International Journal of Engineering & Technology, 7 (4.35) (2018) 885-891

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Artificial Bee Colony for Features Selection Optimization in Increasing T-Method Accuracy

N Harudin[1] , Jamaludin, K R[2] , M Nabil Muhtazaruddin[3] , Ramlie F[4] , Wan Zuki Azman Wan Muhamad[5], NN Jaafar[6]

[1,2,3,4]*Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Jalan Sultan Yahaya Petra, 54100 Kuala Lumpur*
[1]*Department of Mechanical Engineering, Universiti Tenaga Nasional, 43000 Kajang Selangor, Malaysia*
**Corresponding author E-mail: nolia.harudin@gmail.com*

Abstract

The study of prediction has drawn great interest in a wide range of field. T-Method which was developed specifically for prediction of the multidimensional case using historical data to develop its baseline model proved that making a prediction is possible even with limited sample size. The element of the signal to noise ratio (SNR) adopted into the T-Method strengthens its robustness. Orthogonal array (OA) in T-Method was used as features selection optimization in improving the analysis speed, cost and computer burden during the analysis. However, the limitation of OA in dealing with higher dimensionality and complex combination factors restraint the optimization accuracy. Artificial Bee Colony (ABC) was adopted in this study to overcome this limitation. The result of this study shows that T-method +ABC provide the best error% accuracy with only 2.45% and 2.53% (3 optimized features out of 15) compared to T-Method +OA which 2.81% and 2.67% and T-Method +Spearman Correlation as 3.16% and 3.06%. The power consumption prediction case study is a good example for cases that deal with high correlation coefficient (R^2) baseline model (>0.8). If the R^2 is lower than 0.8, further enhancement needs to be done to ensure a low risk of high error% prediction.

Keywords: *Artificial Bee colony; Feature selection; Orthogonal Array; Prediction; T-Method*

1. Introduction

The future in industrial systems is always to some extent uncertain and unpredictable. Disruption tends to be a daily challenge to ensure a robust design, the robust operating system as well as liable performance and responsive risk control are all well responded and predicted. Not all the uncertainty is a result of unpredictable consequences. Perhaps the uncertainty is a good sign of better prediction. Much of our current mood of uncertainty has specific causes. How well we leverage it with the advanced cognitive algorithms, predictive modeling and statistical analysis towards better industrial system were aggressively discussed among analyst. Various prediction tools have been used by several practitioners in various areas for the past few decades and it is progressively enhanced up to recent. Even though the argument on traditional statistical tools were aggressively debated and enhanced up to recent, machine learning finds out to be the most preferable approaches in prediction analysis over the decades.

Up to recent, the enhancement on machine learning algorithm is widely practiced and in fact merging the statistical methods or any other prediction methods with machine learning as hybrid tools were mostly the practiced nowadays. J. Drgona et al., (2018) combined the approach of multivariate regression and dimensionality reduction algorithms together with deep time-delay neural networks (TDNN) and regression trees (RT) to derive the dependency of multiple real-valued control inputs on parameters for building control [1]. J. Reynolds et al., (2018) which also done analyses on building energy optimization are combining Artificial Neural Network (ANN) and Genetic algorithm (GA) in their prediction model [2]. S. Mohsen et. al., (2014) developing a Generalized Feed Forward (GFF) network model for predicting annual flood depth of Johor River in Peninsular Malaysia [3]. Besides modern and classical statistical methods, other prediction methods are also widely applied such grey prediction model used by S. Ding et al., (2018) to predict China's electricity consumption which proved to provide superiority as a benchmark model [4]. N.H Abd Rahman et al., (2016) used time series methods namely seasonal autoregressive integrated moving average (SARIMA) and fuzzy time series (FTS) as an effective decision making of the modern statistical tool in forecasting air pollution index and predicting air quality in Johor [5].

However, all methods available recently were highly depended on data trend and normality assumption. Methods that can provide most accurate prediction is the best but not to forget methods that can easily understand and less time consuming for the calculation with an acceptable amount of accuracy will definitely bring benefit to the industry that relies on fast response and prediction results. On that purposes, there comes the idea of this paper which intentionally to highlight an interesting prediction method called T-Method which rarely shared by analyst yet very helpful in making fast and acceptable range of prediction with some enhancement of feature selection element been absorbed into it. T-Method is one of the numerous tools within Mahalanobis Taguchi System (MTS) theory which weightage by signal to noise ratio (SNR) in making a prediction by relying on historical data. The fundamental T-Method theory consist the concept of unit space and zero-proportional linear regression while orthogonal array as feature selection optimization tool. The benefits of T-Method are its

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

straightforwardness to be understood as well as its ability to do a prediction even with a very limited sample size. When developing a product or determining the optimum manufacturing condition of a new product, it is recommended to acquire knowledge through systematically organized experimentation with the help of orthogonal array. But if one has had experience in manufacturing condition modification or gradual improvement of the manufacturing process, then T-Method is better suited to acquiring useful information [6]. There are various cases study conducted which applied T-Method as the tool for dealing with certain prediction analysis such as [7]–[11].

An additional key element embedded into T-Method rather than making a prediction with all features consideration was the parameter optimization or also known as feature selection which used for making a prediction based on important features only. The orthogonal array (OA) was the current practiced. The uniqueness of this OA element in T-Method was it robustness in assessing variability among noise factors through the signal to noise ratio (SNR) analysis. SNR helps in identifying the most insensitive features to variation as well as cost-effective gain. However, the operational and utilization aspect of OA as an optimization medium were judgmentally argued and claimed as inadequate since it produces a sub-optimal solution [12]–[15]. The main intention of this study is to share an adoption of swarm intelligent method called Artificial Bee Colony (ABC) which used as same intention as OA in optimizing feature selection while maintaining the SNR element as the key indicator for optimum results identification. The optimized parameters were then will be used to predict the power consumption of district thermal cooling storage system.

2. Alternative Feature Selection Method Replacing Orthogonal Array Within T-Method

Feature selection is a method optimally used in selecting an optimum subset of good features with least error. In this work, it is needed to fulfill the least error for prediction accuracy. Techniques of feature selection introduced up to recent help in reducing overfitting during the learning loop. In the current method, orthogonal array (OA) with a signal to noise ratio (SNR) is the one been used for this purpose. However, W.H.Woodall et al., (2003) critique that the run for which all factors at their low level in the MT method is not a valid run since at least one variable must be used in the analysis. Thus, Orthogonal array (OA) containing this run could not be used and claimed that solution governs by this might lead to inadequate and inaccurate analysis [12]. Issues on OA been highlighted as well by several experts and practitioners especially on the fact that OA design tends to assume that there is no interaction between variables which might lead to an inconsistency in the identification of the important variable, especially high dimensionality issues [13], [14], [16]. The enhancement on OA are widely discussed and been emphasized continuously up to recent but most of its application is only for adaption into MT Method which been called Mahalanobis Taguchi system (MTS) by most of the practitioner. The enhancement since the past 15 years covered a broad area of feature selection optimization tools [17]–[25]. In the context of T-method, no published work on the enhancement of OA within T-method being reported yet up to this study been written.

Recently, various methods introduced including Ant Colony Optimization (ACO), Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Simulated Annealing (SA) and many more which used heuristic and random search methods in searching for the optimum subset. This paper focused on the application of Artificial Bee Colony (ABC) algorithm replacing OA for fulfilling the intention in making a prediction in T-Method by relying on optimum features only rather than full features consideration. A. L. Bolaji et al., (2013) Summarized that besides having fewer control parameters, ABC also easy to be implemented, robust and having high flexibility as well as efficiently solved engineering issues with high dimensionality [26]. There are quite numbers of practitioners that applied ABC as single and hybrid approaches in finding their optimum feature for classification case [27]–[29]. E. Zorarpaci et al., (2015) Modified the ABC algorithm for feature selection purposes into a binary form that affects the way of producing the new neighborhood of food source operator while the fitness function was done separately using Weka J48 classifier. The team proved that ABC is performing better than other algorithm and can be applied efficiently in solving hard combination problems with high dimensionality [30].

3. Methodology

This section will explain the theory behind T-method as well as the replacement of Orthogonal array (OA) with Artificial Bee Colony (ABC) algorithm as feature selection optimization tools prior making a prediction on the expected output which mainly relies on optimized feature rather than the whole feature for future prediction.

3.1. T-Method

Combining the concept of Mahalanobis Distance theory with Signal to noise ratio (SNR) which emphasize on three core elements (sensitivity, linearity, and variability) bring the formulation of T-Method a practical approach to be applied in a dynamic environment particularly. Following the Zero-proportional concept with linear regression line pass through zero-point (origin) helps in simplifying the regression formulation. The core element in T-Method is the establishment of unit space and signal space for the normalization analysis. Selection of unit space is one of the most crucial decision-making before making further analysis. Relying on two main rules which are homogenous as possible and taken from the high-density population requires some level of expertise to properly defined the target region prior to selecting the unit space. Taguchi clearly mentioned in his main reference book by following these rules, the selection of unit space will always be in the middle position between low and high data of the selected region [6]. Figure 1 illustrates the unit space concept in a simple manner. Prediction accuracy, as well as a strong justification, can be made once the rule of unit space is fulfilled as well as putting some effort into defining the main population data trend.

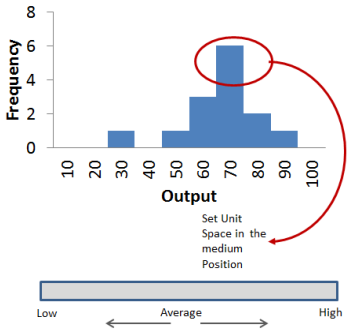


Fig. 1: The concept of unit space selection for T-Method

To fulfill the prediction model which is the Integrated estimate output value ($\hat{M}$) as in equation 1, the proportional coefficient (β) and SN ratio (η) need to be computed item by item basis with the use of normalized data (X_{ij}) calculated by equation 2. The original data was taken from a historical track record on the specific analysis, which then the determination of unit space is done, and the selected data excluded and average out a single unit space value of each variable and output. The remaining unchosen data were treat as signal data. Here is where equation 2 been used as explained earlier. Equations 3 until equation 9 are the formulation of

the proportional coefficient (β) and SN ratio (η). If the value of the SN ratio η calculated to be a negative value, it needs to be assumed as zero. It is clearly seen that higher SN ratio of an item, will contribute to a greater degree of contribution on overall model estimation.

Formulas to calculate the SN ratio η (db) for optimization on feature selection are shown below:

$$M_i = \frac{\eta_1 \times \left(\frac{X_{i1}}{\beta_1}\right) + \eta_2 \times \left(\frac{X_{i2}}{\beta_2}\right) + \cdots . + \eta_j \times \left(\frac{X_{ij}}{\beta_j}\right)}{\eta_1 + \eta_2 + \cdots + \eta_j} \quad (1)$$

$$normalized\ data\ (X) = \text{signal data} - \text{average of unit space} \quad (2)$$

$$\text{Effective Divider, r} = M_1^{\,2} + M_2^{\,2} + \cdots + M_i^{\,2} \quad (3)$$

$$\text{Total Variation, } S_T = X_1^{\,2} + X_2^{\,2} + \cdots + X_i^{\,2} \quad (4)$$

$$\text{Variation of proportional term, } S_\beta = \frac{(M_1X_{11} + M_2X_{21} + \cdots + M_lX_{l1})^2}{r} \quad (5)$$

$$\text{Error variation, } S_e = S_T - S_\beta \quad (6)$$

$$\text{Error Variance, } V_e = \frac{S_e}{l-1} \quad (7)$$

$$\text{Duplicate SN ratio, } \eta = \left(\frac{(S_\beta - V_e)}{r\,V_e}\right) \quad (8)$$

$$\beta_M = \frac{M_1X_{11} + M_2X_{21} + \cdots . + M_lX_{l1}}{r} \quad (9)$$

$$\text{Linear equation, L} = M_1\hat{M}_1 + M_2\hat{M}_2 + \cdots + M_i\hat{M}_i \quad (10)$$

$$\text{Effective Divider, r} = M_1^{\,2} + M_2^{\,2} + \cdots + M_i^{\,2} \quad (11)$$

$$\text{Total Variation, } S_T = \hat{M}_1^{\,2} + \hat{M}_2^{\,2} + \cdots + \hat{M}_i^{\,2} \quad (12)$$

$$\text{Variation of proportional term, } S_\beta = \frac{L^2}{r} \quad (13)$$

$$\text{Error variation, } S_e = S_T - S_\beta \quad (14)$$

$$\text{Error Variance, } V_e = \frac{S_e}{n-1} \quad (15)$$

$$\text{SN ratio, } \eta = 10 \log\left(\frac{(S_\beta - V_e)}{r\,V_e}\right) \quad (16)$$

Depending on a total number of features, the type of orthogonal array (OA) to be used will need to select accordingly with 2 level types. Value of estimation for each sample calculated using equation 1 will need to be multiplied by the actual value of output data as in equation 10. Following the combination of L_{12} as in table 1, if the control factors are no 1, it means that factor will be included as X_{ij} in equation 1 else if the control factor are no 2 it will not includ and assume as zero value. Equations 11 until equation 16 are the remaining formula to generate the SN ratio for the final optimization result. If the different between used item and not used item in each variable is high, it is high possibility showing that contribution of that factor is relatively significant. Once the overall model elements been identified, the integrated estimate value ($\hat{Y}$) for unknown data can be calculated easily. Average unit space value of the output called (M_0) will need to be added with the Integrated estimate output value $\mathbf{M}_i$ which can be relying on optimum feature selection or considering all features will provide the estimate value of ($\hat{Y}_1$) for the unknown data No.1. The procedure repeated accordingly for the remaining unknown data available.

$$\hat{Y}_1 = \hat{M}_1 + M_0 \quad (17)$$

Table 1: L_{12} Orthogonal Array combination

	Control Factors											
sample no.	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**SNR (db)**
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 "Item will be used" and 2 "Item will *not* be used".												

3.2. Artificial Bee Colony As Feature Selection Optimization In T-Method

By following the procedure of T-Method and equation shared in the previous section, the overall prediction can be calculated if analysts tend to assume all factors are important. However, since the requirement on most of the prediction analysis is to focus only on significant variables and knowing the limitation of OA in dealing with high-dimensionally data, this is where the Artificial Bee Colony (ABC) will be considered into the loop of T-Method replacing the OA. Figure 3 shows the overall review of the application of ABC in T-Method which specifically focus on optimization stage prior final prediction been calculated. Applying the ABC with the binary method in this study ease the overall programming flow since it is likely OA concept. This study purposely maintaining the SNR as the objective function, due to its insensitivity to variation. What follows is a brief overview of how ABC been practiced within this study.

ABC algorithm will randomly produce initial distributed position based on population size and a total number of features defined until initial fitness amount (SNR value) calculated. The investigation will continue by employed bees which will make a little change based on their local knowledge memory and generate new source. The fitness amount has then been compared to initial sources and higher value will be memorized while the lower will be forgotten. If previous fitness is higher, the value will maintain as it is. Once the job completed, the employed bee will then share the information of new position to onlooker bees which are in the dance area. The Onlooker bee will then evaluate the new position (nectar) and choose the food source relying on the probability value. The onlookers will modify the position if the criteria fulfilled and fitness amount will be calculated. Same as employed bee, higher fitness value will be memorized while lower will be forgotten. While onlooker and employed bee performing the exploitation process, scout bee will discover for new food source randomly in the search area. The cycle repeated until the requirement met.

Overall procedure of T-method is basically following the S.Teshima et al., (2012) [6] but for 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 basically following the approach introduced by D. Karaboga and B. Bastuk (2007) [31]. Equation 18 summarized the approaches used in fulfilling this intention but also an enhancement approaches made by E. Zorarpaci et al. (2015) [30]. X_{r1} and X_{r2} in equation 18 are the random numbers in terms of (0,1) while *r1* and *r2* is just a random number between 0 to 1 and X_i is the initial food source. After different component been identified, the random component of *r1* and *r2* used as the coefficient of the neighbor solution. The fitness value is calculated, and minimum fitness value will lead to better SNR value in fulfilling the objective function to maximize the SNR. Figure 2 gives a better illustration of how ABC is adapted within T-Method. The objective function stated in figure 2 is similar to the SNR equation used in table 1 as mentioned earlier the robustness of SNR is the key element to be maintained and applied within this combination approach.

$$if\,ference\ component = \begin{cases} 0 & , if\left({X_{r1}}^{jrand} = {X_{r2}}^{jrand}\right) \\ {X_{r1}}^{jrand} & , otherwise \end{cases}$$

$$neighbor = X_i \qquad (18)$$

$$neighbor^{jrand} = \begin{cases} 1 & , if(difference\ component = 1\ and\ r1(0,1) > r2(0,1)) \\ {X_i}^{jrand} & , otherwise \end{cases}$$

March 2017 for the prediction of unknown data while data in 2010 was used for the prediction of

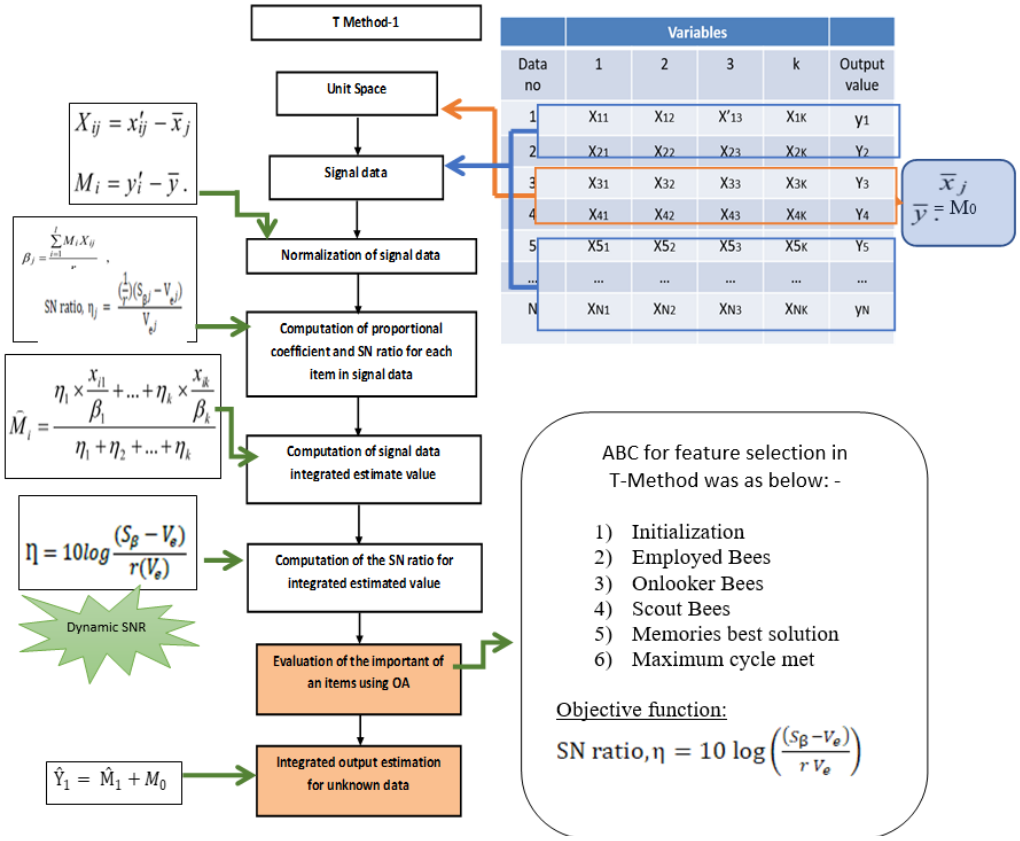


Fig. 2: ABC algorithm adaption into T-Method feature selection stage

4. Result and Discussion

The construction of T-Method, OA and ABC algorithm was done using Matlab R2017a application software. Analysis and discussion within this paper were mainly on a case study conducted at the Nuclear Agency Malaysia in predicting the power consumption of thermal energy storage system (TES) in their plant area.

4.1 Case Study: District Cooling Thermal Energy Storage System

District cooling thermal energy storage (DC-TES) contains the provision of Chilled water to the air-conditioning system as the efficient way of cooling the multiple buildings and facilities through a central plant. Thermal Energy Storage helps in reducing the peak power demand which leads to lower operating energy costs due to lower demand charges and/or lowers energy usage time frame charges. Aware on the district cooling system cost saving gain, the study conducted intentionally for making a prediction on daily power consumption so that an optimum parameter control can be relied on to achieve better performance. Data from 15 parameters *(DP- speed (%), delta-pressure(bar), STL-supply temp (°C), STL-return temp(°C), STL-PHE supply (°C), Chiller1-supply temp*(°C), *Chiller1-return temp*(°C), *Chiller2-supply temp*(°C), *Chiller2-return temp*(°C), *Chiller2-PHE supply (°C), Ts(°C), Tr(°C), KW(r)/KW(e), Chiller1(KW), Chiller2(KW))* was recorded in hourly basis which taken from month of January until signal data model creation. In order to simplify the analysis, an average of daily usage is calculated. The data classification of DC-TES within this study was divided into three stages *(charging stage, discharge stage, and transition stage*) and been analyzed separately. In the charging state, the storage tank (STL) been cooled by chillers system to store cold energy. This cold energy was then discharged to the system. The transition stage is the stage between the transition of charging state to discharge state and vice versa. During the transition, the temperature will extremely drop or increase until it reaches stable temperature for charging and discharge state. Figure 3 below shows the time frame for each state within 24 hours. Charging state will be done during night shift while discharge state during the day shift. Since the charging state is the one consumes the highest power or energy consumption, this research is mainly covered analysis of charging state only.

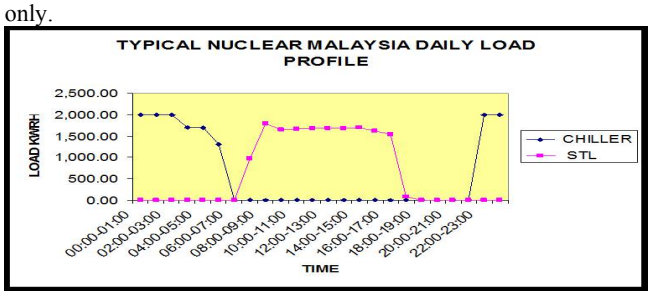


Fig. 3: Thermal Energy Storage (TES) - Load Profile

In measuring the prediction result accuracy, it is important to ensure the accuracy value is higher as possible. For this study, the accuracy been calculated relies on the formula calculating the mean relative error (MRE) as stated in equation 19.

$$MRE = 100 . \frac{1}{s} \sum_{i=1}^{s} \frac{|y_i - \hat{y}_i|}{y_i} \quad (19)$$

In order to see the strength between actual and estimated value of signal data to validate the model in equation 1, the value of correlation coefficient (R^2) is calculated resulted to the value of 0.8264 as stated in figure 4 below. This reflected that the baseline model created able to provide a strong correlation between predicted and actual data and it should able to provide good prediction value for the unknown output data in future. Prediction analysis of the two unknown output data which considering all features as important features shows that the error percentages accuracy using MRE equation was basically fallen within 3.19% and 3.31%. This low error rate is possible by referring to the good correlation coefficient (R^2) value that has been mentioned just now.

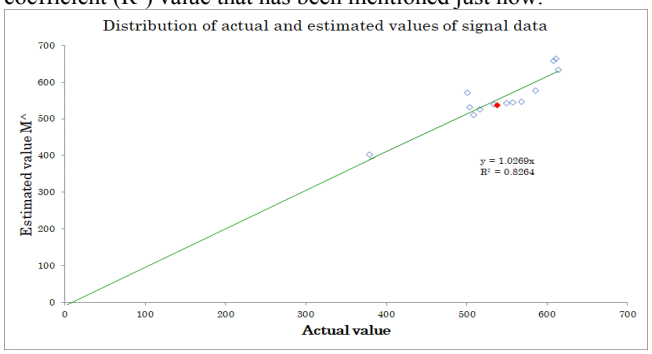


Fig. 4: Distribution of actual and estimated values of signal data

Figure 5 shows the scatter plot for all the features. If looking at the spread of the raw data, *KW(r)/KW(e), Chiller1(KW)* and *Chiller2(KW)* are having less variation since its fall within ± the projected line. However, relying only on the scatter plot data spread might not be a good reference since the normality and outliers identification is a need to understand the behavior of the data population. Despite that, for this case study, the analysis on R^2 should be enough to represent the strengths of the model equation since the R^2 value can be considered high. The same variables (*KW(r)/KW(e), Chiller1(KW) and Chiller2(KW)*) are the selected features in T-Method + Spearman Correlation analysis with error percentages of 3.06% and 3.16%. The analysis for Spearman Correlation was done using the minitab17 software while the prediction accuracy was evaluated using Matlab code.

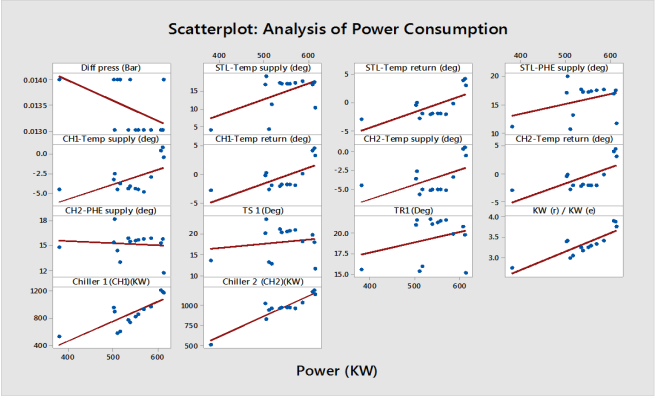


Fig. 5: Scatter plot for Power consumption

Figure 6 illustrates the OA factorial effect on SNR gain using L_{16}. Level 1 in x-axis represents the variable is used (included) within the analysis, while level 2 represents the variable is not used (excluded) within the analysis. The y-axis represents the average SNR value for each level in the respective variables following the combination shared in table 1. Among all 15 parameters, parameter 15th is obviously contributing to a negative effect on power consumption since the average SNR is higher when not considering the feature rather than considering it. Having said that, supposedly considering the 15th parameter into the analysis will directly affect the whole prediction but the MRE (considering all features) is still low even considering the 15th feature. As for this study, a features with higher level 1 average SNR value will be considered as important features in T-Method+OA analysis which are *(DP- speed (%), delta-pressure(bar), STL-return temp(°C), Chiller1-supply temp*(°C), *Chiller1-return temp*(°C), *Chiller2-supply temp*(°C), *Chiller2-return temp*(°C), *Tr(°C))*. By relying on these features only, estimation of the two unknown data leads to an error percentage of 2.67% and 2.81%.

As for the ABC analysis, maximum cycle number used for this study was 1000 with 100 run and 50 number of colony size. Mohammad Shokouhifar Shahid (2014) used a colony size of 50 for their case study due to convergence speed acceptance [32]. While E. Zorarpaci et al., (2015) used colony size 50 and the maximum cycle of 500 on their case [30]. M.S. Uzer et al., (2013) uses 100 runs with 300 maximum run cycle and 20 number of colony size [33]. The parameter defined within this study is using the greater value of maximum cycle number, run and colony size which should cater the risk of high error rate due to non-rigorous exploration. The selection of optimum features was based on the one selected more than 90 times out of 100 runs. Figure 7 shows the evolution of the optimum global best of SNR value. The final optimum parameter was given by (*delta-pressure(bar), STL-supply temp (°C) and Chiller2(KW)).* After considering all these, the result of the two numbers of unknown data shows that the error percentages are between 2.45% ~2.53% accuracy.

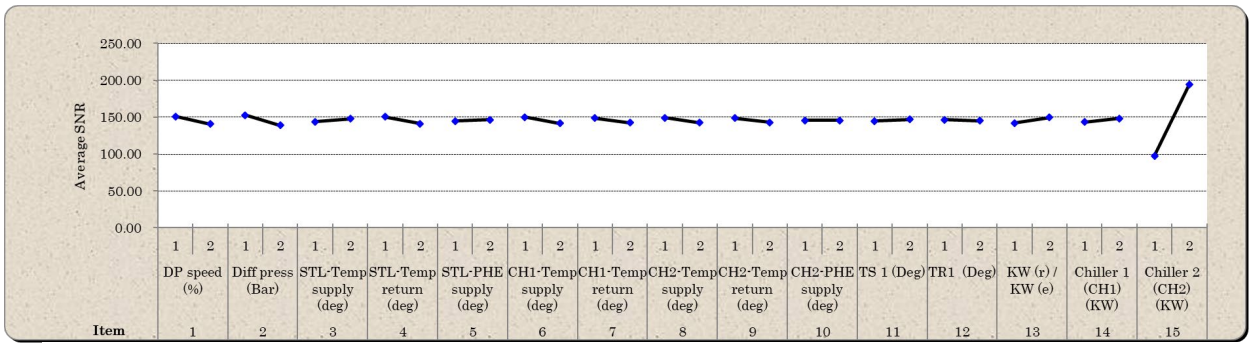


Fig. 6: Orthogonal array factorial effects for 15 features

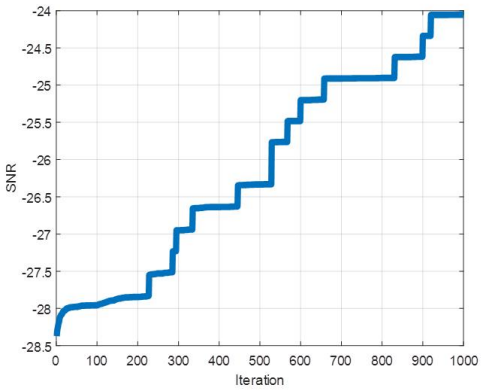


Fig. 7: Evolution of most optimum global best value of SNR function

Results show that T-Method + ABC with only three features selection giving fewer error percentages compared to T-Method + Spearman Correlation and T-Method + OA. Table 2 shows the summary across 4 different combinations including considering all features as important.

Table 2: Prediction accuracy and feature selection for four combinations

	T-Method	T-Method + Spearman Correlation	T-Method + OA	T-Method + ABC
Best/optimum features	All (15 features)	#13,14,15	# ,2,6,7, 8,9,12	#2,3,15
Prediction accuracy: Unknown data 1	3.31%	3.16%	2.81%	2.45%
Prediction accuracy: Unknown data 2	3.19%	3.06%	2.67%	2.53%

The result shows some similarity between T-Method + ABC and T-Method + Spearman Correlation in selecting features 15th as an important feature but it's been contradicted to T-Method + OA. By looking at the overall trend, the prediction accuracy results didn't vary much but the variation is due to the different method used. T-Method +OA represent better result which full fill the main intention of this paper which is to share an adaption of swarm intelligent method called Artificial Bee Colony (ABC) as feature selection tools to find optimum features in making a prediction on the power consumption of district cooling thermal energy storage system.

However, by relying on this case study only might not able to conclude that T-Method + ABC will work well for all case study involving prediction accuracy but if the baseline model is having a high correlation value which more than 0.8, it is believed that the method will provide good result compared to the other three methods shared in this study. Despite that, we believe that there are various enhancement needed in emphasizing the T-Method + ABC, especially in adapting the cross-validation method in increasing the accuracy. Test on several case studies will be very much needed in ensuring the validity and accuracy of this method.

5. Conclusion

Overall, this study strengthens the idea that prediction on certain issues can be determined by using optimum feature selection only instead of overall features that lead to cost and computer burden during the analysis stage. T-Method + ABC proved in this case study conducted as the best tool in dealing with parameter optimization with low error % impact to overall prediction. However, further work needs to be done in emphasizing the critical element within ABC algorithm as well as the T-method procedures so that when dealing with low correlation coefficient data, the risk towards high error % accuracy is lesser and avoidable. This method is applicable when dealing with high correlation coefficient baseline model, but if beyond that, it might need something extra to ponder.

Acknowledgment

This study was financially supported by the Ministry of Higher Education (MOHE), Malaysia under the Fundamental Research Grant Scheme (FRGS) Vote: 15H43 and Research Management Centre (RMC), Universiti Teknologi Malaysia.

References

[1] J. Drgoňa, D. Picard, M. Kvasnica, and L. Helsen, "Approximate model predictive building control via machine learning," *Appl. Energy*, vol. 218, pp. 199–216, May 2018.

[2] J. Reynolds, Y. Rezgui, A. Kwan, and S. Piriou, "A zone-level, building energy optimisation combining an artificial neural network, a genetic algorithm, and model predictive control," *Energy*, vol. 151, pp. 729–739, May 2018.

[3] S. Mohsen, Z. Yusop, J. Milad, and Y. Fadhilah, "Development of Generalized Feed Forward Network for Predicting Annual Flood (Depth) of a Tropical River," *Sains Malaysiana*, vol. 43, no. 12, pp. 1865–1871, 2014.

[4] S. Ding, K. W. Hipel, and Y. Dang, "Forecasting China's electricity consumption using a new grey prediction model," *Energy*, vol. 149, pp. 314–328, Apr. 2018.

[5] N. H. A. Rahman, M. H. Lee, M. T. Latif, and Suhartono, "Evaluation Performance of Time Series Approach for Forecasting Air Pollution Index in Johor , Malaysia," *Sains Malaysiana*, vol. 45, no. 11, pp. 1625–1633, 2016.

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

[7] S. Negishi, Y. Morimoto, S. Takayama, and A. Ishigame, "Daily Peak Load Forecasting by Taguchi's T Method," *Electr. Eng. Japan (English Transl. Denki Gakkai Ronbunshi)*, vol. 201, no. 1, pp. 57–65, 2017.

[8] P. Dasneogi, E. Cudney, A. Adekpedjou, and R. Kestle, "Comparing the Predictive Ability of T-Method and Cobb-Douglas Production Function for Warranty Data," in *Proceedings of the ASME 2009 International Mechanical Engineering Congress & Exposition*, 2015, pp. 1–6.

[9] H. Kawada and Y. Nagata, "An application of a generalized inverse regression estimator to Taguchi ' s T-Method," *J. Japanese Soc. Qual. Control*, no. 1985, pp. 3–4, 2015.

[10] J. Inoh, Y. Nagata, K. Horita, and M. Arisa, "PredictionAccuracies of lmproved Taguchi s T Methods Compared to those of MultipleRegression Analysis" *J. japanese Soc. Qual. Control*, vol. 42, no. 2, pp. 265–277, 2012.

[11] E. A. Cudney and P. A. Shah, "Predicting Annual Precipitation Using the T-Method Missouri University of Science and Technology," in *Proceedings of the 2010 Industrial Engineering Research Conference*, 2010.

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

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

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

[15] S. Taguchi, R. Jugulum, G. Taguchi, and J. O. Wilkins, "Discussion," *Technometrics*, vol. 45, no. 1, pp. 16–21, 2003.

[16] K. Tsui, T. Sukchotrat, and V. C. P. Chen, "A comparison study and discussion of the Mahalanobis-Taguchi System Seoung Bum Kim," *Int. J. Ind. Syst. Eng.*, vol. 4, no. 6, pp. 631–644, 2009.

[17] P. Saraiva, N. Faísca, R. Costa, and A. Gonçalves, "Fault identification in chemical processes through a modified mahalanobis-taguchi strategy," *Comput. Aided Chem. Eng.*, vol. 18, no. C, pp. 799–804, 2004.

[18] C. R. Foster, R. Jugulum, and D. D. Frey, "Evaluating an adaptive

One-Factor-At-a-Time search procedure within the Mahalanobis-Taguchi System," *Internatinal J. Ind. Syst. Eng.*, vol. 4, no. 6, pp. 600–614, 2009.
[19] 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 Syst. Appl.*, vol. 37, no. 2, pp. 1286–1293, 2010.
[20] E. Reséndiz and C. a. Rull-Flores, "Mahalanobis-Taguchi system applied to variable selection in automotive pedals components using Gompertz binary particle swarm optimization," *Expert Syst. Appl.*, vol. 40, no. 7, pp. 2361–2365, 2013.
[21] E. Reséndiz, L. a. Moncayo-Martínez, and G. Solís, "Binary ant colony optimization applied to variable screening in the Mahalanobis-Taguchi System," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 634–637, 2013.
[22] A. S. Iquebal and A. Pal, "Artificial bee colony optimisation-based enhanced Mahalanobis Taguchi system for classification," *Int. J. Intell. Eng. Informatics*, vol. 2, no. 2–3, pp. 1–14, 2014.
[23] F. Ramlie, K. R. Jamaludin, R. Dolah, and W. Z. A. W. Muhamad, "Optimal feature selection of Taguchi character recognition in the Mahalanobis-Taguchi system using bees algorithm," *Glob. J. Pure Appl. Math.*, vol. 12, no. 3, pp. 2651–2671, 2016.
[24] C. G. Mota-Gutiérrez, E. O. Reséndiz-Flores, and Y. I. Reyes-Carlos, "Mahalanobis-Taguchi system: state of the art," *Int. J. Qual. Reliab. Manag.*, vol. 35, no. 3, pp. 596–613, 2018.
[25] E. O. Reséndiz-Flores and M. E. López-Quintero, "Optimal identification of impact variables in a welding process for automobile seats mechanism by MTS-GBPSO approach," *Int. J. Adv. Manuf. Technol.*, vol. 90, no. 1–4, pp. 437–443, 2017.
[26] A. L. Bolaji, A. T. Khader, M. A. Al-Betar, and M. A. Awadallah, "Artificial bee colony algorithm, its variants and applications: A survey," *J. Theor. Appl. Inf. Technol.*, vol. 47, no. 2, pp. 434–459, 2013.
[27] E. Hancer, B. Xue, D. Karaboga, and M. Zhang, "A binary ABC algorithm based on advanced similarity scheme for feature selection," *Appl. Soft Comput. J.*, vol. 36, pp. 334–348, 2015.
[28] F. G. Mohammadi and H. Sajedi, "Region based Image Steganalysis using Artificial Bee Colony," *J. Vis. Commun. Image Represent.*, vol. 44, pp. 214–226, 2017.
[29] P. Shunmugapriya and S. Kanmani, "A hybrid algorithm using ant and bee colony optimization for feature selection and classification (AC-ABC Hybrid)," *Swarm Evol. Comput.*, vol. 36, no. February 2016, pp. 27–36, 2017.
[30] E. Zorarpaci, S.A Ozel and S. Gungor, "An Artificial Bee Colony Based Algorithm for Feature Selection," vol. 30, no. June, pp. 25–31, 2015.
[31] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm," *J. Glob. Optim.*, vol. 39, no. 3, pp. 459–471, 2007.
[32] Mohammad Shokouhifar Shahid, "An Artificial Bee Colony Optimization for Feature Subset Selection using Supervised Fuzzy C-Means Algorithm," in *International Conference on Information Security and Artificial Intelligence (ISAI 2010)*, 2010, pp. 427–432.
[33] M. S. Uzer, N. Yilmaz, O. Inan, M. S. Uzer, N. Yilmaz, and O. Inan, "Feature selection method based on artificial bee colony algorithm and support vector machines for medical datasets classification.," *ScientificWorldJournal.*, vol. 2013, p. 419187, 2013.