A Hybrid Method to Improve Forecasting Accuracy Utilizing Genetic Algorithm –An Application to the Data of Operating equipment and supplies –

Asami Shitara
Tax Corporation Arknet, Shizuoka City, Japan, e-mail: a-shitara@arknet.info

Daisuke Takeyasu
The Open University of Japan, e-mail: take1111@hotmail.co.jp

Kazuhiro Takeyasu
Tokoha University, e-mail: takeyasu@fj.tokoha-u.ac.jp

Abstract - In industries, how to improve forecasting accuracy such as sales, shipping is an important issue. There are many researches made on this. In this paper, a hybrid method is introduced and plural methods are compared. Focusing that the equation of exponential smoothing method(ESM) is equivalent to (1,1) order ARMA model equation, new method of estimation of smoothing constant in exponential smoothing method is proposed before by us which satisfies minimum variance of forecasting error. Trend removing by the combination of linear and 2^{nd} order non-linear function and 3^{rd} order non-linear function is executed to the data of Operating equipment and supplies for three cases (An injection device and a puncture device, A sterilized hypodermic needle and A sterilized syringe). Genetic Algorithm is utilized to search the optimal weight 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. Then forecasting is executed on these data. The new method would be useful for the time series that has various trend characteristics.

Key Words: minimum variance, forecasting, operating equipment and supplies

1. INTRODUCTION

Many methods for time series analysis have been presented such as Autoregressive model (AR Model), Autoregressive Moving Average Model (ARMA Model) and Exponential Smoothing Method (ESM).

In this paper, a revised forecasting method is proposed. In making forecast such as production data, trend removing method is devised. Trend removing by the combination of linear and 2^{nd} order non-linear function and 3^{rd} order non-linear function is executed to the data of Operating equipment and supplies for three cases (An injection device and a puncture device, A sterilized hypodermic needle and A sterilized syringe). These Operating equipment and supplies are used for medical use. Genetic Algorithm is utilized to search the optimal weight 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. Then forecasting is executed on these data. This is a revised forecasting method. The new method would be useful for the time series that has various trend characteristics. The rest of the paper is organized as follows. In section 2, ESM is stated by ARMA model and estimation method of smoothing constant is derived using ARMA model identification. The combination of linear and non-linear function is introduced for trend removing in section 3. The Monthly Ratio is referred in section 4. Forecasting Accuracy is defined in section 5. Optimal weights are searched in section 6. Forecasting is carried out in section 7, 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

α : 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 σ_e^2 variance

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

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

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

$$\begin{aligned} \hat{x}_{t+1} &= \hat{x}_t + (1-\beta)e_t \\ &= \hat{x}_t + (1-\beta)\left(x_t - \hat{x}_t\right) \end{aligned} \tag{7}$$

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. Comparing with (4) and (5), we obtain

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

From (1), (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 estimation of smoothing constant after we identify the parameter of MA part of ARMA model. But, generally MA part of ARMA model become non-linear 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 $\tilde{r}_k$ and from (9), (10), we get the following non-linear equations which are well known.

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

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) (11), we get

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

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, ρ_1 must satisfy

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

From invertibility condition, b_1 must satisfy

$$|b_1| < 1$$

From (14), using the next relation,

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

(16) always holds.
As

$$\alpha = b_1 + 1$$

b_1 is within the range of

$$-1 < b_1 < 0$$

Finally we get

$$\left.\begin{aligned} b_1 &= \frac{1-\sqrt{1-4\rho_1^2}}{2\rho_1} \\ \alpha &= \frac{1+2\rho_1-\sqrt{1-4\rho_1^2}}{2\rho_1} \end{aligned}\right\} \tag{17}$$

which satisfies 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 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 non-linear function.

[1] Linear function
We set

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

as a linear function.

[2] Non-linear 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 non-linear 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 non-linear functions which are estimated by using least square method.

[3] The combination of linear and non-linear function.
We set

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

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

as the combination linear and 2nd order non-linear and 3rd order non-linear 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 shipping data of i -th year, j -th month. Then, monthly ratio $\tilde{x}_j$ $(j = 1, \cdots, 12)$ is calculated as follows.

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

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

5. FORECASTING ACCURACY

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

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

Where, forecasting error is expressed as:

$$\varepsilon_i = \hat{x}_i - x_i \quad (25)$$

$$\bar{\varepsilon} = \frac{1}{N}\sum_{i=1}^{N}\varepsilon_i \quad (26)$$

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 α_1 and α_2. ${\sigma_\varepsilon}^2$((24)) is a function of α_1 and α_2, therefore we express them as $\sigma_\varepsilon^2(\alpha_1, \alpha_2)$. Now, we pursue the following:

$$\begin{aligned} &\text{Minimize: } \sigma_\varepsilon^2(\alpha_1,\alpha_2) \\ &\text{subject to: } 0 \le \alpha_1 \le 1, 0 \le \alpha_2 \le 1, \alpha_1 + \alpha_2 \le 1 \end{aligned} \quad (27)$$

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 binary bit strings <bit6, ～,bit0> is decoded to the [0,1] domain real number by the following procedure.[2]

Procedure 1:Convert the binary number to the binary-coded decimal.

$$\begin{aligned} &\left(\langle bit_6, bit_5, bit_4, bit_3, bit_2, bit_1, bit_0 \rangle\right)_2 \\ &= \left(\sum_{i=0}^{6} bit_i 2^i\right)_{10} \\ &= X' \end{aligned} \quad (28)$$

Procedure 2: Convert the binary-coded decimal to the real number.
The real number
= (Left hand starting point of the domain)
+ X'((Right hand ending point of the domain)/ $(2^7 - 1)$) (29)

1 variable is expressed by 7 bits, therefore 2 variables needs 14 bits.

6.3 The flow of Algorithm

The flow of algorithm is exhibited in Figure 6-1.

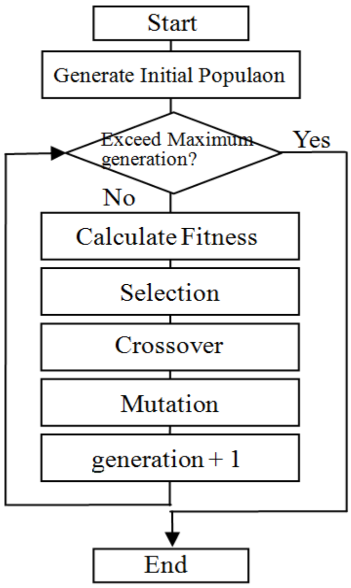

Figure 6-1: The flow of algorithm

A. Initial Population

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

B. Calculation of Fitness

First of all, calculate forecasting value. There are 36 monthly data for each case. We use 24 data(1st to 24th) and remove trend by the method stated in section 3. Then we calculate monthly ratio by the method stated in section 4. After removing monthly trend, the method stated in section 2 is applied and Exponential Smoothing Constant with minimum variance of 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 forecast of 36th data. To examine the accuracy of forecasting, variance of forecasting error is calculated for the data of 25th to 36th data. Final forecasting data is obtained by multiplying monthly ratio and trend. Variance of forecasting error is calculated by (24). Calculation of fitness is exhibited in Figure 6-2.

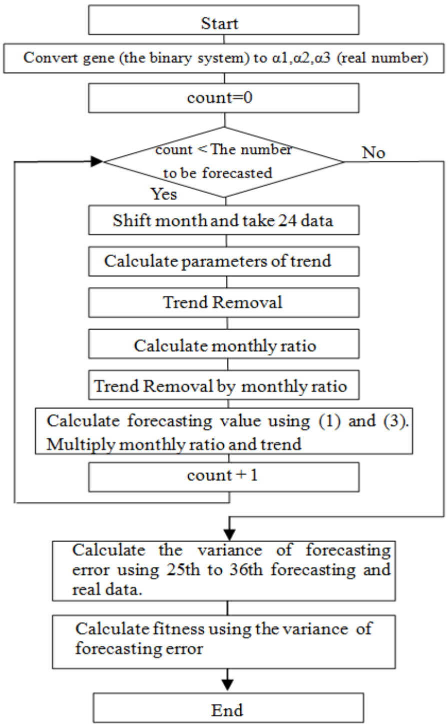

Figure 6-2:The flow of calculation of fitness

Scaling[3] is executed such that fitness becomes large when the variance of forecasting error becomes small. Fitness is defined as follows.

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

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.

C. 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.

D. Crossover

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

$$P_c = 0.7 \tag{31}$$

E. Mutation

Mutation rate is set as follows.

$$P_m = 0.05 \tag{32}$$

Mutation is executed to each bit at the probability P_m, therefore all mutated bits in the population M becomes $P_m \times M \times 14$.

7. NUMERICAL EXAMPLE

7.1 Application to the original production data o f Wheelchairs

The data of Operating equipment and supplies for three cases (An injection device and a puncture device, A sterilized hypodermic needle and A sterilized syringe) from January 2010 to December 2012 are analyzed. These data are obtained from the Annual Report of Statistical Investigation on Statistical-Survey-on-Trends-in-Pharmaceutical-Production by Ministry of Health, Labour and Welfare in Japan. Furthermore, GA results are compared with the calculation results of all considerable cases in order to confirm the effectiveness of GA approach.

7.2 Execution Results

GA execution condition is exhibited in Table 7-1.

Table7-1: 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

We made 10 times repetition and the maximum, average, minimum of the variance of forecasting error and the average of convergence generation are exhibited in Table 7-2 and 7-3.

Table7-2: GA execution results(Monthly ratio is not used)

Food No	The variance of forecasting error			Average of convergence generation
	Maximum	Average	Minimum	
An injection device and a puncture device	551,855,685,384	504,204,854,970	497,434,740,003	15.7
A sterilized hypodermic needle	91,638,679,323	37,319,843,068	28,489,531,611	9.5
A sterilized syringe	176,511,823,650	93,652,636,003	82,555,206,063	14.1

Table7-3: GA execution results(Monthly ratio is used)

Food No	The variance of forecasting error			Average of convergence generation
	Maximum	Average	Minimum	
An injection device and a puncture device	162,051,318,390	106,546,694,680	95,793,948,965	7.3
A sterilized hypodermic needle	79,422,467,024	47,074,155,352	42,493,594,397	11.1
A sterilized syringe	65,371,396,358	38,622,248,849	35,196,139,960	12.4

The variance of forecasting error for the case monthly ratio is not used is smaller than the case monthly ratio is used in A sterilized hypodermic needle. Other cases had good results in the case monthly ratio was used.

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.

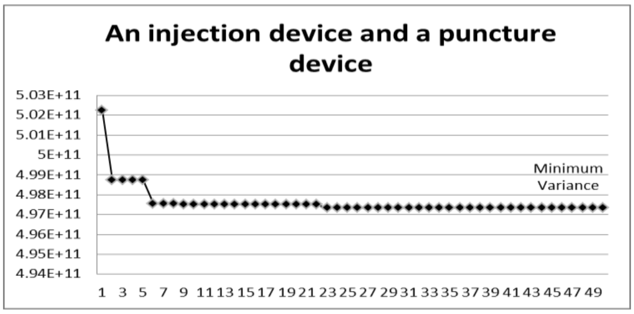

Figure7-1:Convergence Process in the case of An injection device and a puncture device (Monthly ratio is not used)

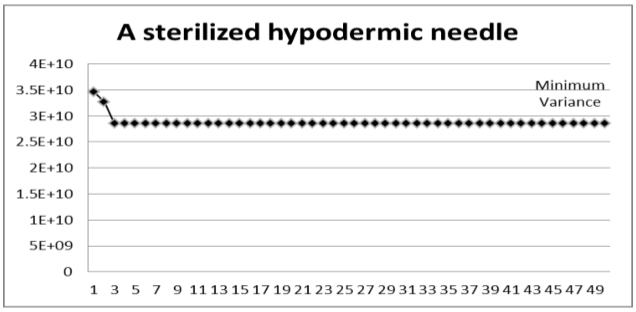

Figure7-2:Convergence Process in the case of A sterilized hypodermic needle (Monthly ratio is used)

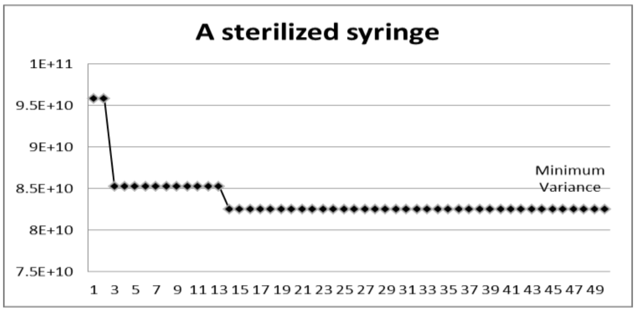

Figure7-3:Convergence Process in the case of A sterilized syringe (Monthly ratio is not used)

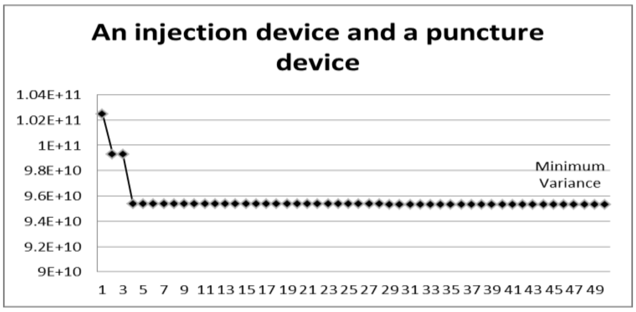

Figure7-4:Convergence Process in the case of An injection device and a puncture device (Monthly ratio is used)

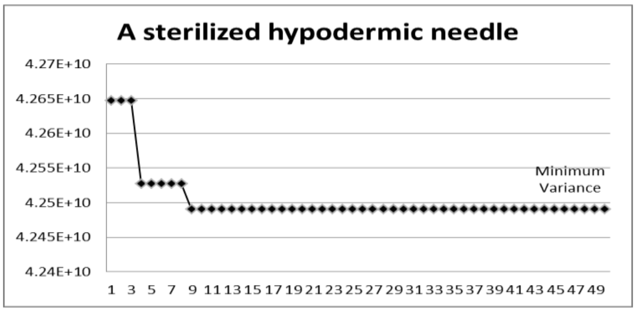

Figure7-5:Convergence Process in the case of
A sterilized hypodermic needle (Monthly ratio is not used)

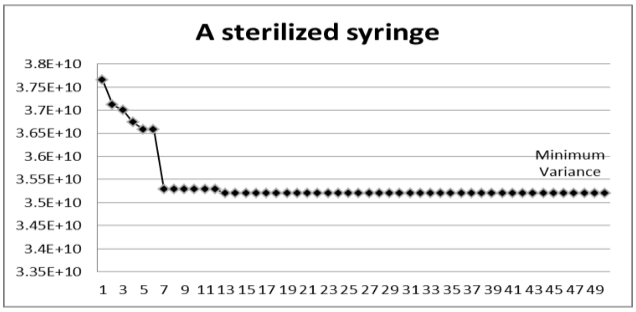

Figure7-6:Convergence Process in the case of
A sterilized syringe (Monthly ratio is used)

The linear function model is best in A sterilized syringe. An injection device and a puncture device selected 1st+2nd order function as the best one. A sterilized hypodermic needle selected 1st +2nd +3rd order function as the best one. These results were same for both of "Monthly ratio is not used" and "Monthly ratio is not used". Parameter estimation results for the trend of equation (21) using least square method are exhibited in Table 7-4 for the case of 1st to 24th data.

Table7-4: Parameter estimation results for the trend of equation (21)

Data	a_1	b_1	a_2	b_2	c_2	a_3	b_3	c_3	d_3
An injection device and a puncture device	8836.40	3971839.34	3616.96	-81587.61	4363676.73	466.59	-13880.33	96978.10	3954240.07
A sterilized hypodermic needle	7625.06	717466.59	369.11	-1602.74	757453.7	242.23	-8714.55	91099.12	544895.87
A sterilized syringe	-2041.09	1469255.56	1353.07	-35867.75	1615837.78	161.02	-4685.33	25756.08	1474539.34

Trend curves are exhibited in Figure 7-7 - 7-9.

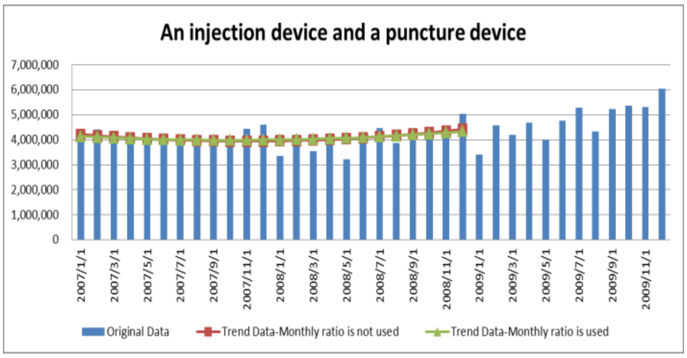

Figure7-7:Trend of An injection device and a puncture device

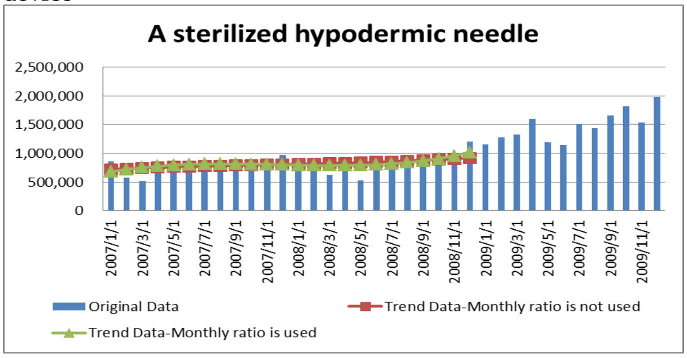

Figure7-8:Trend of A sterilized hypodermic needle

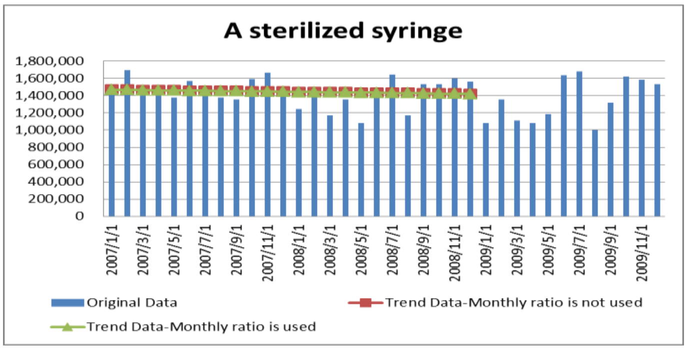

Figure7-9:Trend of A sterilized syringe

Calculation results of Monthly ratio for 1st to 24th data are exhibited in Table 7-5.

Table7-5: Parameter Estimation result of Monthly ratio

Date.	1	2	3	4	5	6	7	8	9	10	11	12
An injection device and a puncture device	0.901	1.038	0.960	1.015	0.868	0.997	1.039	0.968	0.999	0.995	1.063	1.158

A sterilized hypodermic needle	1.146	0.900	0.735	1.072	0.866	1.033	1.094	0.984	0.993	0.993	0.986	1.198
A sterilized syringe	0.941	1.076	0.887	0.95	0.84	1.02	1.078	0.883	1.004	1.088	1.137	1.071

Forecasting results are exhibited in Figure 7-10 - 7-12.

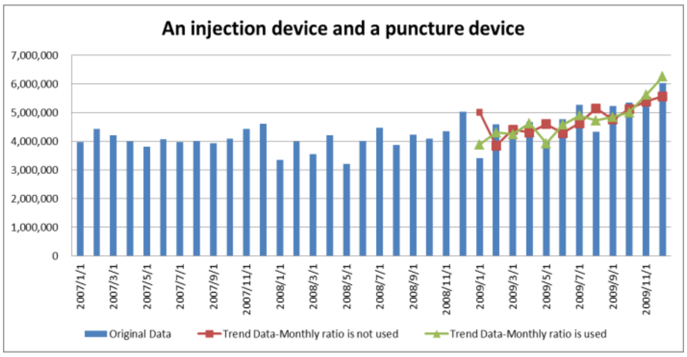

Figure 7-10: Forecasting Result of An injection device and a puncture device

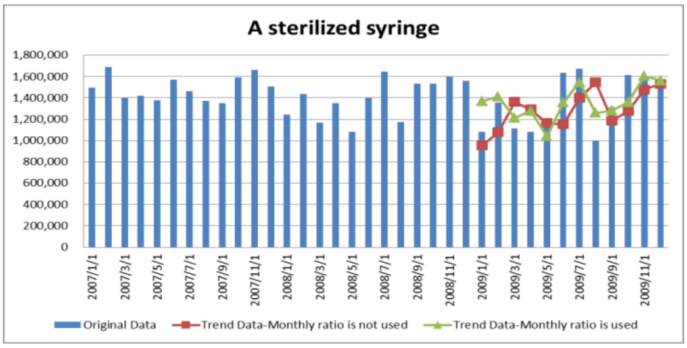

Figure 7-11: Forecasting Result of A sterilized syringe

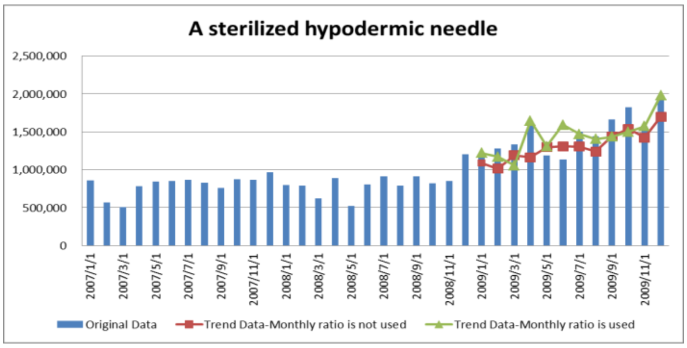

Figure 7-12: Forecasting Result of A sterilized hypodermic needle

7.3 Remarks

The linear function model in the case Monthly ratio was used was best for A sterilized syringe case. $1^{st}+2^{nd}$ function model in the case Monthly ratio was used was best for An injection device and a puncture device case. $1^{st}+2^{nd}+3^{rd}$ function model in the case Monthly ratio was not used was best for A sterilized hypodermic needle case.

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.

8.CONCLUSION

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 data of Operating equipment and supplies for three cases (An injection device and a puncture device, A sterilized hypodermic needle and A sterilized syringe). The combination of linear and non-linear function was also introduced in trend removal. Genetic Algorithm was utilized to search the optimal weight for the weighting parameters of linear and non-linear function. For the comparison, monthly trend was removed after that. Theoretical solution of smoothing constant of ESM was calculated for both of 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

[1] Kazuhiro Takeyasu and Kazuko Nagao.(2008) *Estimation of Smoothing Constant of Minimum Variance and its Application to Industrial Data*, Industrial Engineering and Management Systems, vol.7, no. 1, pp. 44-50.

[2] Masatosi Sakawa. Masahiro Tanaka. (1995) *Genetic Algorithm*、Asakura Pulishing Co., Ltd.

[3] Hitoshi Iba.（2002）*Genetic Algorithm*、Igaku Publishing.

[4] H.Takeyasu, Y.Higuchi and K.Takeyasu. (2013) A Hybrid Method to Improve Forecasting Accuracy Utilizing Genetic Algorithm –An Application to the Data of Processed Cooked Rice-, Industrial Engineering and Management Systems Vol.12, No.3, pp.244-253