \alpha_2 \le 1,\ 0 \le \alpha_3 \le 1,\ \alpha_1 + \alpha_2 + \alpha_3 = 1 \tag{22}$$

as the combination linear and 2nd order nonlinear and 3rd order nonlinear function. Trend is removed by dividing the original data by (21). The optimal weighting parameter $\alpha_1, \alpha_2, \alpha_3$ are determined by utilizing GA. GA method is precisely described in Section 6.

## 4. MONTHLY RATIO

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

$$\{x_{ij}\}\ (i = 1, \cdots, L)\ (j = 1, \cdots, 12)$$

where $x_{ij} \in R$ in which $j$ means month and $i$ means year and $x_{ij}$ is a production data of $i$-th year, $j$-th month. Then, monthly ratio $\tilde{x}_j\ (j = 1, \cdots, 12)$ is calculated as follows.

$$\tilde{x}_j = \frac{\dfrac{1}{L}\sum_{i=1}^{L} x_{ij}}{\dfrac{1}{L} \cdot \dfrac{1}{12}\sum_{i=1}^{L}\sum_{j=1}^{12} x_{ij}} \tag{23}$$

Monthly trend is removed by dividing the data by (23). Numerical examples both of the monthly trend removal case and the non-removal case are discussed in 7.

## 5. FORECASTING ACCURACY

Forecasting accuracy is measured by calculating the variance of the forecasting error. Variance of the forecasting error is calculated by:

$$\sigma_\varepsilon^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\varepsilon_i - \bar{\varepsilon})^2 \tag{24}$$

where the forecasting error is expressed as:

$$\varepsilon_i = \hat{x}_i - x_i \tag{25}$$

$$\bar{\varepsilon} = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i \tag{26}$$

## 6. SEARCHING OPTIMAL WEIGHTS UTILIZING GA

### 6.1 Definition of the Problem

We search $\alpha_1, \alpha_2, \alpha_3$ of (21) which minimizes (24) by utilizing GA. By (22), we only have to determine $\alpha_1$ and $\alpha_2$. $\sigma_\varepsilon^2$ (24) is a function of $\alpha_1$ and $\alpha_2$, therefore we express them as $\sigma_\varepsilon^2(\alpha_1, \alpha_2)$. Now, we pursue the following:

Minimize: $\sigma_\varepsilon^2(\alpha_1, \alpha_2)$

subject to: $0 \le \alpha_1 \le 1,\ 0 \le \alpha_2 \le 1,\ \alpha_1 + \alpha_2 \le 1$ (27)

We do not necessarily have to utilize GA for this problem which has small member of variables. Considering the possibility that variables increase when we use logistics curve, etc., in the near future, we want to ascertain the effectiveness of GA.

### 6.2 The Structure of the Gene

Gene is expressed by the binary system using $\{0,1\}$ bit. Domain of variable is $[0,1]$ from (22). We suppose that variables take down to the second decimal place. As the length of domain of variable is $1 - 0 = 1$, seven bits are required to express variables. The decimal number, the binary number, and the corresponding real number in the case of 7 bits are expressed in Table 1.

**Table 1.** Corresponding table of the decimal number, the binary number, and the real number

| Decimal number | Binary number (position of the bit) | | | | | | | Corresponding real number |
|---|---|---|---|---|---|---|---|---|
| | 6 | 5 | 4 | 3 | 2 | 1 | 0 | |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.00 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0.01 |
| 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0.02 |
| 3 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0.02 |
| 4 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0.03 |
| 5 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0.04 |
| 6 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0.05 |
| 7 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0.06 |
| 8 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0.06 |
| … | | | | | | | | … |
| 126 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0.99 |
| 127 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.00 |

**Table 2.** The gene structure

| | $\alpha_1$ | | | | | | $\alpha_2$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Position of the bit | | | | | | | | | | | | | |
| 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 | 0-1 |
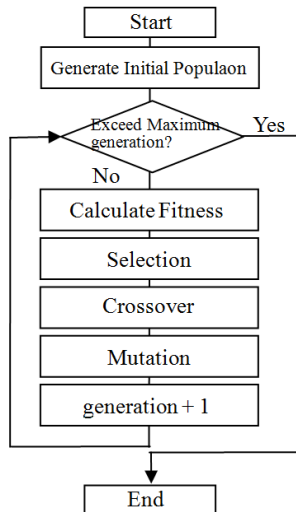


**Figure 1.** The flow of algorithm.

1 variable is expressed by 7 bits, therefore 2 variables needs 14 bits. The gene structure is exhibited in Table 2.

## 6.3 The Flow of Algorithm

The flow of algorithm is exhibited in Figure 1.

### 6.3.1 Initial population
Generate $M$ initial population. Here, $M = 100$. Generate each individual so as to satisfy (22).

### 6.3.2 Calculation of fitness
First, calculate the forecasting value. There are 36 monthly data for each case. We use 24 data (1st to 24th) and remove trend by the method described in Section 3. Then we calculate the monthly ratio by the method described in Section 4. After removing the monthly trend, the method explained in Section 2 is applied and exponential smoothing constant with minimum variance of the forecasting error is estimated. Then 1 step forecast is executed. Thus, data is shifted to 2nd to 25th and the forecast for 26th data is executed consecutively, which finally reaches the forecast of 36th data. To examine the accuracy of forecasting, variance of the forecasting error is calculated for the data of 25th to 36th data. The final forecasting data is obtained by multiplying the monthly ratio and the trend. Variance of the forecasting error is calculated by (24). Calculation of fitness is exhibited in Figure 2.
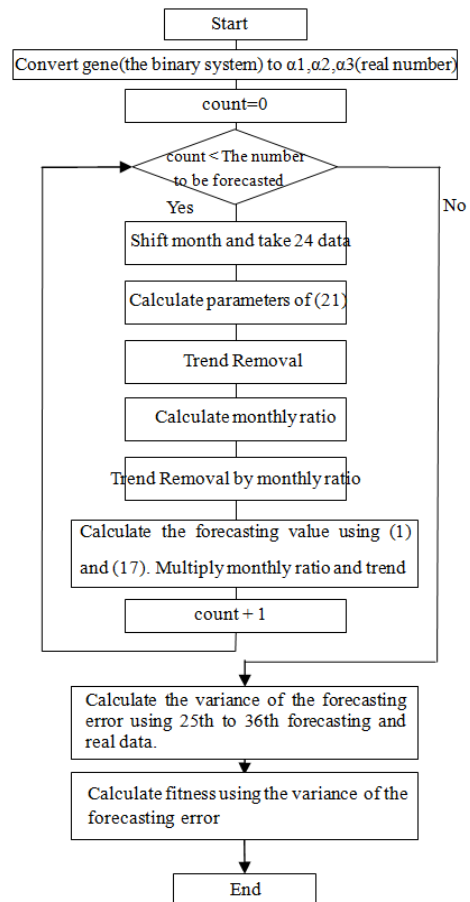


**Figure 2.** The flow of calculation of fitness.

Scaling (Ogasahara and Inoue, 2009) is executed such that fitness becomes large when the variance of forecasting error becomes small. Fitness is defined as follows.

$$f(\alpha_1, \alpha_2) = U - \sigma_\varepsilon^2(\alpha_1, \alpha_2) \qquad (30)$$

where $U$ is the maximum of $\sigma_\varepsilon^2(\alpha_1, \alpha_2)$ during the past $W$ generation. Here, $W$ is set to be 5.

### 6.3.3 Selection
Selection is executed by the combination of the general elitist selection and the tournament selection. Elitism is executed until the number of new elites reaches the predetermined number. After that, tournament selection is executed and selected.

### 6.3.4 Crossover
Crossover is executed by the uniform crossover. Crossover rate is set as follows.

$$P_c = 0.7 \qquad (31)$$

### 6.3.5 Mutation
Mutation rate is set as follows.

$$P_m = 0.05 \qquad (32)$$

Mutation is executed to each bit at the probability $P_m$,

therefore all mutated bits in the population $M$ becomes $P_m$ $\times M \times 14$.

We examined one point crossover, two points of crossover and uniform crossover, and we found that the uniform crossover was the best in convergence. Therefore, we took the uniform crossover in this case. We varied the mutation rate and found that this value was the best in performance (Takeyasu and Kainosho, 2012).

# 7. NUMERICAL EXAMPLE

## 7.1 Application to the Original Production Data of Processed Cooked Rice

We analyzed the original production data of processed cooked rice for 2 cases (Data of chilled cooked rice and those of frozen cooked rice: Annual Report of Statistical Research, Ministry of Agriculture, Forestry and Fisheries, Japan) from January 2008 to December 2010. Furthermore, GA results are compared with the calculation results of all considerable cases in order to confirm the effectiveness of GA approach. First, graphical charts of these time series data are exhibited in Figures 3 and 4.
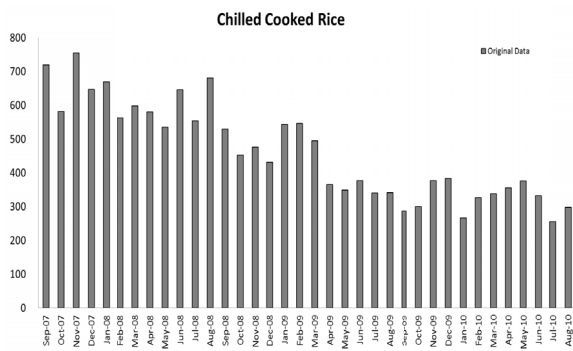


**Figure 3.** Data of chilled cooked rice.
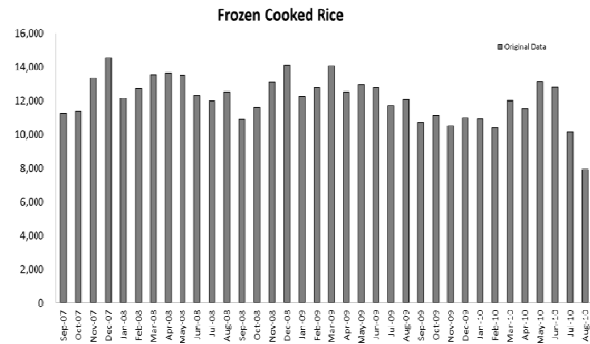


**Figure 4.** Data offrozen cooked rice.

**Table 3.** Genetic algorithm (GA) execution condition

| GA execution condition | |
|---|---|
| Population | 100 |
| Maximum generation | 50 |
| Crossover rate | 0.7 |
| Mutation ratio | 0.05 |
| Scaling window size | 5 |
| The number of elites to retain | 2 |
| Tournament size | 2 |

## 7.2 Execution Results

GA execution condition is exhibited in Table 3.

We made 10 times repetition and the maximum, average, minimum of the variance of the forecasting error and the average of convergence generation are exhibited in Tables 4 and 5.

The variance of forecasting error in the case that the monthly ratio is not used is smaller than those the monthly ratio is used for frozen cooked rice. It may be because frozen cooked ricedoes not have definite seasonal

**Table 4.** Genetic algorithm execution results (monthly ratio is not used)

| Food no. | Variance of the forecasting error | | | Average of convergence generation |
|---|---|---|---|---|
| | Maximum | Average | Minimum | |
| Chilled cooked rice | 3,138.70168 | 3,138.70168 | 3,138.70168 | 8.9 |
| Frozen cooked rice | 1,893,350.37 | 1,893,350.37 | 1,893,350.37 | 13.1 |

**Table 5.** Genetic algorithm execution results (monthly ratio is used)

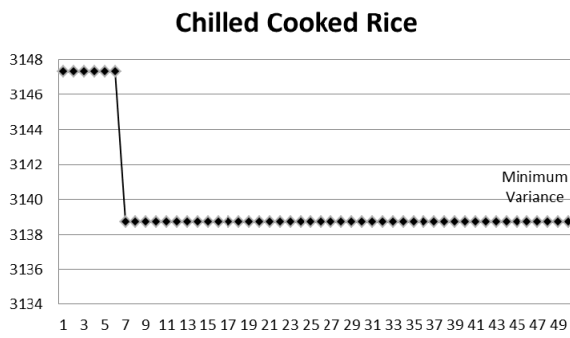| Food no. | Variance of the forecasting error | | | Average of convergence generation |
|---|---|---|---|---|
| | Maximum | Average | Minimum | |
| Chilled cooked rice | 4,026.15041 | 4,026.15041 | 4,026.15041 | 9.5 |
| Frozen cooked rice | 1,640,686.337 | 1,640,686.337 | 1,640,686.337 | 12.2 |

**Figure 5.** Convergence process in the case of chilled cooked rice (monthly ratio is not used).
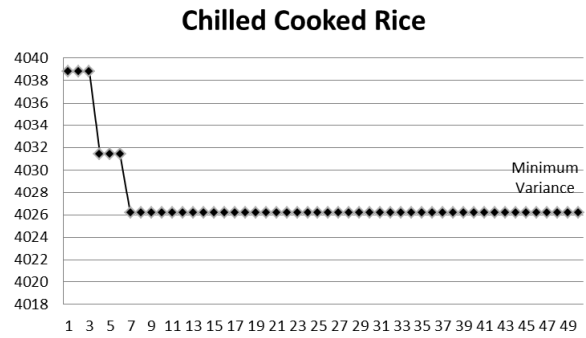


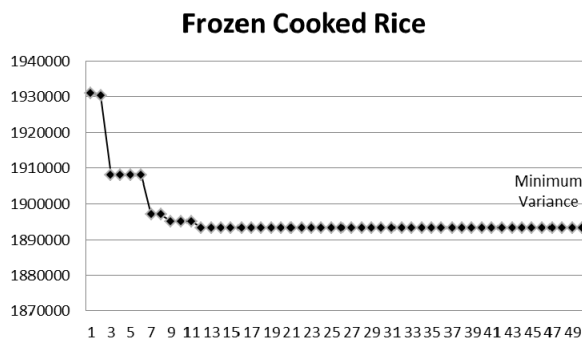**Figure 6.** Convergence process in the case of chilled cooked rice (monthly ratio is used).



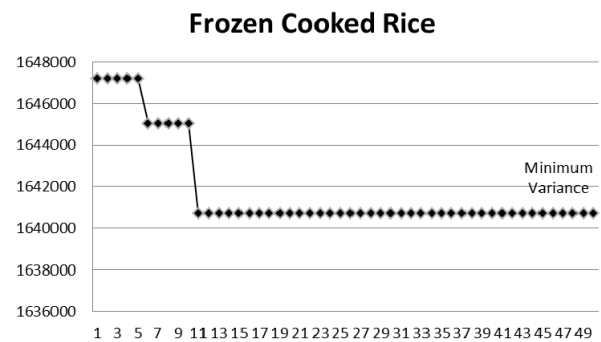**Figure 7.** Convergence process in the case of frozen cooked rice (monthly ratio is not used).



**Figure 8.** Convergence process in the case of frozen cooked rice (monthly ratio is used).

**Table 6.** Optimal weights and their genes (monthly ratio is not used)

| Data | $\alpha_1$ | $\alpha_2$ | Position of the bit | | | | | | | | | | | | | |
|------|-----------|-----------|----|----|----|----|---|---|---|---|---|---|---|---|---|---|
| | | | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| Chilled cooked rice | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Frozen cooked rice | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Table 7.** Optimal weights and their genes (monthly ratio is used)

| Data | $\alpha_1$ | $\alpha_2$ | Position of the bit | | | | | | | | | | | | | |
|------|-----------|-----------|----|----|----|----|---|---|---|---|---|---|---|---|---|---|
| | | | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| Chilled cooked rice | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Frozen cooked rice | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

trend in general.

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 should examine the complex problems hereafter.

Next, optimal weights and their genes are exhibited in Tables 6 and 7.

In the case that monthly ratio is not used, the linear function model is best in both cases. In the case that monthly ratio is used, the linear function model is best in both cases. Parameter estimation results for the trend of Eq. (21) using least square method are exhibited in Table

8 for the case of 1st to 24th data.
Trend curves are exhibited in Figures 9 and 10.

Calculation results of Monthly ratio for 1st to 24th data are exhibited in Table 9.

**Table 8.** Parameter estimation results for the trend of Eq. (21)

| Data | $a_1$ | $b_1$ | $a_2$ | $b_2$ | $c_2$ | $a_3$ | $b_3$ | $c_3$ | $d_3$ |
|---|---|---|---|---|---|---|---|---|---|
| Chilled cooked rice | -14.96 | 718.93 | -0.33 | -6.72 | 683.21 | -0.03 | 0.80 | -18.22 | 709.60 |
| Frozen cooked rice | -1.79 | 12,671.35 | -4.11 | 101.05 | 12,225.71 | 0.40 | -19.11 | 254.09 | 11,874.78 |
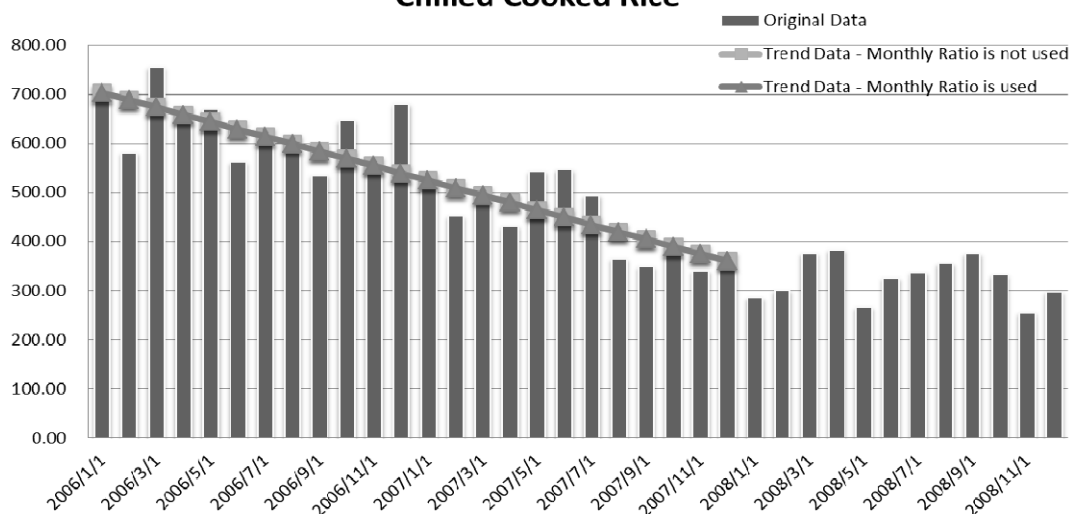


**Figure 9.** Trend of chilled cooked rice.
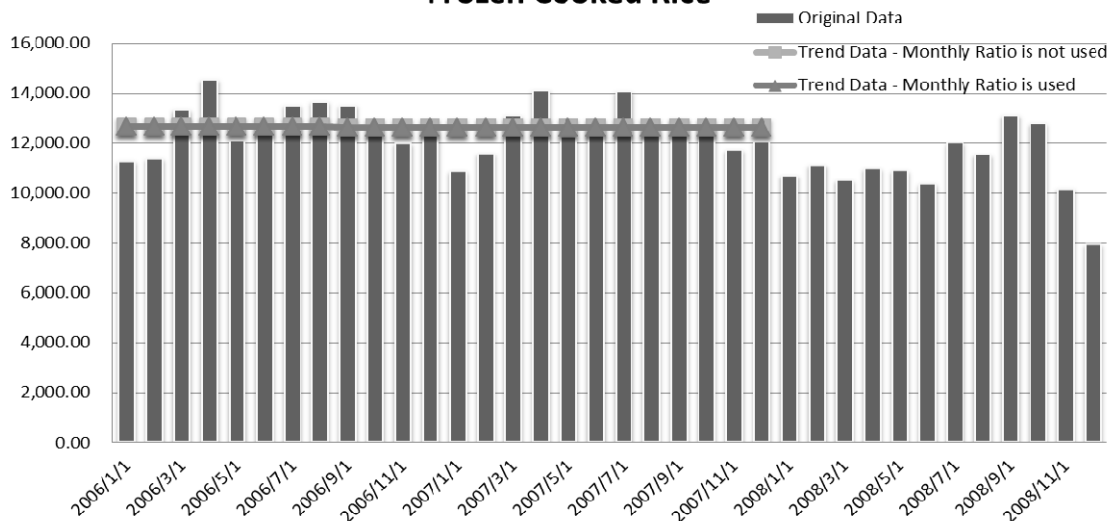


**Figure 10.** Trend of frozen cooked rice.

**Table 9.** Parameter estimation result of monthly ratio

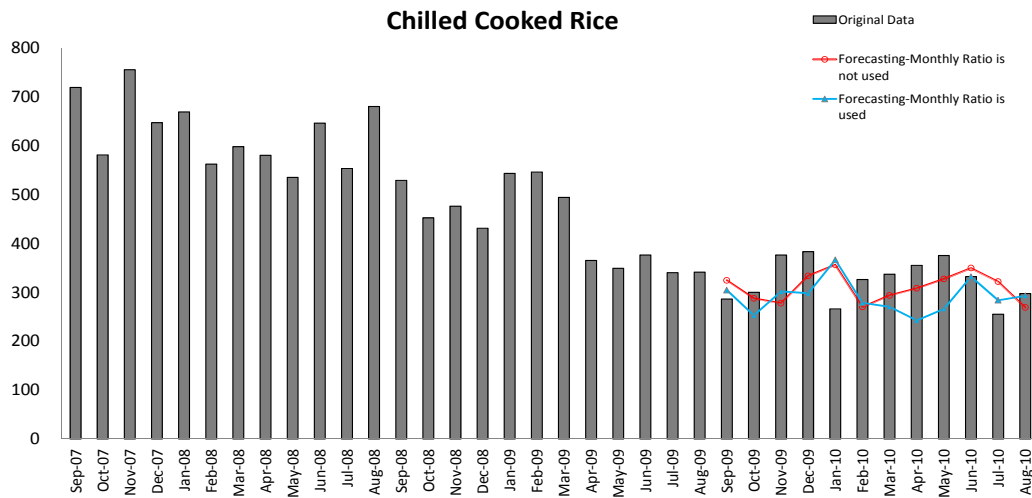| Date | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Chilled cooked rice | 1.02 | 0.87 | 1.04 | 0.94 | 1.10 | 1.05 | 1.06 | 0.92 | 0.89 | 1.05 | 0.95 | 1.11 |
| Frozen cooked rice | 0.87 | 0.91 | 1.04 | 1.13 | 0.96 | 1.01 | 1.09 | 1.03 | 1.05 | 0.99 | 0.94 | 0.97 |

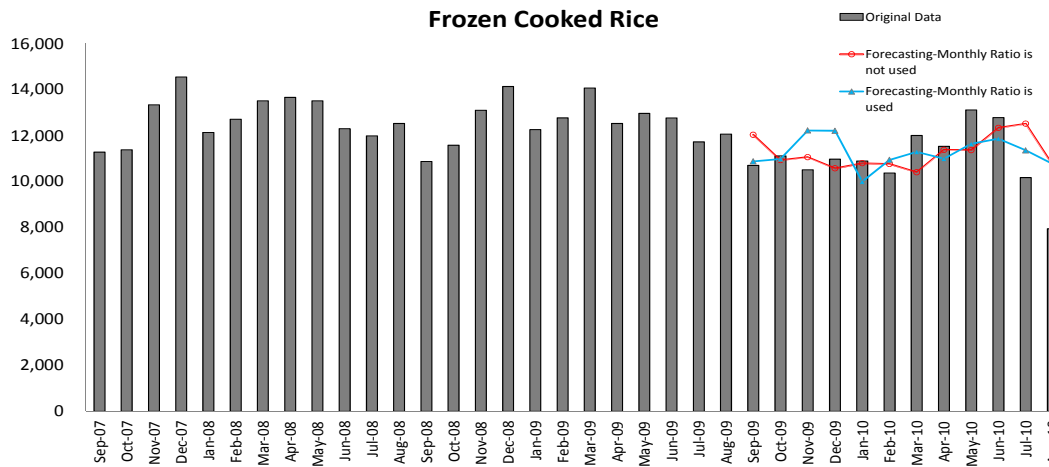**Figure 11.** Forecasting result of chilled cooked rice.



**Figure 12.** Forecasting result of frozen cooked rice.

Estimation result of the smoothing constant of minimum variance for the 1st to 24th data are exhibited in Tables 10 and 11.

**Table 10.** Smoothing constant of minimum variance of Eq. (17) (monthly ratio is not used)

| Date | ρ1 | α |
|---|---|---|
| Chilled cooked rice | -0.429544 | 0.431757 |
| Frozen cooked rice | -0.179138 | 0.814712 |

Forecasting results are exhibited in Figures 11 and 12.

**Table 11.** Smoothing constant of minimum variance of Eq. (17) (monthly ratio is used)

| Date | ρ1 | α |
|---|---|---|
| Chilled cooked rice | -0.118693 | 0.879586 |
| Frozen cooked rice | -0.179721 | 0.814066 |

### 7.3 Remarks

In the case of chilled cooked rice, it had a better forecasting accuracy in the case that the monthly ratio was not used. On the other hand, frozen cooked rice had a better forecasting accuracy when the monthly ratio was used. Both cases had a good result in the 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 should examine the complex problems hereafter.

## 8. CONCLUSION

Focusing on the idea that the equation of ESM is equivalent to (1,1) order ARMA model equation, a new

method of estimating smoothing constant in exponential smoothing method had been proposed previously by us which satisfied minimum variance of forecasting error. Generally, the smoothing constant was selected arbitrarily. But in this paper, we utilized the above stated theoretical solution. Firstly, we made estimation of the ARMA model parameter and then estimated the smoothing constant. Thus theoretical solution was derived in a simple way, and it may be utilized in various fields.

Furthermore, combining the trend removal method with this method, we aimed to improve forecasting accuracy. An approach to this method was executed in the following method. Trend removal by a linear function was applied to the original production data of processed cooked rice. The combination of linear and nonlinear function was also introduced in trend removal. GA was utilized to search the optimal weight for the weighting parameters of linear and nonlinear function. For comparison, the monthly trend was removed after that. Theoretical solution of the smoothing constant of ESM was calculated for both the monthly trend removing data and the non monthly trend removing data. Then forecasting was executed on these data. The new method shows that it is useful for the time series that has various trend characteristics. The effectiveness of this method should be examined in various cases.

## REFERENCES

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994), *Time Series Analysis: Forecasting & Control* (3rd ed.), Prentice-Hall, Englewood Cliffs, NJ.

Brown, R. G. (1963), *Smoothing, Forecasting and Prediction of Discrete Time Series*, Prentice-Hall, Englewood Cliffs, NJ.

Ekern, S. (1982), Adaptive exponential smoothing revisited, *Journal of the Operational Research Society*, **32**(9), 775-782.

Ishii, N., Iwata, A., and Suzumura, N. (1981), Bilateral exponential smoothing of time series, *International Journal of Systems Science*, **12**(8), 977-988.

Johnston, F. R. (1993), Exponentially weighted moving average (EWMA) with irregular updating periods, *Journal of the Operational Research Society*, **44**(7),

711-716.

Kobayashi, K. (1992), *Sales Forecasting for Budgeting*, Chuokeizai-Sha Publishing, Tokyo, Japan.

Maeda, K. (1984), Smoothing constant of exponential smoothing method, *Report Seikei University, Faculty of Engineering*, **38**, 2477-2484.

Makridakis, S. and Winkler, R. L. (1983), Averages of forecasts: some empirical results, *Management Science*, **29**(9), 987-996.

Miki, T. and Yamakawa, T. (1999), Analog implementation of neo-fuzzy neuron and its on-board learning, *Proceedings of the 3rd IMACS/IEEE International Multiconference on Circuits, Systems, Communications, and Computers*, Athens, Greece.

Ogasahara, T. and Inoue, H. (2009), Efficient hybrid neural network for chaotic time series prediction, *Proceedings of the Forum on Information Technology (FIT2009)*, Sendai, Japan, 597-598.

Takaho, H., Arai, T., Otake, T., and Tanaka, M. (2002), Prediction of the next stock price using neural network: extraction the feature to predict next stock price by filtering, *IEICE Technical Report: Nonlinear Problems*, **102**(432), 13-16.

Takeyasu, K. (1996), *System of Production, Sales and Distribution*, Chuokeizai-Sha Publishing, Tokyo, Japan.

Takeyasu, K. and Kainosho, M. (2012), Optimization Technique by Genetic Algorithms for International Logistics, presented at 2012 International Symposium on Semiconductor Manufacturing Intelligence (ISMI 2012), Hsinchu, Taiwan.

Takeyasu, K. and Nagao, K. (2008), Estimation of smoothing constant of minimum variance and its application to industrial data, *Industrial Engineering and Management Systems*, **7**(1), 44-50.

Tokumaru, H., Soeda, T., Nakamizo, T., and Akizuki, K. (1982), *Analysis and Measurement: Theory and Application of Random Data Handling*, Baifukan Publishing, Tokyo, Japan.

West, M. and Harrison, J. (1989), *Bayesian Forecasting and Dynamic Models* (1st ed.), Springer-Verlag, New York, NY.

Winters, P. R. (1960), Forecasting sales by exponentially weighted moving averages, *Management Science*, **6**(3), 324-343.