Research Article

Binary Bitwise Artificial Bee Colony as Feature Selection Optimization Approach within Taguchi's T-Method

Nolia Harudin [ID],[1] **Faizir Ramlie** [ID],[2] **Wan Zuki Azman Wan Muhamad** [ID],[3] **M. N. Muhtazaruddin** [ID],[2] **Khairur Rijal Jamaludin** [ID],[2] **Mohd Yazid Abu** [ID],[4] **and Zulkifli Marlah Marlan** [ID][2]

[1]*Department of Mechanical Engineering, Universiti Tenaga Nasional, Kajang 43000, Selangor, Malaysia*
[2]*Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, Kuala Lumpur 54100, Malaysia*
[3]*Institute of Engineering Mathematics, Universiti Malaysia Perlis, Kampus Pauh Putra, Arau 02600, Perlis, Malaysia*
[4]*Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, Pekan 26600, Malaysia*

Correspondence should be addressed to Wan Zuki Azman Wan Muhamad; wanzuki@unimap.edu.my

Taguchi's T-Method is one of the Mahalanobis Taguchi System- (MTS-) ruled prediction techniques that has been established specifically but not limited to small, multivariate sample data. The prediction model's complexity aspect can be further enhanced by removing features that do not provide valuable information on the overall prediction. In order to accomplish this, a matrix called orthogonal array (OA) is used within the existing Taguchi's T-Method. However, OA's fixed-scheme matrix and its drawback in coping with the high-dimensionality factor led to a suboptimal solution. On the contrary, the usage of SNR (dB) as its objective function was a reliable measure. The application of Binary Bitwise Artificial Bee Colony (BitABC) has been adopted as the novel search engine that helps cater to OA's limitation within Taguchi's T-Method. The generalization aspect using bootstrap was a fundamental addition incorporated in this research to control the effect of overfitting in the analysis. The adoption of BitABC has been tested on eight (8) case studies, including large and small sample datasets. The result shows improved predictive accuracy ranging between 13.99% and 32.86% depending on cases. This study proved that incorporating BitABC techniques into Taguchi's T-Method methodology effectively improved its prediction accuracy.

1. Introduction

Taguchi's T-Method, which was explicitly developed for predictive analysis, is one of the Mahalanobis Taguchi System's (MTS) variants that has been increasingly used by researchers and industrial practitioners in Japan and other countries. Taguchi's T-Method was proposed for multivariate estimation to predict the integrated estimated output value. In the 1980s, Dr. Genichi Taguchi developed the Mahalanobis Taguchi System (MTS) as a pattern recognition technique that blends Mahalanobis Distance (MD) theory and Taguchi Robust Engineering concept to systematically and effectively classify and predict data in a multidimensional environment [1–6].

MTS establishes a multivariate measurement scale that recognizes a normal or healthy observation from an abnormal or an unhealthy observation and integrates it with the concept of signal-to-noise ratio (SNR) and orthogonal array (OA). Beginning with the introduction of the MT-Method as a classification technique that has so far gained much attention among scholars [7–14], Taguchi's T-Method has been established since then, which has utilized the same integration principles. The unit-space concept, the duplicate signal-to-noise ratio (SNR) adaptation as a weighting factor, zero-proportional theory, and OA as the feature selection optimization are the main elements that have been adopted in reinforcing Taguchi's T-Method robustness.

One of Taguchi's T-Method significant advantages is its ability to predict even with limited sample data. In multiple regression analyses, a limitation exists in which the sample size has to be higher than the number of variables. On the contrary, the said limitation does not apply to Taguchi's T-Method. Additionally, Taguchi's T-Method has no direct influence from multicollinearity since individual regression has been considered [2, 15, 16]. Based on the number of papers published in the literature, Taguchi's T-Method studies' progress is moving towards optimizing parameters and optimizing feature selection rather than just application purposes since the year 2012 [17–20]. The increasing pattern has indirectly triggered that there are indeed a variety of enhanced approaches towards parameter and feature selection optimization available out there that can be further explored and incorporated into Taguchi's T-Method as a hybridization or integration element.

1.1. Taguchi's T-Method for the Feature Selection Optimization Problem. In MTS, the orthogonal array (OA) is a feature selection search mechanism that has been established between a series of MTS, including Taguchi's T-Method, which share standard procedures but vary in their objective function determination. The OA element within MTS has been debated and is believed to be insufficient as it offers a suboptimal solution [21, 22]. Most OA's concerns are based on its restriction in having appropriate combinations of features to be assessed and evaluated in the search for optimality, as it relies on a fixed scheme [20, 23]. The authors of [24] argued that the fixed combination in OA is not optimal since the results may vary significantly if the column-to-column information is rearranged [6]. In [25], the authors agreed with the authors of [24] after proving 1000 random variables to the column assignment. Issues in OA have been highlighted as well by [26, 27], especially the fact that the OA design has a limitation in handling the higher-order interaction between variables, which might lead to an inconsistency in the identification of the significant variables [24, 25, 27–29]. Therefore, developing a hybrid methodology for better accuracy is a preferred solution to these concerns that drove this research's primary motivation.

Until recently, the OA element in the MTS classification approaches has been continuously improved by numerous machine learning algorithms. However, enhancing the OA element within Taguchi's T-Method as a prediction tool is still at an initial stage. In [30], the authors applied a stepwise forward and backward selection procedure for this purpose which showed an increase in accuracy in many cases conducted [30]. The author of [31] suggested a Binary Artificial Bee Colony (BABC) algorithm, and the findings revealed that T-Method + BABC worked better than T-Method + OA in a particular case study conducted [31]. The most recent reported study by [32] has specifically addressed OA's downside and suggested Binary Particle Swarm Optimization (BPSO), which indicates an increase in accuracy for specific case studies [32]. The published literature on OA improvement in Taguchi's T-Method is found not utilizing the generalization aspect thoroughly and focused on a somewhat limited case study. The previous research by [31, 32] was further expanded in this study by proposing the other variant of binary ABC called Binary Bitwise ABC algorithms with proper generalization aspect been amended into it, which is the application of bootstrap cross-validation.

2. Methodology

2.1. Taguchi's T-Method. Regression analysis aims to construct a mathematical model that describes and explains the relationship between variables for prediction or a study of causal relationships [33]. Taguchi's T-Method, which is driven by similar purposes, was built to forecast the unknown value of the output variable concerning the established value of the input variables by statistically evaluating the relevant correlation and functional relationship between those variables through a specific developed linear regression model to compute the integrated estimate output value. The integrated estimate output model in Taguchi's T-Method consists of some additional elements that differentiated it from standard linear regression: (1) zero-point proportional term, (2) inverse regression model, (3) unit-space concept, and (4) weightage SNR. All these elements have been embedded into the existing Taguchi's T-method model described by [34] to generate the specified integrated estimated model, as shown in equation (1). Taguchi's T-Method as well utilizes the ordinary least squares approach to calculate the proportional coefficient, β which is a common approach in linear regression. Equations (2)–(7) govern the inclusion of dynamic SNR as a weightage factor for each feature within the model [35]:

$$\text{Integrated estimate value } \widehat{M}_{ii} = \frac{\eta_1 \check{M}_{ii,1} + \eta_2 \check{M}_{ii,2} + \ldots + \eta_D \widehat{M}_{ii,D}}{\eta_1 + \eta_2 + \ldots + \eta_D}, \text{ where } ii \text{ (number of sample)}$$

(1)

$$= 1, 2, \ldots, l \text{ and } \widehat{M} = \frac{Z_{ii,j}}{\beta_j}, \text{ while } Z \text{ is the normalized input variables,}$$

$$\text{Effective divider}, r = M_1^2 + M_2^2 + \ldots + M_l^2,$$

(2)

$$\text{Total variation}, S_{Tj} = Z_{1j}^2 + Z_{2j}^2 + \ldots + Z_{lj}^2 \text{ for } j \text{ (number of features)} = 1, 2, \ldots, D,$$

(3)

$$\text{variation of proportional term, } S_{\beta j} = \frac{\left(M_1 Z_{1j} + M_2 Z_{2j} + \ldots + M_l Z_{lj} \right)^2}{r}, \tag{4}$$

$$\text{Error variation, } S_{ej} = S_{Tj} - S_{\beta j}, \tag{5}$$

$$\text{Error variance, } V_{ej} = \frac{S_{ej}}{l-1}, \tag{6}$$

$$\text{SNR, } \eta = \begin{cases} \dfrac{1/r\left(\text{Variation of proportional term, } S_{\beta j} - \text{Error variance, } V_{ej}\right)}{\text{Error variance, } V_{ej}}, & \left(\text{when } S_{\beta j} > V_{ej}\right), \\[2mm] 0, & \left(\text{when } S_{\beta j} \leq V_{ej}\right). \end{cases} \tag{7}$$

It is seen that the higher SNR of an item will contribute to a greater degree of contribution to overall model estimation. The integrated estimate SNR (dB) is computed based on the result obtained using equation (1). The integrated estimate SNR, η (dB), is a performance measure to evaluate the input variable's relative importance towards the output variable. To further increase the model accuracy, optimization concerning the selection of features is considered a value-added approach within Taguchi's T-Method. Equations (8)–(13) are used for calculating the SNR (dB) for feature selection optimization, which as shown below. The evaluation of the relative importance of features is conducted using the two-level orthogonal array (OA). OA with a predetermined combination of "use" and "not use" of features allows for comparison of integrated estimate SNR (dB) under the setting. Table 1 shows the example of L_{12} orthogonal array with Level 1 in the array indicates that the variable will be used, while Level 2 indicates that the variable will not be used during the simulation study. Evaluation of relative importance of features is performed by computing the new integrated estimate SNR (dB) when the features are not used in computation and observed the increment or deterioration of the value. A higher integrated estimate SNR (dB) value is preferred, and a combination of input variables that yields optimal integrated estimate SNR (dB) is selected as an optimal combination:

$$\text{Linear equation, } L = M_1 \widehat{M}_1 + M_2 \widehat{M}_2 + \ldots + M_l \widehat{M}_l, \tag{8}$$

$$\text{Variation of proportional term, } S_\beta = \frac{L^2}{r}, \tag{9}$$

$$\text{Total variation, } S_T = \widehat{M}_1^2 + \widehat{M}_2^2 + \ldots + \widehat{M}_l^2, \tag{10}$$

$$\text{Error variation, } S_e = S_T - S_\beta, \tag{11}$$

$$\text{Error variance, } V_e = \frac{S_e}{l-1}, \tag{12}$$

$$\text{Integrated estimate } SN \text{ ratio, } \eta = 10 \log\left(\frac{1/r\left(S_\beta - V_e\right)}{V_e}\right), (dB), \tag{13}$$

2.2. Binary Bitwise Artificial Bee Colony (BitABC) into Taguchi's T-Method for Feature Selection Optimization. This research's binary approach is similar to the orthogonal array (OA) concept in existing Taguchi's T-Method. The Binary ABC was explicitly developed for the feature selection optimization process by changing the information of each identified food source update to the discrete-binary data type to be "1" or "0." The primary food source *(Xi)* is randomly initialized by following the identified bee's population size (NP/2 = N) and the total number of features (D) using discrete-binary data (1 or 0). The primary objective function, which is to maximize the SNR (dB) value, is then computed. The best SNR (dB) are selected as Global_max and its binary combination as Global_para. The employed bees will continue searching for a better food source, which will make a little change based on their nearby information memory and create a new source. The objective function, SNR (dB) value, is computed then and been compared to primary sources. The higher SNR (dB) value will be memorized, while the lower will be forgotten. If the previous SNR (dB) value is

TABLE 1: L_{12} orthogonal array combination.

No.	Control factors											Integrated estimate SNR (dB)
	A	B	C	D	E	F	G	H	I	J	K	
1	1	1	1	1	1	1	1	1	1	1	1	SNR1
2	1	1	1	1	1	2	2	2	2	2	2	SNR2
3	1	1	2	2	2	1	1	1	2	2	2	SNR3
4	1	2	1	2	2	1	2	2	1	1	2	SNR4
5	1	2	2	1	2	2	1	2	1	2	1	SNR5
6	1	2	2	2	1	2	2	1	2	1	1	SNR6
7	2	1	2	2	1	1	2	2	1	2	1	SNR7
8	2	1	2	1	2	2	2	1	1	1	2	SNR8
9	2	1	1	2	2	2	1	2	2	1	1	SNR9
10	2	2	2	1	1	1	1	2	2	1	2	SNR10
11	2	2	1	2	1	2	1	1	1	2	2	SNR11
12	2	2	1	1	2	1	2	1	2	2	1	SNR12

(Note: 1 denotes "Item will be used," 2 denotes "Item will not be used").

higher than the existing candidates, the value will remain. This decision process is called greedy selection. The employed bees will then share the information on the new position to onlooker bees once they return to their hive in the dance area. The onlooker bees will then evaluate the new position and choose to emphasize the food source's information, relying on the probability rate calculated. The onlooker bees will modify the position if the criteria are fulfilled, and SNR (dB) amount will be recalculated and updated following greedy selection criteria. The employed bee that cannot improve their position up to the defined limit will be abandoned and become a scout bee. A scout bee will randomly search for a new food source near the surrounding area of its hive. The cycle is repeated until it reaches the maximum number of cycles. The Global_max and Global_para at the maximum cycle are updated accordingly.

The method used by the bees (employed and onlooker) to search for the new food source which having more nectar amount within its neighborhood are following the approach introduced by Jia, Duan, and Khan [36] called binary Bitwise ABC (BitABC). Bitwise operators often transform an image into a binary number and represent a series of 0s and 1s on the computer [36]. However, only the logic operator results are adopted in the study conducted by Jia, Duan, and Khan [36], as it has similar characteristics with the binary space (0 and 1). The bitwise operator ($\wedge$, &, and |) to describe the trajectory of the food source within this study is illustrated in Table 2 and equations (14) and (15).

$$v_{ij} = X_{ij} \wedge \varnothing_{ij} \& \left(X_{ij} | X_{kj} \right), \quad i = \text{number of bees, while } j = \text{number of features,} \tag{14}$$

$$\text{where, } \varnothing_{ij} = \begin{cases} 1 & \text{rand}(0,1) < (r = 0.5), \\ 0 & \text{rand}(0,1) \geq (r = 0.5). \end{cases} \tag{15}$$

2.3. Data Preparation and Selection. The optimum features are selected based on the total number of use items ("1") produced by each feature across the run's total number. The combination of use item ("1") at each run represents the combination features contributed to the most optimum SNR (dB) value across the maximum cycle iteration. In demonstrating the proposed algorithm's stability and consistency, 70% of the training dataset from 20 different independent runs were set, and features that appear to be selected more than 10 times (more than 50%) are selected as the optimum features. The optimum features will be used to validate the remaining 30% validation dataset. The 70% training dataset follows the bootstrap cross-validation analysis during the training phase, which segregates the training and test set into 63.2% and 36.8%, respectively, with 1000 bootstrap cycles. The risk of overfitting is being considered and monitored accordingly within this study.

For better comparative purposes, despite the current Taguchi's T-method, the outcome of Bitwise ABC's optimum features has also been compared to another metaheuristic algorithm variant called Probability Binary Particle Swarm Optimization (PBPSO) [32] as well as the existing Taguchi's T-Method with full features and Taguchi's T-Method with optimal features provided by OA analysis [35]. Several simulations were performed on eight real-world datasets on prediction and regression with multivariate cases in assessing the suggested algorithm. Six out of eight datasets were obtained from the University of California at Irvine (UCI) Machine Learning Benchmark Repository [37]. The other two datasets were taken from the actual case study.

Both the BitABC and PBPSO are being set by the parameter configuration listed in Table 3. The optimization of all the algorithms within this study was constructed using Matlab R2018a application software. The programming algorithm compiled on 64 bits Sony VAIO VPCCA notebook with Intel i5 (2.3 GHz) 4 Gigabytes RAM capability and 212 GB data storage. The pseudocode of the proposed BitABC algorithm into Taguchi's T-Method is shown in Figure 1.

2.4. Performance Measure. Prediction is an iteration method involving model creation before performance evaluation, then proceeds to repeat the cycle until a satisfying solution is encountered. Throughout this study, two performance criteria are used to evaluate the developed algorithm's performances: the prediction accuracy and convergence rate of training, testing, and validation dataset.

In machine learning, especially on the regression analysis, the standard prediction error performance measures are computed using the mean absolute error (MAE), root

TABLE 2: Bitwise logical index sample operation.

		Bitwise operator and description		
		And a & b	Or a | b	Exclusive or aˆb
Input				
A	B		Results	
1	1	1	1	0
0	1	0	1	1
1	0	0	1	1
0	0	0	0	0

TABLE 3: Parameter setting for all BitABC and PBPSO algorithm.

Parameter	BABC	BPSO
Population size, NP	50	30
Acceleration coefficient, C1	NA	2
Acceleration coefficient, C2	NA	2
Inertia weight, w_{min}	NA	0.4
Inertia weight, w_{max}	NA	0.9
Maximum number of cycle	500	100
Limit	100	NA
V_{max}	NA	5
Number of run		20
Training dataset		70%
Validation dataset		30%

Note. NA: not applicable.

mean squared error (RMSE), mean absolute percentage error (MAPE), and several others. In practice, the regression prediction model accuracy must be estimated over the training and validation sets and are independent of one another. In this study, after the optimum features have been identified, the integrated estimate value, $\widehat{M}_{predicted}$ will be calculated as indicated by Equation (1). MAE formula was applied for the prediction model accuracy as shown in Equation (16). The MAPE measure has also been applied in this study to provide the final increment percentages of the optimal approach toward existing Taguchi's T-Method that uses full features, as shown by equation (17):

$$\text{MAE} = \frac{1}{\text{no.of sample}} \sum_1^n \left| M_{actual} - \widehat{M}_{predicted} \right|, \tag{16}$$

$$\text{MAPE} = 100 \left[\frac{1}{\text{no.of sample}} \sum_1^n \left| M_{actual} - \widehat{M}_{predicted} \right| / M_{actual} \right]. \tag{17}$$

3. Results and Discussion

The feature selection analysis findings are addressed according to the respective case studies presented in this research using the defined integrated estimate model shown by equation (1) previously. Despite focusing on the MAE results and its SD value, the discussion is also guided with several other performance measures such as the convergence plot of the SNR (dB) value as the objective function and also MAE for the training and testing phase. Table 4 and Figure 2

illustrate the example of the performance analysis for the heating load case study. Researchers often use this dataset to interact with several other techniques that rely on regression analysis [38, 39]. Similar procedures were applied to the remaining seven datasets applied within this study. The explanation of the heating load case study will provide a general idea of how the other case studies are analyzed in terms of their MAE trend for both training and testing, as well as the SNR (dB) convergence plot. The validation phase is summarizing the overall case studies considered within this research.

Table 4 reveals that F2, F3, F6, F7, and F8 are the dominant features for both *T* Method-BitABC and *T* Method-PBPSO. The T-Method with OA shows conflicting results, with F1 identified as one of the dominant features instead of F3 and F8 as other methods.

In providing a more explicit description of how each outcome reflects the overall prediction analysis, the effects of the SNR (dB) and MAE for the training and testing are illustrated by the convergence plot shown in Figures 1(a) and 1(b). The result reveals that the *T* Method-BitABC is the most optimum approach with the highest SNR (dB) value compared to the *T* Method-PBPSO, *T* Method, and *T* Method-OA. The trend aligned with MAE's trend for the training and testing phase, with *T* Method-BitABC performing better prediction accuracy with lower MAE value than *T* Method-PBPSO.

As seen in Table 5, the validation phase results indicate the result of the trained model performance towards the validation dataset with the case studies having more than 30 sample data (large dataset), while Table 6 summarized for the case study having less than 30 sample data (small dataset). Table 5 indicates that the result of *T* Method-BitABC and *T* Method-PBPSO reflect the same MAE performance. This is possible due to similar optimal features' selection results gained from the training and testing phase. The improvement percentages range from 13.99% to 32.86% across three different case studies (Abalone, Heating, and Cooling). Body fat and Concrete Compressive Strength cases show that Taguchi's T-Method maintains the best compared to others, while *T* Method-OA is the best for the Auto MPG case study, which contributes to 45.71% improvement compared to Taguchi's T-Method. The trend for the small sample case studies is a little bit of contrast. The result for both *T* Method-BitABC and *T* Method-PBPSO seems to differ from each other. *T* Method-BitABC provides better performance for the JD dataset with 9.07% improvement compared to Taguchi's T-Method. *T* Method-OA provides the best result for the Chiller dataset with 9.54% improvement compared to Taguchi's T-Method.

The analysis results shared explicitly represent how well the *T* Method-BitABC approach is well reflected in several case studies. A few findings could be further investigated, which implicitly represent the analysis results identified. The findings shall be summarized as follows:

> The adoption of BitABC into existing Taguchi's T-Method replacing the OA is found not suitable for the body fat case study. Body fat is a case study with a normal distribution trend and has a stable output

Enumerate unit-space, η , β, r, l as fix value in evaluating the objective function (snr_est) and...
          ...evaluating the fitness values.
Input: *dataTraining (Z (ii,j) , M(ii,j)), N, D, limit,lb,ub,maxcycle and BS*
Output: a set of optimal solutions obtained by the algorithm

1   **For** bootstrap_runcycle = 1 : BS

2   Randomly **initialize N x D** food sources, $X_i = \begin{bmatrix} 1 & 1 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 1 & 1 & 0 & \dots & 1 \end{bmatrix}$ following uniform distribution [0,1]

3   Evaluate the functions of the initialized food source, and compute their fitness values

4   $snr\_est_i \ (maximize) = f(X_i) = 10 log_{10}\left[\frac{1}{r}\frac{S_\beta - V_e}{V_e}\right]$ *% using objective function pseudocode*

5   $Fit\_F_i = \begin{cases} \dfrac{1}{1+f(X_i)} \\ 1 + abs\, f(X_i) \end{cases}$

6   Set the initial Globe_max and Globe_para value   *% Max(snr_est) & the combination Xi at max index*

7   **While** *Stopping Criteria (**maxcycle**) not satisfied* **do**

          %% *Employed Bee phase*
8   **for** *i =1 : N*

                  r is set to 0.5, while k is the neighbouring location in the range 1 to N...
9   *... update the food source trajectory using bitwise operator sequentially...*
                        ...following uniform distribution [0,1] .

                  jrand=fix(rand(D))+1;*   and   *k=fix(rand*N)+1;*
                        if jrand >= D+1   while (k==i)
10   *jrand=D;   k=fix(rand*(N))+1;*
                        elseif jrand <=0   end
                           jrand=1;
                           end
11                  $v_{ijrand} = x_{ijrand} \wedge \emptyset_{ijrand} \& (x_{ijrand}|x_{kjrand})$ where, $\emptyset_{ijrand} = \begin{cases} 1 & rand(0,1) < r \\ 0 & rand(0,1) \geq r \end{cases}$

12   **end**

13   Calculate the SNR function and fitness as in step 4 and 5

14   Greedy selection is applied between current and candidate solution based on maximum…
          …function and minimum fitness value as in step 4 & 5

15   Keep the best solution between current and candidate solution, update *Xi* position
16   Calculate the probability , $NomFit_i = Fit\_F_i /\ sum\ (Fit\_F_i)$

          %% Onlooker Bee Phase
17   A random value between 0 to 1 is generated for an onlooker bee to compare with...
                  … the calculated probability, *Nomfit* value of a food source.
18   **if**   *rand < Nomfit (i)*
19   This food source (*i*) is selected by onlooker bee and step 8 to 12 is followed…
                              … for new food source
20   **end**

21   Greedy selection between current and candidate *snr_est* value based on maximum ...
                  … function and minimum fitness value as in step 4 & 5.
22   Keep the best solution between current and candidate solution, update ***Xi*** position
23   Update the Globe_max value *(the maximum SNR(dB))* and ...
          ...Global_para *(the best among Xi position with maximum SNR (dB))*

          %% Scout Bee Phase
24   If the counter value of a food source is the maximum among those of food sources and
                  … exceeds limit,
25   *[Max_bus ind] = max(trial);*
26   **if**   *trial(ind) > limit*
27   *trial(ind) = 0;*
28   *X(ind,:) = randi([0,1],1,D);*   *% a new food source for (ind) is created by a scout*
                  Evaluate the functions of the new food source, and compute the fitness values.
29   **end**   *% bee using binary random values 0 or 1.*
          cycle = cycle + 1;
31   **End While**
32 **End**
33          $optimum\_para = \sum\limits_{1}^{BS} Global\_para \geq 50\%\ BS$

Note: Mean Absolute Error (MAE) is calculated at each cycle to see the accuracy of Global_para in improving the model accuracy
          while optimum_para is tested using validation data.

FIGURE 1: The pseudocode of the proposed BitABC algorithm into Taguchi's T-Method.

TABLE 4: Times of the appearance of each feature over 20 runs for heating load.

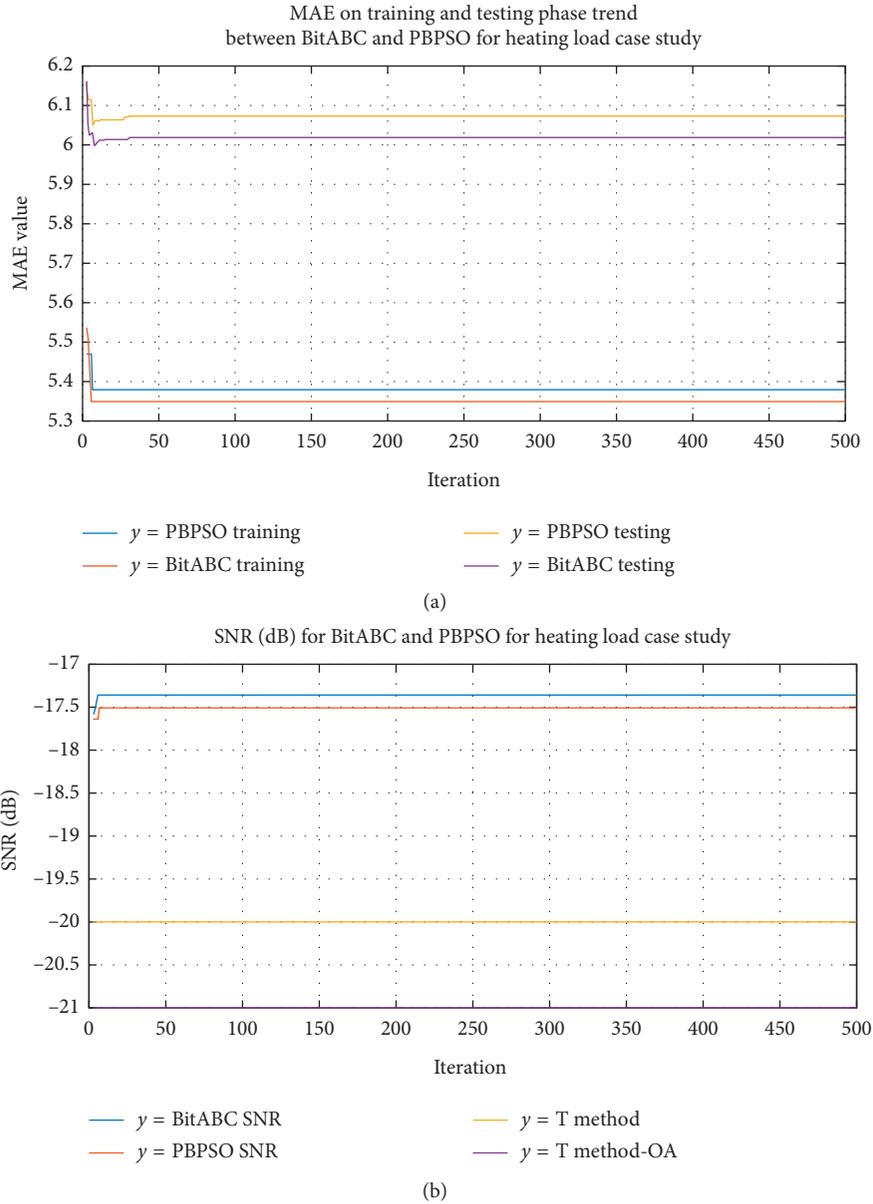| Methods	Heating load							
	F1	F2	F3	F4	F5	F6	F7	F8
T Method-PBPSO	1	19	20	0	0	11	20	10
T Method-BitABC	0	20	20	0	0	12	20	15
T Method-OA	√	√	x	x	x	√	√	x

(a)

(b)

FIGURE 2: (a) MAE for the training and testing phase between Bitwise ABC and PBPSO for the heating load case study; (b) SNR (dB) trend between Bitwise ABC and PBPSO for the heating load case study.

performance than other cases [40]. The adoption of feature selection optimization does not provide a better trend on this type of data since the combination features are already appropriate for the model.

The Concrete Compressive Strength dataset shows how the quality of the data within each analysis affects the

analysis result. By considering randomness and variation effect within datasets, it is possible to have slightly different trend results. From the result in Table 5, the slightly different trend between *T* Method, *T* Method-BitABC, and *T* Method-PBPSO shows that the proposed algorithm should provide a better deal since just

TABLE 5: The prediction accuracy (MAE) compilation across all proposed enhanced methods and existing Taguchi's T-Method on validation datasets for large sample data.

Dataset	# sample training	# sample testing	# features	Measure	T-Method [28]	T Method-BitABC and T Method-PBPSO [26]	T-Method + OA [28]	% MAE improvement of best result vs. T-Method
Body fat	176	76	14	MAE	0.3868	1.5663	1.651	NA
				SD	0.3250	1.0743	1.056	
				# optimum features	14	10	9	
Abalone	2924	1253	8	MAE	4.2377	3.6447	3.757	13.99%
				SD	3.3917	2.5686	2.791	
				# optimum features	8	3	4	
Heating	538	230	8	MAE	8.6033	5.7762	6.317	32.86%
				SD	8.5813	3.5191	3.831	
				# optimum features	8	5	4	
Cooling	538	230	8	MAE	8.1106	5.9515	6.312	26.62%
				SD	4.5612	4.0778	4.217	
				# optimum features	8	5	2	
Concrete Compressive Strength	721	309	7	MAE	11.4115	11.8847	12.382	NA
				SD	11.3301	12.3216	11.677	
				# optimum features	7	6	6	
Auto MPG	274	118	7	MAE	6.0035	3.5716	3.259	45.71%
				SD	3.0157	2.6609	2.717	
				# optimum features	7	2	5	

TABLE 6: The prediction accuracy (MAE) compilation across all proposed enhanced methods and existing Taguchi's T-Method on validation datasets for the small sample.

Dataset	# sample training	# sample testing	# Features	Measure	T-Method [28]	T Method-PBPSO [26]	T Method-BitABC	T-Method + OA [28]	% MAE improvement of best result vs. T-Method
Chiller	10	4	44	MAE	4.7925	5.3121	5.6858	4.3354	9.54%
				SD	6.2396	3.6373	4.3237	4.0135	
				# optimum features	44	12	14	16	
JD power	14	6	44	MAE	0.7866	0.9127	0.7153	0.7299	9.07%
				SD	0.7235	0.4276	0.5361	0.7431	
				# optimum features	44	24	27	23	

relying on 6 features instead of 7 total number of features. A similar situation occurs to the Auto MPG and Chiller case studies with T Method-BitABC, and T Method-PBPSO requiring fewer features compared to the T Method-OA with minimum MAE differences.

In this study, Taguchi's T-Method proved capable of computing a prediction analysis involving sample data much lower than the number of features than multiple linear regression that cannot compute the analysis within a similar state. This served as one of the main advantages of Taguchi's T-Method.

Adopting the BitABC replacing the OA within Taguchi's T-Method for small sample data with many features seems feasible, even though risk towards model accuracy still exists, requiring further monitoring. A considerable number of features are able to be reduced by implementing this approach. However, overfitting might be one of the risks to deal with for this small sample datasets' cases.

The adoption of BitABC seems not to differ from PBPSO for the large sample data within this study but varies for the small sample dataset. The better

exploration and exploitation search mechanism within the ABC algorithm might be the main reason for this trend since small sample data are susceptible to variation. The bootstrap, adopted as the cross-validation element, helps in reducing the risk of overfitting across training, testing, and validation dataset.

4. Conclusion

The adoption of BitABC into Taguchi's T-Method replacing the OA is shown feasible in this study. The result analysis shows that 4 out of 8 case studies reflect that BitABC adoption provides better performance than existing Taguchi's T-Method. The other case studies vary with minimal MAE differences and provide fewer significant features to be considered. Even though the trend result for both BitABC and PBPSO is similar for the large dataset, the small data samples reflected that BitABC provides much better prediction results. It was apparent that the merging of the BitABC into the current Taguchi's T-Method optimization technique to increase the SNR (dB) and predict the accuracy of the predicted integrated model was indeed practical. Further development studies should also focus on improving parameter estimates' robustness to ensure an established integrated estimated output model is reliable, especially for small sample data analysis.

Data Availability

Data are available within the repository of the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] G. Taguchi, S. Chowdhury, and Y. Wu, *The Mahalanobis-Taguchi System*, McGraw-Hill, New York, NY, USA, 1st edition, 2001.

[2] S. Teshima, Y. Hasegawa, and K. Tatebayashi, "Pattern recognition and the MT system," in *Quality Recognition and Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System*, pp. 1–13, Momentum Press, New York, NY, USA, 1st edition, 2012.

[3] F. Ramlie, W. Z. A. Wan Muhamad, K. R. Jamaludin, E. Cudney, R. Dollah, and R. Dollah, "A significant feature selection in the mahalanobis taguchi system using modified-bees algorithm," *International Journal of Engineering Research and Technology*, vol. 13, no. 1, pp. 117–136, 2020.

[4] W. Z. A. W. Muhamad, K. R. Jamaludin, F. Ramlie, N. Harudin, and N. N. Jaafar, "Criteria selection for an MBA programme based on the mahalanobis Taguchi system and the Kanri Distance Calculator," in *Proceedings of the 2017 IEEE 15th Student Conference on Research and Development (SCOReD)*, pp. 220–223, Kuala Lumpur Malaysia, December 2017.

[5] S. K. M. Saad, M. H. M. Razali, M. Y. Abu et al., "Optimizing the MFlex monitoring system using Mahalanobis-Taguchi system," *IOP Conference Series: Materials Science and Engineering*, vol. 1092, no. 1, Article ID 012009, 2021.

[6] N. N. N. M. Kamil, S. N. A. M. Zaini, and M. Y. Abu, "Feasibility study on the implementation of Mahalanobis-Taguchi system and time driven activity-based costing in electronic industry," *International Journal of Industrial Management*, vol. 10, pp. 160–172, 2021.

[7] H. Sakeran, N. A. Abu Osman, and M. S. Abdul Majid, "Gait classification using Mahalanobis-Taguchi system for health monitoring systems following anterior cruciate ligament reconstruction," *Applied Sciences*, vol. 9, no. 16, 16 pages, Article ID 3306, 2019.

[8] T. Asakura, W. Yashima, K. Suzuki, and M. Shimotou, "Anomaly detection in a logistic operating system using the mahalanobis-taguchi method," *Applied Sciences*, vol. 10, no. 12, pp. 4376–12, 2020.

[9] A. M. Yazid, J. K. Rijal, M. S. Awaluddin, and E. Sari, "Pattern recognition on remanufacturing automotive component as support decision making using Mahalanobis-Taguchi system," *Procedia CIRP*, vol. 26, pp. 258–263, 2015.

[10] N. N. Nik Mohd Kamil and M. Y. Abu, "Integration of Mahalanobis-Taguchi System and activity based costing for remanufacturing decision," *Journal of Modern Manufacturing Systems and Technology*, vol. 1, no. 1, pp. 39–51, 2018.

[11] J. Zhan, W. Chen, L. Cheng, Q. Wang, F. Han, and Y. Cui, "Diagnosis of asthma based on routine blood biomarkers using machine learning," *Computational Intelligence and Neuroscience*, vol. 20208 pages, 2020.

[12] X. Xiao, D. Fu, Y. Shi, and J. Wen, "Optimized mahalanobis-taguchi system for high-dimensional small sample data classification," *Computational Intelligence and Neuroscience*, vol. 202015 pages, 2020.

[13] S. Chen, Y. Liu, L. Wei, and B. Guan, "PS-FW: a hybrid algorithm based on particle swarm and fireworks for global optimization," *Computational Intelligence and Neuroscience*, vol. 201827 pages, 2018.

[14] M. El-Banna, "Modified mahalanobis taguchi system for imbalance data classification," *Computational Intelligence and Neuroscience*, vol. 201715 pages, 2017.

[15] K. Nishino and A. Suzuki, "Taguchi's T-method using median-median line for small sample with outliers," *IEEJ Transactions on Industry Applications*, vol. 138, no. 7, pp. 598–604, 2018.

[16] S. Negishi, Y. Morimoto, S. Takayama, and A. Ishigame, "Daily peak load forecasting by Taguchi's T method," *Electrical Engineering in Japan*, vol. 201, no. 1, pp. 57–65, 2017.

[17] N. Harudin, K. R. Jamaludin, M. Nabil Muhtazaruddin, F. Ramlie, and W. Z. A. W. Muhamad, "A feasibility study in adapting Shamos Bickel and Hodges Lehman estimator into T-Method for normalization," *IOP Conference Series: Materials Science and Engineering*, vol. 319, Article ID 012033, 2018.

[18] Z. M. Marlan, K. R. Jamaludin, F. Ramlie, N. Harudin, and N. N. Jaafar, "Determination of optimal unit space data for taguchi's T-method based on homogeneity of output," *Open International Journal of Informatics*, vol. 7, pp. 167–179, 2019, special issue.

[19] N. Harudin, "Increasing T-method accuracy through application of Robust M-estimator," *International Journal of Engineering & Technology*, vol. 7, no. 3, pp. 44–48, 2018.

[20] W. Z. A. W. Muhamad, F. Ramlie, and K. R. Jamaludin, "Mahalanobis-taguchi system for pattern recognition: a brief review," *Far East Journal of Mathematical Sciences (FJMS)*, vol. 102, no. 12, pp. 3021–3052, 2017.

[21] A. Pal and J. Maiti, "Development of a hybrid methodology for dimensionality reduction in Mahalanobis-Taguchi system using Mahalanobis distance and binary particle swarm optimization," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1286–1293, 2010.

[22] W. H. Woodall, R. Koudelik, K.-L. Tsui, S. B. Kim, Z. G. Stoumbos, and C. P. Carvounis, "Response," *Technometrics*, vol. 45, no. 1, pp. 29-30, 2003.

[23] W. Z. A. W. Muhamad, K. R. Jamaludin, S. A. Saad, Z. R. Yahya, and S. A. Zakaria, "Random binary search algorithm based feature selection in Mahalanobis Taguchi system for breast cancer diagnosis," *AIP Conference Proceedings*, vol. 1974, no. 1, Article ID 020027, 2018.

[24] B. Abraham and A. M. Variyath, "Discussion," *Technometrics*, vol. 45, no. 1, pp. 22–24, 2003.

[25] C. R. Foster, R. Jugulum, and D. D. Frey, "Evaluating an adaptive one-factor-at-a-time search procedure within the mahalanobis-taguchi system," *International Journal of Industrial and Systems Engineering*, vol. 4, no. 6, pp. 600–614, 2009.

[26] D. M. Hawkins, "Discussion," *Technometrics*, vol. 45, no. 1, pp. 25–29, 2003.

[27] K. Tsui, T. Sukchotrat, and V. C. P. Chen, "A comparison study and discussion of the mahalanobis-taguchi system seoung bum kim," *International Journal of Industrial and Systems Engineering*, vol. 4, no. 6, pp. 631–644, 2009.

[28] D. M. Hawkins, "Discussion," *Technometrics*, vol. 45, no. 1, pp. 25–29, 2003.

[29] H. Kawada and Y. Nagata, "An application of a generalized inverse regression estimator to Taguchi's T-Method," *Total Quality Science*, vol. 1, no. 1, pp. 12–21, 2015.

[30] H. Kawada and Y. Nagata, "Studies on the item selection in taguchi's T-method," *Journal of the Japanese Society for Quality Control*, vol. 45, no. 2, pp. 179–193, 2015.

[31] N. Harudin, M. N. Muhtazaruddin, F. Ramlie, and W. M. Wan Zuki Azman, "Artificial bee colony for features selection optimization in increasing T-method accuracy," *International Journal of Engineering & Technology*, vol. 7, no. 4, pp. 885–891, 2018.

[32] N. Harudin, K. R. Jamaludin, F. Ramlie, M. N. Muhtazarudin, C. M. Che Razali, and W. Z. A. Wan Muhamad, "Binary particle swarm optimization for variables selection optimization in Taguchi's T-Method," *Matematika*, vol. 36, no. 1, pp. 69–84, 2020.

[33] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*, Wiley Series in Probability and Statistics, Hoboken, NJ, USA, 2nd edition, 2003.

[34] S. Teshima, Y. Hasegawa, and K. Tatebayashi, *Quality Recognition and Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System*, Momentum Press, New York, NY, USA, 2012.

[35] S. Teshima, Y. Hasegawa, and K. Tatebayashi, "T-Method application procedures and Key points," in *Quality Recognition and Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System*, pp. 87–104, Momentum Press, New York, NY, USA, 1st edition, 2012.

[36] D. Jia, X. Duan, and M. K. Khan, "Binary Artificial Bee Colony optimization using bitwise operation," *Computers & Industrial Engineering*, vol. 76, pp. 360–365, 2014.

[37] M. Lichman, *UCI Machine Learning Repository*, University of California, School of Information and Computer Science, Irvine, CA, USA, 2013, http://archive.ics.uci.edu/ml.

[38] C. M. C. Razali, S. F. M. Hussein, N. Harudin, and S. S. Abdullah, "Estimation of building energy efficiency performance using radial basis function neural network," *International Journal of Engineering & Technology*, vol. 7, no. 4.35, pp. 755–759, 2018.

[39] A. Tsanas and A. Xifara, "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools," *Energy and Buildings*, vol. 49, pp. 560–567, 2012.

[40] N. Harudin, "An overview of taguchi' S T-method as A prediction tool for multivariate analysis," *Open International Journal of Informatics*, vol. 7, no. 1, 2019.