An application of a generalized inverse regression estimator to Taguchi's T-Method

Hiroshi Kawada*,1, Yasushi Nagata1

1. Waseda University : 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan.
*contact author's e-mail address: mzs.t-d.m.m@akane.waseda.jp

Abstract:
Taguchi’s T-method is a part of the Mahalanobis-Taguchi (MT) system. The T-method is a technique for making predictions, and it has applications in various fields, including economic forecasting and predicting demand. The T-method can be summarized as follows: Prepare many weak learners that each predict one output value. Then, use the signal to noise ratio (S/N) to perform a weighted integration of these predictions and derive an overall estimate.

If the accuracies of each of the individual estimates are increased, the accuracy of the integrated estimate will be increased. Therefore, in this paper, we propose a way to improve the accuracies of the individual estimates. Our method uses the concept of a generalized inverse regression estimator for each of the individual estimators.

We use Monte Carlo simulations under various conditions to compare the prediction accuracies of the proposed method and that of the existing method. In many cases, our results show that the prediction accuracy of the proposal method is better than that of the original T-method. In particular, when the number of samples is small, the prediction accuracy of the proposal method is much better than that of the original T-method. Thus, we conclude that our proposed method is effective.

Keywords
MT system, T-method, Prediction, Generalized inverse regression estimator, Monte Carlo simulation.

1. Introduction

The Mahalanobis-Taguchi (MT) system is a series of techniques developed by Genichi Taguchi. The details of the methods in the MT system can be found in Taguchi and Jugulum (2002), Tatebayashi et al. (2008), and Nagata (2013). These methods can be divided into two groups. The first group consists of methods that determine which data points are abnormal or normal. Typical methods in this group are the MT method, the MT adjoint (MTA) method, and the Recognition-Taguchi (RT) method. The second group consists of methods that predict future data values; the Taguchi method (T-method) and the Taguchi-Schmidt (TS) method belong to this group.

In this paper, we consider the T-method, which was proposed by Taguchi (2005). The T-method analyzes the same types of data as those suitable for multiple regression analysis (MRA). Because the T-method can be applied to problems in various fields, such as economic forecasting, demand forecasting, and weather forecasting, it has been used in many studies. One well-known case study is that of Yoshino et al. (2006). Theoretical aspects of the T-method were considered by Inou et al. (2012), and they proposed two improved T-methods.

For the application of the T-method, two or more items are prepared, and an estimated output value is calculated for each item. After that, by weighted integration with the corresponding signal-to-noise ratio (S/N), an overall estimated output value is obtained. Therefore, if the accuracies of the estimated values for each item can be increased, the accuracy of the integrated estimated value will also be increased. Hence, in this paper, we point out the features of estimated output values for each item in the T-method. And then, we use the generalized inverse regression estimator, which was proposed by Miwa (1985), to increase the accuracies of the values estimated by each item in the T-method.

This paper consists of five sections. In Section 2, we present an outline of the T-method and discuss some
previous studies. In Section 3, we focus on a method for predicting the estimated output value from each item in the T-method, and we propose an improved procedure. In Section 4, we present the results of Monte Carlo simulations and compare the performances of these methods. In Section 5, we present our conclusions.

2. The Calculation Procedures of the T-Method and Improved T-Methods

2.1 T-Method

In this section, we explain how to calculate an overall estimated output value in the T-method, using the methods presented in Taguchi (2005) and Tatebayashi et al. (2008). A sample data set is shown in Table 1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Item 1</th>
<th>Item 2</th>
<th>...</th>
<th>Item k</th>
<th>Output Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>...</td>
<td>$x_{1k}$</td>
<td>$y_1$</td>
</tr>
<tr>
<td>2</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>...</td>
<td>$x_{2k}$</td>
<td>$y_2$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$n$</td>
<td>$x_{n1}$</td>
<td>$x_{n2}$</td>
<td>...</td>
<td>$x_{nk}$</td>
<td>$y_n$</td>
</tr>
</tbody>
</table>

Step 1: Determination of the Unit Space

One or more sets of sample data are defined as the unit space data. The data of the unit space must be chosen from those data which the output values are contained within the space defined by the sample data. Let the number of sets in the unit space data be $a$. Each output value in the unit space is denoted as $y_i$ ($i = 1, 2, \ldots, a$). The value of each item in the unit space is denoted as $x_{ij}^w$ ($i = 1, 2, \ldots, a; j = 1, 2, \ldots, k$). For the sample data shown in Table 1, the unit space data are shown in Table 2. The average of the output values in the unit space is denoted as $\bar{y}$, and the average values of each item are denoted as $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$. The remaining data in the sample data are the signal space data. The number of sets of signal space data is $l = n - a$. Each output value in

<table>
<thead>
<tr>
<th>No.</th>
<th>Item 1</th>
<th>Item 2</th>
<th>...</th>
<th>Item k</th>
<th>Output Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$x_{11}^w$</td>
<td>$x_{12}^w$</td>
<td>...</td>
<td>$x_{1k}^w$</td>
<td>$y_1^w$</td>
</tr>
<tr>
<td>2</td>
<td>$x_{21}^w$</td>
<td>$x_{22}^w$</td>
<td>...</td>
<td>$x_{2k}^w$</td>
<td>$y_2^w$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$a$</td>
<td>$x_{aw1}$</td>
<td>$x_{aw2}$</td>
<td>...</td>
<td>$x_{awk}$</td>
<td>$y_a^w$</td>
</tr>
<tr>
<td>Average Values</td>
<td>$\bar{x}_1$</td>
<td>$\bar{x}_2$</td>
<td>...</td>
<td>$\bar{x}_k$</td>
<td>$\bar{y}$</td>
</tr>
</tbody>
</table>

Table 3: Signal Space

<table>
<thead>
<tr>
<th>No.</th>
<th>Item 1</th>
<th>Item 2</th>
<th>...</th>
<th>Item k</th>
<th>Output Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>...</td>
<td>$x_{1k}$</td>
<td>$y_1$</td>
</tr>
<tr>
<td>2</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>...</td>
<td>$x_{2k}$</td>
<td>$y_2$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$l$</td>
<td>$x_{l1}$</td>
<td>$x_{l2}$</td>
<td>...</td>
<td>$x_{lk}$</td>
<td>$y_l$</td>
</tr>
</tbody>
</table>
Applying GIR Estimator to T-method, Kawada et al.

signal space is denoted as \( y_i' (i = 1, 2, \ldots, l) \). The value of each item in the signal space data is denoted as \( x_{ij} \) \((i = 1, 2, \ldots, l; j = 1, 2, \ldots, k)\). The signal space data for this example are shown in Table 3.

**Step 2: Normalization of the Signal Space**

Each of the values of the signal space is normalized by subtracting the average values of the unit space \((\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k, \bar{y})\) from them, as shown Equations (1) and (2):

\[
X_{ij} = x_{ij}' - \bar{x}_j , \quad (1)
\]

\[
M_i = y_i' - \bar{y} . \quad (2)
\]

The value of the items in the normalized signal space data are denoted by \( X_{ij} (i = 1, 2, \ldots, l; j = 1, 2, \ldots, k) \), and the output values in the normalized signal space data are denoted by \( M_i (i = 1, 2, \ldots, l) \). The normalized signal data are shown in Table 4.

![Table 4: Normalized Signal Space](image)

**Step 3: Calculation of \( \beta \) and \( \eta \) for Each Item**

We denote by \( \beta_j \) the proportionality constant between an item \( j \) and its output value, and the corresponding S/N is denoted by \( \eta_j \). The values of \( \beta_j \) and \( \eta_j \) \((j = 1, 2, \ldots, k)\) are calculated as follows:

\[
\beta_j = \frac{\sum_{i=1}^{l} M_i X_{ij}}{r} , \quad (3)
\]

\[
\eta_j = \begin{cases} 
\frac{1}{r} (S_{\beta_j} - V_{\epsilon j}) & (S_{\beta_j} > V_{\epsilon j}) \\
0 & (S_{\beta_j} \leq V_{\epsilon j})
\end{cases} , \quad (4)
\]

where

\[
r = \sum_{i=1}^{l} M_i^2 , \quad (5)
\]

\[
S_{\beta j} = \sum_{i=1}^{l} X_{ij}^2 , \quad (6)
\]

\[
S_{\beta j} = \frac{\left(\sum_{i=1}^{l} M_i X_{ij}\right)^2}{r} , \quad (7)
\]
Step 4: An Overall Prediction

Let us suppose new item values \( x_{new1}, x_{new2}, \cdots, x_{newk} \) are obtained, and \( M_{new} \) is unknown. By subtracting the average of the unit space data \( (\bar{x}_1, \cdots, \bar{x}_k, \bar{y}) \) from each of the new values, we obtain new normalized signal data \( \tilde{x}_{new1}, \tilde{x}_{new2}, \cdots, \tilde{x}_{newk} \). Estimated output values for each item are calculated as follows:

\[
\hat{M}_{new,j} = \frac{X_{new,j}}{\beta_j}.
\]  

(9)

An overall estimated output value \( \hat{M}_{new} \) is given by the weighted integration of every estimated output value \( \hat{M}_{new1}, \hat{M}_{new2}, \cdots, \hat{M}_{newk} \) with the individual corresponding S/Ns \( \eta_1, \eta_2, \cdots, \eta_k \).

\[
\hat{M}_{new} = \frac{\eta_1 \times \hat{M}_{new1} + \eta_2 \times \hat{M}_{new2} + \cdots + \eta_k \times \hat{M}_{newk}}{\eta_1 + \eta_2 + \cdots + \eta_k}.
\]  

(10)

2.2 Improved T-Method

Inou et al. (2012) proposed two improved T-methods, the Ta-method and Tb-method. The steps for normalization were revised in both of the improved methods. Inou et al. (2012) reported that, in many cases, the prediction accuracies of the Ta-method and Tb-method were better than that of the T-method. The normalization procedures in the Ta-method and Tb-method are as follows.

2.2.1 Ta-method

In the Ta-method, the signal space includes all of the sample data; this is done without first defining the unit space. Thus, the number of sets of data in the signal space is \( l = n \). The average values for each item and the output value are calculated, based on all of the sample data. By subtracting the average value of each item and its output value from the signal space data, the normalized signal data are obtained. In the Ta-method, it is possible to fit a proportional line through the origin for each item and its output value.

2.2.2 Tb-method

As in the Ta-method, in the Tb-method, the signal space includes all of the sample data, which again implies that the number of sets of data in the signal space is \( l = n \). In the Tb-method, the normalized signal data are obtained by subtracting the sample that maximizes the S/N in each item from the initial value. The data used for the normalization are determined independently in each item. Since the S/N indicates the strength of the signal relative to that of the noise, using, the Tb-method increases the linearity between the items and their output values. There is a possibility that during normalization, different samples are selected for different items.

3. Inverse Estimation in the T-method

3.1 The values estimated by each item in the T-method

When using multiple regression to analyze the data shown in Table 1, it is assumed that there are causal relations between the explanatory variables \( x_1, x_2, \cdots, x_j \) and an objective variable \( y \). On the other hand, in the T-method, the Ta-method, and the Tb-method, an inverse estimate is obtained for each item, by simple linear regression, and an integrated output value is predicted by unifying these results. The T-method, the Ta-method and the Tb-method assume that each item has the structure \( X_{ij} = \beta_j M_i \) \( (j = 1, 2, \cdots, k) \) and that \( \beta_j \) is to be estimated. An inverse estimate is then obtained by using the newly obtained items and the normalized values, \( X_{new, j} \) and \( M_{new, j} \) are calculated for every item, as in Equation (9).

Obtaining the inverse estimate by formula conversion is called the classical method of inverse regression, and it is widely used. However, Williams (1969) showed that the estimate produced by the classical estimator is unstable, and its mean square error (MSE) could be infinite, since it is the ratio of two random variables that
follow normal distributions.

### 3.2 Generalized inverse regression estimator

We now consider the issue of calibration as it applies to single regression analysis. For linear calibration, Miwa (1985) proposed the generalized inverse regression (GIR) estimator, which is an improvement over both the classical estimator and the Krutchkoff (1967) estimator. Below, we present a procedure for using the GIR estimator in regression analysis; this is based on Miwa (1985) and Yamada (1994). We note that in linear calibration, it is assumed that the level of each item will be determined by using newly obtained output values. To begin, \( p \) pairs of \( Y \) and \( X \) are observed:

\[
Y_i = a + bX_i + \epsilon_i ,
\]

where \( \epsilon_i \sim N(0,\sigma^2) \), and \( i \) is the number of samples \( (i = 1,2,\ldots,p) \). Next, the unknown \( X_0 \) is estimated by using \( q \) new observations \( Y_{0j} \):

\[
Y_{0j} = a + bX_0 + \epsilon_{0j} ,
\]

where \( \epsilon_{0j} \sim N(0,\sigma^2) \), and \( j \) is the number of new samples \( (j = 1,2,\ldots,q) \). The classical estimators, based on the inverse estimate of \( X_0 \) from \( q \) new observations of \( Y_0 \), is defined to be

\[
\hat{X}_0C = \frac{\overline{Y}_0 - \hat{a}}{\hat{b}} = \overline{X} + \frac{\overline{Y}_0 - \overline{Y}}{\hat{b}} .
\]

The classical estimator is derived by the conversion of the formula \( Y = \hat{a} + \hat{b}X \). On the other hand, the GIR estimator is

\[
\hat{X}_{0G} = \overline{X} + \frac{\hat{b}}{\hat{b}^2 + c\hat{\phi}}(\overline{Y}_0 - \overline{Y}) ,
\]

where

\[
\hat{\phi} = \frac{\hat{\sigma}^2}{S_{XX}} ,
\]

\[
\hat{\sigma}^2 = \frac{S'_{c}}{\phi} ,
\]

\[
S'_{c} = \sum_{i=1}^{p} (Y_i - \hat{a} - \hat{b}X_i)^2 + \sum_{i=1}^{q} (Y_{0j} - \overline{Y}_0)^2 ,
\]

\[
\phi = p - 2 + q - 1 ,
\]

and \( c \) is a constant. We should let \( c = 4\phi/(\phi + 2) \) to minimize the MSE of \( \hat{X}_{0G} \), and we should let \( c = 1 \) to minimize the mean absolute value error of \( \hat{X}_{0G} \). Miwa (1985) showed that the MSE of the estimated value based on the GIR estimator is smaller than that based on the classical estimator.

### 3.3 Application to the T-method

The GIR estimator can be used as a linear calibration when new data are obtained. As in Section 3.2, the inverse estimation was initially intended to be used to estimate the level of the factors when new experimental
data became available.

On the other hand, in the T-methods, the inverse estimation is performed for the purpose of predicting new data from each item. When a new data set \( (x_{new1}, x_{new2}, \ldots, x_{newk}) \) is obtained, the T-method and improved T-methods should predict one output value \( (M_{new}) \) for each item.

In this section, we extend the idea of Miwa (1985), and we use the GIR estimator to predict the output value for each item in the T-methods. We expect to improve the integrated estimated output value by improving the estimated output values from each of the items. The following procedures are explained as they apply to the T-method, and they are also applied to the Ta-method and Tb-method.

In order to apply the GIR estimator to the T-method, we first define the model assumed to be behind the T-method, as follows (Inou et al. 2012):

\[
X_{ij} = \beta_j M_i + \varepsilon_{ij} \quad (i = 1, 2, \ldots, l; \ j = 1, 2, \ldots, k),
\]

(19)

where \( M_i \sim N(\theta, \tau^2) \), and \( \varepsilon_{ij} \sim N(0, s^2_i) \). First, we suppose that \( n \) pairs of \( X \) and \( M \) are observed. After dividing the sample data into unit space and signal space, \( \hat{\beta}_j \) and \( \eta_j \) are calculated for each item value and output value in the normalized signal space. Next, we suppose that new item values \( (x_{new1}, x_{new2}, \ldots, x_{newk}) \) in accordance with the assumed model are obtained and are normalized using the averages of the unit space data. The new normalized signal data are \( X_{new1}, X_{new2}, \ldots, X_{newk} \). The inverse estimator of each item in the T-method, as shown in Equation (9), is changed to

\[
\hat{M}_{newj} = \frac{\hat{\beta}_j}{\hat{\beta}_j^2 + c \hat{V}(\hat{\beta}_j)} X_{newj},
\]

(20)

where

\[
\hat{V}(\hat{\beta}_j) = \frac{s^2_j}{r},
\]

(21)

\[
s^2_j = \frac{\hat{S}_{\hat{\beta}_j}^2}{\phi},
\]

(22)

\[
\hat{S}_{\hat{\beta}_j} = \hat{S}_{\hat{\beta}_j} + X^2_{newj},
\]

(23)

\[
\phi = l + 1 - 1,
\]

(24)

and \( c \) is a constant. Based on Miwa (1985), we will let \( c = 4\phi/(\phi + 2) \) to minimize the MSE of the output estimator of each item, and we will let \( c = 1 \) to minimize the mean absolute value error of the output estimator of each item.

In this way, we use the new estimator based on the GIR estimator for each item in the T-method. The remaining calculations are the same as those in the original T-method. In this way, it is possible to overcome the weaknesses discussed in Section 3.1; moreover, the fluctuations in the new data are incorporated into the predictions for each item.

4. Comparison by Monte Carlo simulations

4.1 Methods

We compared the performances of the methods listed in Table 5. MRA was added in order to compare the performances.
Applying GIR Estimator to T-method, Kawada et al.

### Table 5: The Methods Used

<table>
<thead>
<tr>
<th>Method</th>
<th>Methods used to predict the value from each item</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-method 1</td>
<td>Method 1 includes the various normal T-methods.</td>
</tr>
<tr>
<td>Ta-method 1</td>
<td></td>
</tr>
<tr>
<td>Tb-method 1</td>
<td></td>
</tr>
<tr>
<td>T-method 2</td>
<td>Method 2 includes the T-methods that use the GIR estimator in the predictions from each item, with $c = 1$.</td>
</tr>
<tr>
<td>Ta-method 2</td>
<td></td>
</tr>
<tr>
<td>Tb-method 2</td>
<td></td>
</tr>
<tr>
<td>T-method 3</td>
<td>Method 3 includes the T-methods that use the GIR estimator in the predictions from each item, with $c = 4\phi/(\phi + 2)$.</td>
</tr>
<tr>
<td>Ta-method 3</td>
<td></td>
</tr>
<tr>
<td>Tb-method 3</td>
<td></td>
</tr>
</tbody>
</table>

#### 4.2 Settings

(A) Model A

Simulation model A is

$$y_j = 2x_{j1} + 2x_{j2} + 2x_{j3} + \varepsilon_j \quad (i = 1,2,\cdots,n ; j = 1,2,3),$$

where $x_{ij} \sim N(1,i^2)$, $\varepsilon_i \sim N(0,\sigma^2)$, and $n$ is the number of samples (5 or 20). We assumed $\sigma = 1.0,3.0,5.0$.

(B) Model B

Simulation model B is

$$y_j = 2y_i + \varepsilon_j \quad (i = 1,2,\cdots,n ; j = 1,2,3),$$

where $y_i \sim N(1,i^2)$, $\varepsilon_i \sim N(0,\sigma_i^2)$, and $n$ is the number of samples (5 or 20). We let $\sigma = (\sigma_1,\sigma_2,\sigma_3) = (1.0,1.0,1.0),(3.0,3.0,3.0),(5.0,5.0,5.0)$.

The simulations were conducted as follows:

1. Fix a model and parameters.
2. Generate the data according to the model, and calculate $\eta$ and $\beta$ for each of the T-methods.
3. Generate new data ($X_{new1},X_{new2},\cdots,X_{newk},y_{new}$) in accordance with the model, and predict the output value $y_{new}$ by using the values of each of the item $x_{new1},x_{new2},\cdots,x_{newk}$, for each of the T-methods.
4. Compare the results of 5,000 runs to estimate the precision.

We evaluated the prediction error ($PE$) as follows:

$$PE = E[(\hat{y}_{new} - y_{new})^2].$$

$PE$ is thus an index that measures prediction accuracy, and a smaller value indicates a better performance. The $PE$s for each technique are then compared in order to evaluate the prediction accuracies of each technique.

#### 4.3 Results

The performance comparisons are shown in Figs. 1 through 6.

From Figs. 1 through 6, we see that for both models, using the GIR estimator (methods 2 and 3) improves the total prediction accuracies compared to those of the original T-method (method 1). When the number of samples is small compared to the number of items, Inou et al. (2012) pointed out that the prediction accuracies of the various T-methods are better than those of MRA. Our results also show the same tendency. Furthermore, by using the procedure introduced in this paper, the accuracies are further improved.

The prediction accuracy is slightly better when using $c = 4\phi/(\phi + 2)$ (method 3) than it is when using $c = 1$ (method 2). Furthermore, even when the number of samples is much larger than the number of items, the...
Fig. 1: Results of the T-method with model A (N is the number of samples)

Fig. 2: Results of the Ta-method with model A (N is the number of samples)

Fig. 3: Results of the Tb-method with model A (N is the number of samples)

Fig. 4: Results of the T-method with model B (N is the number of samples)
prediction accuracy of method 3 is comparable to the prediction accuracy of MRA. The degree of improvement in the prediction accuracy of the Tb-method is smaller than those of the T-method or the Ta-method. One reason for this is that in the Tb-method, good linearity for each item is achieved by the normalization prior to using the GIR estimator for each item. We note that similar trends are seen with the other data patterns.

5. Conclusions

In this paper, we have pointed out the statistical features of estimators for each item in the T-methods, and apply the GIR estimator to the T-methods to overcome the weaknesses. The results presented in Section 4 show that introducing the GIR estimator into the calculations of the T-methods has good effects. In future studies, we plan to further investigate the features of our methods, by using mathematical approaches and using realistic data sets. We note that in Section 4, when introducing the GIR estimator, the accuracy of the predictions depends on the value of the constant $c$. Therefore, as another area of future work, we intend to develop a method for determining the value of $c$.

References:


[DOI : 10.17929/tqs.1.12] Copyright © 2015 Journal of the Japanese Society for Quality Control. All rights reserved.


**Acknowledgements**

We would like to thank the anonymous referees for their valuable comments. This work was partly supported by JSPS Grants-in-Aid for Scientific Research Grant Number 24500351.

**Authors’ biographical notes**

Hiroshi Kawada is a graduate student in Department of Industrial and Management System Engineering, Graduate School of Creative Science and Engineering at Waseda University.

Yasushi Nagata is a professor in Department of Industrial and Management System Engineering, School of Creative Science and Engineering at Waseda University.

[DOI : 10.17929/tqs.1.12]

Received, May 20, 2014
Revised: November 27, 2014
Accepted: January 26, 2015

Copyright © 2015 Journal of the Japanese Society for Quality Control. All rights reserved.