Bulletin of Electrical Engineering and Informatics
Vol. 11, No. 5, October 2022, pp. 2828~2835
ISSN: 2302-9285, DOI: 10.11591/eei.v11i5.4350

Enhanced Taguchi's T-method using angle modulated Bat algorithm for prediction

Zulkifli Marlah Marlan[1], Faizir Ramlie[1], Khairur Rijal Jamaludin[1], Nolia Harudin[2]
[1]Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[2]Department of Mechanical, College of Engineering, Universiti Tenaga Nasional, Kajang, Malaysia

Article Info

Article history:

Received Jun 4, 2022
Revised Jul 7, 2022
Accepted Jul 26, 2022

Keywords:

Angle modulation
Binary Bat algorithm
Feature selection
Mahalanobis-taguchi system
Taguchi's t-method

ABSTRACT

Analysis of multivariate historical information in predicting future state or unknown outcomes is the core function of Taguchi's T-method. Introduced by Dr. Genichi Taguchi under Mahalanobis-Taguchi system, the T-method combines regression principle and robust quality engineering element in formulating a predictive model and employs taguchi's orthogonal array design in optimizing the model through feature or variable selection process. There is a concern regarding the sub-optimality of the T-method prediction accuracy, particularly when the orthogonal array failed to offer a significant number of combinations in search for an optimal subset of features. This is due to the fixed and limited combination offered for evaluation as well as the lack of higher-order interaction of combination. In response to this issue, this paper proposed an angle modulated Bat algorithm to be integrated with the T-method in optimizing the prediction model. A comparison study was conducted using energy efficiency benchmark datasets with the mean absolute error metric used as the performance measure. The results show that the proposed method improved the prediction accuracy by 10.74%, from 6.05 to 5.4, by integrating only four features over the original eight in the prediction model.

Corresponding Author:

Zulkifli Marlah Marlan
Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia (Kuala Lumpur)
54100 Kuala Lumpur, Malaysia
Email: zulkifli49@graduate.utm.my

1. INTRODUCTION

The availability and accessibility of big data, in general, have spurred the betterment of the quality of life through informed decisions, product advancement, enhanced environmental monitoring and safety, effective pollution control, and many more. This, however, could not be accomplished without an effective analysis medium in extracting valuable patterns and information from raw data, such as statistical tools and predictive modeling techniques [1]. Amongst many methods, Taguchi's T-method (T-method) is a relatively new predictive modeling technique introduced by Dr. Genichi Taguchi to predict future state or unknown output based on historical information [2]. As part of the Mahalanobis-Taguchi system, the T-method was developed based on the integration between the regression principle and taguchi's robust quality engineering elements involving a blend of mathematical-statistical analysis in determining the model's parameters and selection of significant input features or variables. Indirectly, this has become the differentiation factor in distinguishing the T-method from other available predictive modeling techniques. Since the introduction, the T-method has been acknowledged as a practical prediction technique, owing to its straightforward step-by-step computational approach, which provides high prediction accuracy while being simple [3]. Some of the prominent advantages of the T-method include its capability in dealing with multicollinearity issues and computationally possible events when the number of sample data is less than the number of input features. As a matter of applicability, the T-method prediction technique has been employed and proven effective in solving many prediction-based problems covering various sectors and industries such manufacturing industry [4], agriculture [5], automotive industry [6], [7], energy sector [8], and healthcare industry [9].

As a multivariate prediction model, the T-method utilizes taguchi's orthogonal array (OA) design to search for an optimal subset of relevant and significant features towards the prediction outcome. The main objective is to filter and extract out insignificant features from the prediction model. It is imperative as the presence of an irrelevant, redundant, and unimportant feature could lead to deterioration of learning accuracy, hinder the full potential of relevant features, increased data dimensionality, complexity, and computational time [10]. However, the utilization of OA leads to the suboptimal prediction accuracy of the T-method as the OA approach failed to determine the optimal feature combination due to the fixed and limited variable combinations employed during analysis [11] and lack of higher-order interactions between features [12]. These suggested that the utilization of OA for the feature selection process in the T-method is ineffective and requires a better feature selection strategy in obtaining optimal prediction accuracy. Thus, this paper proposed the integration of the angle modulated Bat algorithm (AMBA) with the T-method for a better feature selection process, which eventually enhanced the T-method prediction accuracy. An angle modulated refers to the utilization of an angle modulation trigonometric function to map continuous-valued solutions generated by the Bat algorithm into binary form to solve discrete problems such as feature selection. The motivation to employ a swarm-based metaheuristic such as the Bat algorithm is driven by the exploration and exploitation searchability of the metaheuristic algorithm, as reported in [13]. Specifically, in the T-method domain, very less studies were found that integrate the T-method with metaheuristic algorithms. Studies by [14], [15] that employ an artificial bee colony algorithm and a modified artificial bee colony algorithm to optimize the feature selection process show promising improvement in the T-method prediction accuracy over the conventional OA approach in terms of error rate. In the latest development, [16] reported that integrating the binary Bat algorithm based on the nearest integer binarization scheme with the T-method enhanced the prediction accuracy ranging from 7.2% to 55%, dependent on case studies. In a continuous quest for optimality, the ideal approach is through an exhaustive search approach where every possible combination is assessed and evaluated. However, the limitation surfaced as the dimensionality of data increases, resulting in an exponential increment of possible solutions, ultimately increasing computational cost.

2. METHOD

2.1. Conventional T-method prediction model

Conventionally, the formulation of the T-method prediction model consists of five main steps: data collection and preprocessing, estimation of the model's parameter, formulation of the prediction model, estimation of the model quality index, and optimization of the prediction model using OA. This paper focused on enhancing step 5 by replacing the OA approach with the AMBA methodology. The basic prediction model was formulated using steps 1 to 4 as follows:

Step 1: sample data is gathered and tabulated consisting of input and output features. A set of unit space data based on the homogeneity of output is selected, and the average is computed. The unit space data for input features are chosen accordingly. A normalized signal data is then obtained by subtracting the average value of unit space data from raw data for every feature and the output. The normalized signal data consist of all data except unit space data.

Step 2: the model's parameter, namely proportional coefficient, β, and signal-to-noise ratio (SNR), η, is computed based on normalized signal data using (1) and (2), respectively. Fundamentally, the T-method utilizes a least square methodology to derive both [6].

$$Proportional\ coefficient, \beta_j\ = (M_1\, X_{1j}\ +\ M_2\, X_2\ \ +\ \dots\ +\ M_l\, X_l)/r \tag{1}$$

$$Effective\ divider, r\ =\ M_1^2\ +\ M_2^2\ +\ \dots + M_l^2$$

$$SNR, \eta_j \begin{cases} = ((1/r)(S_{\beta j} - V_{ej}))/V_{ej} & ;(when\ S_{\beta j} >\ V_{ej}) \\ =\ 0 & ;(when\ S_{\beta j}\ \leq\ V_{ej}) \end{cases} \tag{2}$$

were, $Error\ variance, V_{ej}\ \ =\ \ S_{ej}/(l-1)$

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

$Total\ variation, S_{Tj}\ =\ X_{11}^2\ +\ X_{21}^2\ +\ \dots + X_{lj}^2$

$Variation\ of\ proportional\ term, S_{\beta j}\ = (M_1\, X_{1j}\ +\ M_2\, X_{21}\ \ +\ \dots\ +\ M_l\, X_{lj})^2/r$

Step 3: the T-method prediction model is formulated using the integrated estimate output value as shown in (3), where X_i is the value of normalized signal data for i = 1, 2, …, l of signal data, and k is the number of input features. A weightage average methodology based on SNR value is employed in deriving the integrated estimate out value.

$$Integrated\ Estimate\ Output\ Value,$$
$$\widehat{M}_i = ((\eta_1 \times X_{i1}/\beta_1) + (\eta_2 \times X_{i2}/\beta_2) + \cdots + (\eta_k \times X_{ik}/\beta_k))/(\eta_1 + \eta_2 + \cdots + \eta_k) \quad (3)$$

Step 4: the quality index representing the current model performance is computed using (4), where M is the actual normalized output of signal data, $\widehat{M}$ is the predicted output obtained from (3). A better subset of input feature combinations will result in a better or higher integrated estimate SNR (db) value.

$$Integrated\ estimate\ SNR\ (db), \eta_{est} = 10\ log\ ((1/r)(S_{\beta j} - V_{ej})/V_{ej}) \quad (4)$$

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

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

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

$Variation\ of\ proportional\ term, S_\beta = L^2/r$

$Error\ variation, S_e = S_T - S_\beta$

$Error\ variance, V_e = S_e/(l-1)$

Step 5: the procedure for determining the subset of significant features using taguchi's OA design consists of three steps, namely the selection of suitable OA design, estimation of integrated estimated SNR (db), and generation of factorial effect chart. The factorial effect chart reveals the relative importance of features based on the deterioration of the average integrated estimate SNR (db) of features when it is not used in combination. The detail of the steps is elaborated in [16].

2.3. Angle modulated bat algorithm

An AMBA is a binary variant of a Bat algorithm that is based on the angle modulation discretization strategy for discrete optimization problems such as feature selection. Rather than assessing a fixed number of feature combinations as in OA methodology, AMBA generates and offers countless feature subsets for evaluation, although not exhaustive, until an optimal (or sub-optimal) solution is achieved. AMBA fundamentally is a Bat algorithm that was introduced by [17] based on the echolocation ability of microbats, except that in AMBA, bats move in binary space, while in the Bat algorithm, bats are moving in continuous solution space. Idealized by the bat's superior characteristics, the Bat algorithm was developed following three approximation rules as follows:

a. Echolocation is used by all bats to recognize distance, and they also have an innate ability to tell the difference between food/prey and background barriers in a mysterious way;
b. Bats search for prey by flying randomly at a velocity v and a position x with a fixed frequency f_{min}, varying wavelength, and loudness. They are capable of autonomously adjusting the frequency and pace of their transmitted pulses in response to their target's proximity;
c. While the loudness may change in a variety of ways, it is assumed that it ranges between a large (positive) A_0 and a minimum constant value, A_{min}.

In the Bat algorithm, the artificial bat, i fly from one solution to another by updating frequency, f_i, velocity, v_i, and position, x_i using (5), (6), and (7), where $\beta \in [0,1]$ is a uniform generated random number, t is search iteration and x^* is the current global best solution. Next, a random walk is conducted on the chosen optimum solution to explore further the solution for optimality using (8), where $\varepsilon \in [-1,1]$ is a uniform generated random number and $\overline{A}^t$ is the average loudness of all bats at iteration, t.

$$f_i = f_{min} + (f_{max} - f_{min})\beta \quad (5)$$

$$v_i^t = v_i^{t-1} + (x_i^t - x_*)f_i \quad (6)$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (7)$$

$$x_{new} = x_{old} + \varepsilon\bar{A}^t \quad (8)$$

In the Bat algorithm, two important parameters are updated as search iteration grows, which are loudness, A_i and pulse rate, r_i using (9) and (10), respectively, where α and γ are constant parameters. The main objective is to balance the exploration on the global scale and exploitation of solutions in the targeted region.

$$A_i^{t+1} = \alpha A_i^t \tag{9}$$

$$r_i^{t+1} = r_i^0 \, [1 - exp(-\gamma t)] \tag{10}$$

The binary variant of the Bat algorithm was first introduced by [18] by utilizing a sigmoid-based transfer function to map a continuous-valued velocity of the bat into a binary position in the form of a bit string. Subsequently, [19] introduced a v-shaped transfer function instead of the sigmoid-based function to obtain the binary position of the bat. The similarity between the two approaches in binary discretization is the dependency on the velocity of bats in generating the binary position. Later, [20] conducted a study to determine the efficiency of five binary mapping functions using the BAT algorithm: nearest integer, normalization, angle modulation, sigmoid function, and great value priority. Some mapping mechanisms, such as nearest integer and great value priority, convert the continuous-valued position into a binary position instead of continuous-valued velocity as in [18], [19]. The study's outcome shows that the effect of such mapping functions on an algorithm's efficiency depends on the size of the problem and its complexity.

Specifically, for the angle modulation mapping function, a trigonometric-based generator function that is used for signal processing in the telecommunications domain is employed to map continuous-valued solutions into binary form. The generator function is as (11), where i=1, 2, …, N, j=1, 2, …, D, N is the population size, D is the problem dimension, x_{ij} is a single element from a potential solution vector having equal interval and [a, b, c, d] are coefficients for the generator function obtained from four dimensions of the continuous-valued solution vector. Conceptually, rather than optimizing a d-dimensional binary string solution, the search space is condensed into four dimensions, where each dimension represents the coefficient in the generator function.

$$AMBA, g(x) = sin\left(2\pi(x_{ij} - a) * b * cos(2\pi(x_{ij} - a) * c)\right) + d \tag{11}$$

Realizing the deficiency of the existing generator function in the angle modulated binary Bat algorithm, [21] introduced an amplitude angle modulated version that modified the generator function by incorporating an additional variable coefficient, e, as (12). The new coefficient, e, aims to control the amplitude of the sin wave, likewise of coefficient b towards the amplitude of the cos wave. In the recent development, a new version called a phase AMBA was introduced by [22] by adding a phase adjustment to the cosine wave and an amplitude adjustment to the sine wave. While the latter is similar to the amplitude angle modulated version, the new modification is claimed to play an important role in the global search ability of the optimization algorithm. As (13), the enhanced generator function involved an additional coefficient, e, and g, as opposed to the amplitude version with only coefficient, e. For each version of angle modulated binary Bat algorithm, the obtained generated continuous-value will be transformed into binary form through the binary operator as (14), where a generator function value more than zero is transformed into 1, and 0 otherwise.

$$A\text{-}AMBA, g(x) = e * sin\left(2\pi(x_{ij} - a) * b * cos(2\pi(x_{ij} - a) * c)\right) + d \tag{12}$$

$$P\text{-}AMBA, g(x) = e * sin(2\pi(x_{ij} - a) * b * cos(2\pi(x_{ij} - a) * c) + g) + d \tag{13}$$

$$x_{ij} = \begin{cases} 1, & if\ g(x_{ij}) \geq 0 \\ 0, & otherwise \end{cases} \tag{14}$$

2.4. T-method with angle modulated Bat algorithm algorithm

The integration of AMBA with the T-method follows a similar approach as Taguchi's OA, where the AMBA focused on optimizing the feature selection process after the T-method prediction model was successfully formulated. In addition, the objective function of the AMBA with the T-method is also similar: to maximize the integrated estimate SNR (db) value that represents the quality index of model performance. A higher integrated estimate SNR (db) value is better as it represents a low model variability or noise [23]. Through the integration, the AMBA algorithm is expected to evaluate as many feature combinations (solution) and as varies as possible in search of the optimal subset of features. The best fitness value with the respective solution (feature's combination) is stored as the current global best solution, exhibiting greedy selection criteria. In generating a new solution, the current global solution is considered in the computation of the bat's velocity to attract the search direction towards optimality. The search process continues until a stopping criterion is met. The pseudocode for AMBA can be found in [21], and it can be used for A-AMBA and P-AMBA by increasing the random initialization dimension to five and six, respectively.

3. RESULTS AND DISCUSSION

3.1. Experimental design

In this study, the performance of taguchi's T-method with AMBA (and its variants: A-AMBA and P-AMBA) against the conventional T-method with OA is computed and compared using an energy efficiency benchmark dataset obtained from the university of california at irvine (UCI) machine learning repository [24]. In particular, only a single output response, cooling load, is considered. The data was primarily collected to predict the cooling load of residential buildings based on eight contributing factors as independent variables [25]. A 70%-30% hold-out cross-validation is randomly applied to the 768-sample data, resulting in 538 samples used to train the model and 230 samples data for model validation purposes. For this analysis, the subset of unit space data consists of 5 sample data with homogeneous characteristics located in the output feature's densely populated region. Thus, it resulted in the remaining 533 sample data for signal data, which was subsequently normalized by subtracting the average value of unit space data from the signal data. The normalized signal data is then used to determine the T-method model parameters: proportional coefficient, β, and SNR, η, as shown in Table 1. For any feature having computed negative SNR value, the negative value is converted to zero satisfying the SNR (2). The T-method prediction model in the form of an integrated estimate output value consisting of all variables is obtained by substituting the proportional coefficient, β, and SNR, η value into (3).

Table 1. Computed model parameters for cooling load train data set

Parameter	X1	X2	X3	X4	X5	X6	X7	X8
β	0.007	-6.342	1.772	-4.057	0.166	0.001	0.003	0.012
η	0.005	0.005	0.002	0.011	0.007	-1.897 ↓	0.0001	$2.404E^{-05}$
Corrected η	0.005	0.005	0.002	0.011	0.007	0	0.0001	$2.404E^{-05}$

In determining the parameters setting for AMBA, A-AMBA, and P-AMBA, a taguchi method technique is employed as it provides a simple, efficient, and systematic approach in determining the optimum setting [26]–[29]. Seven parameters with three levels each were selected, where the level's value was obtained from a compilation of past research practices in executing the binary Bat algorithm. An L_{27} orthogonal array was employed as an experimental design based on the number of parameters and levels. Instead of a total of 2187 experimental runs, the taguchi method reduced the number of experiments runs to only 27 with a balanced design to ensure that all parameters are considered equally [30], [31]. The objective function of the parameter experiment is to minimize the mean absolute error (MAE) on the training data set, thus signifying that a smaller-the-better signal-to-noise ratio is utilized in analyzing the response. Each experimental run was repeated three times to minimize the influence of noise factors and the MAE percentage as the response is used in determining the significance of the parameter's level. Table 2 shows the finalized optimal parameters setting for AMBA, A-AMBA, and P-AMBA.

Table 2. AMBA, A-AMBA and P-AMBA parameters setting

Parameter	AMBA	A-AMBA	P-AMBA
Population size	50	50	50
Min frequency, *fmin*	0.8	0.8	0
Max frequency, *fmax*	2	2	0.15
Pulse rate, *r*	0.5	0.5	0.9
Loudness, *A*	0.5	0.75	0.75
Alpha, α	0.9	0.1	0.7
Gamma, γ	0.1	0.1	0.9

Upon identifying the optimal parameter setting, the T-method with AMBA, A-AMBA, and P-AMBA algorithms was executed using the respective optimal parameter setting to obtain the optimal feature subset. In this study, a total of 20 independent runs were executed for each algorithm, with 500 internal iterations for each run using normalized signal data. The selection criteria of features to be included in the subset of significant features is to appear 50% or more over the total number of runs. In quantifying the performance of the feature selection process against the original list of features, a reduction rate is computed, where the metric represents the percent decrease in the number of features used in the T-method prediction model. Next, the overall performance of the proposed T-method with AMBA, A-AMBA, and P-AMBA with the optimal parameters setting using the optimal subset of features is evaluated using the validation data set. The MAE metric is used as the performance measure in evaluating the conventional and proposed T-method prediction accuracy. In addition, a comparison study is conducted by comparing the performance of the proposed T-method with AMBA, A-AMBA and P-AMBA against the conventional T-method without feature selection and T-method with OA approach. Throughout this study, all algorithms were constructed and executed using MATLAB R2020a programming application software on a laptop-type computer equipped with an intel core i5-8250U central processor unit, 4 gigabytes of random-access memory, and 1 terabyte of storage capacity.

3.2. Experimental result

Table 3 shows the compilation of results which consists of the number of features, the respective feature combination, and the reduction rate for the T-method without feature selection, T-method with OA approach, T-method with AMBA, T-method with A-AMBA, and the T-method with P-AMBA obtained using train data set. Originally, the T-method is formulated using all eight input features. However, as the feature selection process is applied, the T-method through OA approach, T-method with AMBA, T-method with A-AMBA and T-method with P-AMBA recorded a reduction in the number of features. The T-method with OA recorded a 62.5% reduction rate over the original number of features, with features X3, X4 and X7 selected as significant features. The T-method with AMBA and T-method with P-AMBA recorded a 37.5% reduction rate, with features X1, X3, X6, X7, and X8 selected as significant features. As for T-method with A-AMBA, a 50% reduction rate was recorded, with features X1, X3, X7, and X8 selected as significant features. In general, integrating the feature selection process either through the OA approach or the AMBA algorithm and its variant reduces the dimensionality of data, resulting in a smaller and less complex T-method prediction model. However, the number of significant features and their respective combinations show dissimilarity, except for the T-method with AMBA and P-AMBA. The performance on the validation data set is sought to determine whether these differences can improve or maintain the T-method's prediction accuracy on a new set of data.

Table 3. Optimal features combination

Dataset	Item	T-method	T-method + AO	T-method + AMBA	T-method + A-AMBA	T-method + P-AMBA
Cooling load	No. of feature	8	3	5	4	5
	Opt. Combination	all	3, 4, 7	1, 3, 6, 7, 8	1, 3, 7, 8	1, 3, 6, 7, 8
	Reduction rate, R_r	-	62.5%	37.5%	50.0%	37.5%

Table 4 shows the T-method prediction accuracy based on the MAE metric for the T-method without feature selection process, T-method with OA, T-method with AMBA, T-method with A-AMBA and T-method with P-AMBA computed using validation data set. The result clearly shows that T-method with AMBA, A-AMBA, and P-AMBA recorded the best prediction accuracy at 5.40 compared to T-method with AO and T-method without feature selection at 6.05 and 8.08, respectively. The value in parenthesis is the percent enhancement of MAE of the T-method against the t-method without feature selection. Thus, these suggest that the feature selection process help to enhance the T-method prediction accuracy further when only relevant and significant features are incorporated into the prediction model. As for why the T-method with AMBA, T-method with A-AMBA, and T-method with P-AMBA recorded the same MAE result, the optimal subset of features consists of similar features, which are X1, X3, X7, and X8, except feature X6. The presence of feature X6 in the optimal subset of features for the t-method with AMBA and P-AMBA is found to not contribute to the objective function due to feature X6 having zero SNR value. Thus, it resulted in the same MAE as the T-method with A-AMBA.

Table 4. Prediction accuracy (MAE) on the validation data set

Dataset	T-method	T-method + AO	T-method + AMBA	T-method + A-AMBA	T-method + P-AMBA
Cooling load	8.08	6.05 (25.1%)	5.40 (33.2%)	5.40 (33.2%)	5.40 33.2%)

4. CONCLUSION

This study proposed the integration of an AMBA with the T-method to enhance the T-method prediction accuracy by optimizing the feature selection process. Three variances of the AMBA were considered, and the performance against the T-method without the feature selection process and T-method with OA were compared. The result strongly suggests that the feature selection process is imperative and necessary to improve the T-method prediction accuracy regardless of the approaches. By considering only significant features in the prediction model, the prediction accuracy of the T-method with AMBA improved by 33.2% over the T-method without a feature selection process or 25.1% over the T-method with OA, indicating a better and more reliable prediction model was achieved. In addition to enhanced accuracy, the model with fewer features resulted in a less complex and simple model. Based on observation, the betterment of accuracy was contributed by the recursive and adaptive characteristic of the AMBA in search of the optimal solution. To conclude, the integration is feasible, and the experimental results strongly suggest the utilization of the AMBA with the T-method for better prediction accuracy.

ACKNOWLEDGEMENTS

This work was supported under the fundamental research grant scheme (FRGS) awarded by the Ministry of Education, Malaysia (Ref: FRGS/1/2019/TK03/UTM/02/10) and Universiti Teknologi Malaysia.

REFERENCES

[1] S. Gupta, J. Mateu, A. Degbelo, and E. Pebesma, "Quality of life, big data and the power of statistics," *Statistics & Probability Letters*, vol. 136, pp. 101–104, May 2018, doi: 10.1016/j.spl.2018.02.030.

[2] S. Teshima, Y. Hasegawa, and K. Tatebayashi, *Quality recognition and prediction: smarter pattern technology with the mahalanobis-taguchi system*. Momentum Press, 2012.

[3] N. Harudin *et al.*, "An overview of Taguchi's T-method as a prediction tool for multivariate analysis," *Open Int. J. Informatics*, vol. 7, no. 1, pp. 158–166, 2019.

[4] A. Suzuki and K. Sugimoto, "Robust estimation of heat capacity in cooking devices," in *2010 IEEE International Conference on Control Applications*, Sep. 2010, pp. 1374–1378. doi: 10.1109/CCA.2010.5611231.

[5] P. Dasneogi and E. A. Cudney, "Comparing the predictive Ability of T-Method and linear regression method," in *in Proceedings of the 2009 Industrial Engineering Research Conference Comparing*, 2009, pp. 2176–2182.

[6] K. Nishino and A. Suzuki, "Taguchi's T-method using median-median line for small sample with outliers," *Electronics and Communications in Japan*, vol. 102, no. 1, pp. 49–56, Jan. 2019, doi: 10.1002/ecj.12134.

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

[8] 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, Oct. 2017, doi: 10.1002/eej.22988.

[9] A. Suzuki, "Inverse-model-based cuffless blood pressure estimation using a single photoplethysmography sensor," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 229, no. 7, pp. 499–505, Jul. 2015, doi: 10.1177/0954411915587957.

[10] J. Pirgazi, M. Alimoradi, T. Esmaeili Abharian, and M. H. Olyaee, "An efficient hybrid filter-wrapper metaheuristic-based gene selection method for high dimensional datasets," *Scientific Reports*, vol. 9, no. 1, p. 18580, Dec. 2019, doi: 10.1038/s41598-019-54987-1.

[11] S. B. Kim, K. L. Tsui, T. Sukchotrat, and V. C. P. Chen, "A comparison study and discussion of the Mahalanobis-Taguchi System," *International Journal of Industrial and Systems Engineering*, vol. 4, no. 6, p. 631, 2009, doi: 10.1504/IJISE.2009.026768.

[12] W. H. Woodall, R. Koudelik, K.-L. Tsui, S. B. Kim, Z. G. Stoumbos, and C. P. Carvounis, "A review and analysis of the Mahalanobis—Taguchi System," *Technometrics*, vol. 45, no. 1, pp. 1–15, Feb. 2003, doi: 10.1198/004017002188618626.

[13] K. Hussain, M. N. M. Salleh, S. Cheng, and Y. Shi, "On the exploration and exploitation in popular swarm-based metaheuristic algorithms," *Neural Computing and Applications*, vol. 31, no. 11, pp. 7665–7683, Nov. 2019, doi: 10.1007/s00521-018-3592-0.

[14] N. Harudin, J. K R, M. Nabil Muhtazaruddin, R. F, W. Zuki Azman Wan Muhammad, and N. Jaafar, "Artificial bee colony for features selection optimization in increasing T-Method accuracy," *International Journal of Engineering & Technology*, vol. 7, no. 4.35, p. 885, Nov. 2018, doi: 10.14419/ijet.v7i4.35.26276.

[15] N. Harudin *et al.*, "Binary bitwise artificial bee colony as feature delection optimization approach within Taguchi's T-Method," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–10, May 2021, doi: 10.1155/2021/5592132.

[16] Z. M. Marlan, K. R. Jamaludin, F. Ramlie, and N. Harudin, "Taguchi's T-method with nearest integer-based binary bat algorithm for prediction," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 4, pp. 2215–2224, Aug. 2022, doi: 10.11591/eei.v11i4.3859.

[17] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," 2010, pp. 65–74. doi: 10.1007/978-3-642-12538-6_6.

[18] R. Y. M. Nakamura, L. A. M. Pereira, D. Rodrigues, K. A. P. Costa, J. P. Papa, and X.-S. Yang, "Binary bat algorithm for feature selection," in *Swarm Intelligence and Bio-Inspired Computation*, Elsevier, 2013, pp. 225–237. doi: 10.1016/B978-0-12-405163-8.00009-0.

[19] S. Mirjalili, S. M. Mirjalili, and X.-S. Yang, "Binary bat algorithm," *Neural Computing and Applications*, vol. 25, no. 3–4, pp. 663–681, Sep. 2014, doi: 10.1007/s00521-013-1525-5.

[20] Z. A. E. M. Dahi, C. Mezioud, and A. Draa, "Binary bat algorithm: on the efficiency of mapping functions when handling binary problems using continuous-variable-based metaheuristics," 2015, pp. 3–14. doi: 10.1007/978-3-319-19578-0_1.

[21] X. Huang, P. Li, and Y. Pu, "Amplitude angle modulated bat algorithm with application to zero-one knapsack problem," *IEEE Access*, vol. 7, pp. 27957–27969, 2019, doi: 10.1109/ACCESS.2019.2901988.

[22] J. Dong, Z. Wang, and J. Mo, "A phase angle-modulated bat algorithm with application to antenna topology optimization," *Applied Sciences*, vol. 11, no. 5, p. 2243, Mar. 2021, doi: 10.3390/app11052243.

[23] G. Taguchi, S. Chowdhury, and Y. Wu, *Taguchi's quality engineering handbook*. 2004.

[24] M. Lichman, "UCI machine learning repository," *Irvine, CA: University of California, School of Information and Computer Science*, 2013. https://archive.ics.uci.edu/ml/datasets

[25] 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, Jun. 2012, doi: 10.1016/j.enbuild.2012.03.003.

[26] M. Nalbant, H. Gökkaya, and G. Sur, "Application of Taguchi method in the optimization of cutting parameters for surface roughness in turning," *Materials & Design*, vol. 28, no. 4, pp. 1379–1385, Jan. 2007, doi: 10.1016/j.matdes.2006.01.008.

[27] J. A. Ghani, H. Jamaluddin, M. Nizam Ab., and B. Md. Deros, “Philosophy of Taguchi approach and method in design of experiment,” *Asian Journal of Scientific Research*, vol. 6, no. 1, pp. 27–37, Dec. 2012, doi: 10.3923/ajsr.2013.27.37.

[28] G. Candan and H. R. Yazgan, “Genetic algorithm parameter optimisation using Taguchi method for a flexible manufacturing system scheduling problem,” *International Journal of Production Research*, vol. 53, no. 3, pp. 897–915, Feb. 2015, doi: 10.1080/00207543.2014.939244.

[29] A. A.Malik, S. M.Maharum, M. Z. W. Mohd. Tohid, M. K.Rahmat, and I. H.Hamzah, “Statistical analysis using taguchi method for wind turbine in ducting system,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 15, no. 1, p. 27, Jul. 2019, doi: 10.11591/ijeecs.v15.i1.pp27-35.

[30] N. C. Fei, N. M. Mehat, and S. Kamaruddin, “Practical applications of Taguchi method for optimization of processing parameters for plastic injection moulding: a retrospective review,” *ISRN Industrial Engineering*, vol. 2013, pp. 1–11, Jun. 2013, doi: 10.1155/2013/462174.

[31] R. K. Roy, *A primer on the Taguchi method*. Society of Manufacturing Engineers, 2010.

BIOGRAPHIES OF AUTHORS

Zulkifli Marlah Marlan received a Bachelor’s degree in Mechanical (Industrial) Engineering from Universiti Teknologi Malaysia in 2008 and a Master’s degree in Engineering Management from Universiti Putra Malaysia in 2014. He is a Graduate Member of the Board of Engineers, Malaysia. Currently, he is conducting research on Taguchi’s T-method for prediction as a Ph.D. postgraduate student at Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Malaysia. His research areas include the Mahalanobis-Taguchi system, metaheuristic algorithm, and robust statistics. He can be contacted at email: zulkifli49@graduate.utm.my.

Faizir Ramlie is a senior lecturer at Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Malaysia since 2017. He received a Bachelor’s degree in Mechanical Engineering from the University of Queensland, Australia, in 1997, a Master’s degree in Mechanical Engineering, specializing in Advanced Manufacturing Technology, and a Ph.D. from Universiti Teknologi Malaysia in 2016. Experienced engineer with over 15 years of experience in various multinational companies in Malaysia, involving product development and manufacturing. His research interests are primarily in the area of pattern recognition using Taguchi Method and artificial intelligence, the Mahalanobis-Taguchi system, robust engineering, and manufacturing technology. He can be contacted at email: faizir.kl@utm.my.

Khairur Rijal Jamaludin is an Associate Professor at Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Malaysia. He received a Bachelor’s degree in Mechanical Engineering from Universiti Teknologi Malaysia in 1996, a Master of Science degree in Advanced Mechanical Engineering from the University of Warwick in 1998, and a Ph.D. from Universiti Kebangsaan Malaysia in 2009. He is a Profesional Technologist (Ts.) of Malaysia Technologies Board and a member of various professional membership such as Graduate Member Board of Engineers, Malaysia, Graduate Member Institution of Engineers Malaysia, Member of American Powder Metal Institute, and Quality Engineering Society, Japan. His research interests are primarily in the area of robust engineering (Taguchi Method and Mahalanobis-Taguchi system), remanufacturing, powder metallurgy and particulate materials, quality engineering and quality management, Lean / Toyota Production systems, and end-of-life vehicles/ vehicles recycling, where he is the author/co-author of over 50 research publications. He can be contacted at email: khairur.kl@utm.my.

Nolia Harudin is a senior lecturer at Universiti Tenaga Nasional, Malaysia since 2013. She received a Bachelor’s degree in Mechanical (Industrial) Engineering from Universiti Teknologi Malaysia in 2008, a Master’s degree in Mechanical Engineering from Universiti Teknologi Malaysia in 2012, and a Ph.D. from Universiti Teknologi Malaysia in 2020. She is a Graduate Member of the Board of Engineers, Malaysia. Her research interests are Taguchi Method, Mahalanobis-Taguchi system, pattern recognition, energy efficiency, robust engineering, and quality control. She can be contacted at email: nolia@uniten.edu.my.