HYBRID BOOTSTRAP-BASED APPROACH WITH BINARY ARTIFICIAL BEE COLONY AND PARTICLE SWARM OPTIMIZATION IN TAGUCHI'S T-METHOD

NOLIA HARUDIN

UNIVERSITI TEKNOLOGI MALAYSIA

HYBRID BOOTSTRAP-BASED APPROACH WITH BINARY ARTIFICIAL BEE COLONY AND PARTICLE SWARM OPTIMIZATION IN TAGUCHI'S T-METHOD

NOLIA HARUDIN

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia

SEPTEMBER 2020

DEDICATION

I dedicate this thesis,

To my parent, ***Harudin Abu Aziz & Zakiah Abd.Rahman***. In the light of their relentless prayers, trust and sacrifice, I have had the chance to achieve a significant part of my dream that I have wished for. Thank you very much, Mak&Ayah.

To my sisters ***Haryati Harudin & Harfizah Harudin***. For more than three decades, we grew up together, pursued our dreams, and prayed Allah for the best of each other. You are my source of inspiration. My sisters, thank you very much

To my husband, my soulmate, my best friend, **Zulkifli Marlah@Marlan**. Thank you for your prayers, your compassion, your generosity, your encouragement and your faith in me until today. Together we strive for success. May your bright stars, be more glorious my love. Thank you very much papa.

To my beloved, ***Nur Naira Zamira Zulkifli, Nur Zaara Nafeesa Zulkifli and Muhammad Nafiz Zuhair Zulkifli.*** Thank you for all the joy you have brought to us. May your bright stars be even more glorious, my love. Thank you Loves.

To my family-in-law, ***Mak, Abah, Kakyong&family, Kakya&family, Alang, Anis&familys, ijat&family, Tika and Usop.*** Thank you so much for your prayers and encouragement. May the relation between us stay stronger.

ACKNOWLEDGEMENT

In the name of Allah, the Most Beneficent, the Most Merciful

Alhamdulillah to Allah the Almighty, because of His grace, I have been able to finish my studies. I would like to show my deepest gratitude to my supervisors, AP. Dr. Khairur Rijal Jamaludin, Dr. Faizir Ramlie, and Dr. Nabil Muhtazaruddin for their encouragement, guidance, critique, motivation, and friendship during my hard and fruitful research days. Without their continued support and interest, this thesis would not have been the same as presented here.

Special appreciation to UNITEN and the Ministry of Education (myBrain15 scholarship) for providing me with financial support throughout this research.

My sincere thanks to experts at Genichi Taguchi Centre, especially Prof. Ir. Dr. Sha'ari, Dr. Rozzeta, Engineer Shin Taguchi (son of the late Genichi Taguchi) , and Dr. Rajesh Jugulum for their valuable thought and support given to me. Special gratitude to Shin Taguchi sensei for giving me consent to utilize the JD power dataset for this exploration. Not to overlooked En Yusnan and En Hamzah from Agensi Nuklear Malaysia for granting me permission to use the Chiller's Performance report data for the thesis and publication purposes.

My fellow postgraduate student should also be recognized for their support. My sincere appreciation also extends to all my colleagues, Dr. Wan Zuki, Najmiah, Zulkifli, Munira, and others who have assisted at various occasions. Thank you to my friends who have always given me moral and spiritual encouragement, Noor Minahuda, Dr. Juniza, Che Raidah Ayu, Norfaizah, Azieyanti, Noorilyana, Nurul Azwa, Nur Irmawati and others whom I may have missed to name here.

I would also like to take this opportunity to thank the Warisan Dayang family for their prayers and the encouragement they have provided me during this journey, particularly Busu, Pak usu, Maklang, Makcik, Maklong, Bangah Faiz, who have helped me with the task of proofreading and all my other relatives whom I have not been able to mention here. I really appreciate all your support and prayers.

To all my teachers and lecturers from SK Jalan 6 B.B. Bangi, SK Simpang Durian, MRSM Kuala Klawang, SM Teknik Ampangan, UTMKL, UTM Skudai & Razak Faculty of Technology and informatics, thank you for all the knowledge been shared to me. Only Allah is able to repay your kindness. To those who have supported me directly or indirectly in completing my research work, whose names are not mentioned here, thank you.

Last but not least, thank you to the most important person in my life, my husband, Zulkifli Marlah @ Marlan, my mother, Zakiah Binti Abd Rahman, and my father, Harudin Bin Abu Aziz. Without your love, prayers, support, and inspiration, all of this would never be possible.

ABSTRACT

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. When evaluating data using a system such as the Taguchi's T-Method, bias issues often appear due to inconsistencies induced by model complexity, variations between parameters that are not thoroughly configured, and generalization aspects. In Taguchi's T-Method, the unit space determination is too reliant on the characteristics of the dependent variables with no appropriate procedures designed. Similarly, the least square-proportional coefficient is well known not to be robust to the effect of the outliers, which indirectly affects the accuracy of the weightage of SNR that relies on the model-fit accuracy. The small effect of the outliers in the data analysis may influence the overall performance of the predictive model unless more development is incorporated into the current framework. In this research, the mechanism of improved unit space determination was explicitly designed by implementing the minimum-based error with the leave-one-out method, which was further enhanced by embedding strategies that aim to minimize the impact of variance within each parameter estimator using the leave-one-out bootstrap (LOOB) and 0.632 estimates approaches. The complexity aspect of the prediction model was further enhanced by removing features that did not provide valuable information on the overall prediction. In order to accomplish this, a matrix called Orthogonal Array (OA) was used within the existing Taguchi's T-Method. However, OA's fixed-scheme matrix, as well as its drawback in coping with the high-dimensionality factor, leads to a sub-optimal solution. On the other hand, the usage of SNR, decibel (dB) as its objective function proved to be a reliable measure. The architecture of a Hybrid Binary Artificial Bee Colony and Particle Swarm Optimization (Hybrid Binary ABC-PSO), including the Binary Bitwise ABC (BitABC) and Probability Binary PSO (PBPSO), has been developed as a novel search engine that helps to cater the limitation of OA. The SNR (dB) and mean absolute error (MAE) were the main part of the performance measure used in this research. The generalization aspect was a fundamental addition incorporated into this research to control the effect of overfitting in the analysis. The proposed enhanced parameter estimators with feature selection optimization in this analysis had been tested on 10 case studies and had improved predictive accuracy by an average of 46.21% depending on the cases. The average standard deviation of MAE, which describes the variability impact of the optimized method in all 10 case studies, displayed an improved trend relative to the Taguchi’s T-Method. The need for standardization and a robust approach to outliers is recommended for future research. This study proved that the developed architecture of Hybrid Binary ABC-PSO with Bootstrap and minimum-based error using leave-one-out as the proposed parameter estimators enhanced techniques in the methodology of Taguchi's T-Method by effectively improving its prediction accuracy.

ABSTRAK

Kaedah-T *Taguchi's* adalah salah satu teknik ramalan yang dijadualkan oleh sistem *Mahalanobis Taguchi (MTS)* yang telah dibuat secara khusus tetapi tidak terhad kepada data sampel kecil yang multivariasi. Semasa menilai data menggunakan sistem seperti Kaedah-T *Taguchi*, masalah nilai pincang sering dilihat disebabkan oleh ketidakseragaman yang disebabkan oleh model yang kompleks, variasi antara pemboleh ubah yang tidak dikonfigurasi secara menyeluruh, serta aspek generalisasi. Dalam kaedah-T *Taguchi*, penentuan unit ruang terlalu bergantung pada ciri-ciri pemboleh ubah bersandar, tanpa prosedur yang sesuai dirancang. Begitu juga, kuasa dua terkecil- pekali berkadar diketahui tidak teguh terhadap kesan titik terpencil, yang secara tidak langsung mempengaruhi ketepatan pemberat *SNR* yang bergantung pada ketepatan model. Kesan kecil dari titik terpencil dalam analisis data dapat mempengaruhi prestasi keseluruhan model ramalan kecuali penambahbaikan lebih banyak digabungkan ke dalam kerangka semasa. Dalam penyelidikan ini, mekanisma penentuan unit ruang yang lebih baik dirancang secara eksplisit dengan menerapkan kaedah ralat berdasarkan minimum bersama kaedah *leave-one-out,* yang selanjutnya ditambahbaik dengan menerapkan strategi yang bertujuan untuk meminimumkan impak variasi dalam setiap penganggar parameter menggunakan pendekatan *leave-one-out bootstrap (LOOB)* dan *0.632 estimate*. Aspek kompleksiti dalam model ramalan ditingkatkan dengan membuang pemboleh ubah tidak bersandar yang tidak memberikan maklumat berharga mengenai ramalan keseluruhan. Untuk mencapai matlamat ini, matriks yang disebut tatasusun ortogon *(OA)* digunakan dalam kaedah-T *Taguchi* sedia ada. Walau bagaimanapun, ketetapan matriks tetap *OA*, serta kekurangannya untuk mengatasi faktor dimensi tinggi, membawa kepada penyelesaian yang sub-optimum. Namun, penggunaan *SNR, desibel (dB)* sebagai fungsi objektifnya terbukti menjadi ukuran yang boleh dipercayai. Seni bina *Hybrid Binary Artificle Bee Colony* dan *Particle Swarm Optimization (Hybrid Binary ABC-PSO)*, termasuk *Binary Bitwise ABC (BitABC)* dan *Probability Binary PSO (PBPSO),* telah dibangunkan sebagai enjin carian baru yang membantu memenuhi kekurangan *OA*. *SNR (dB)* dan ralat mutlak purata *(MAE)* adalah sebahagian daripada ukuran prestasi yang digunakan dalam penyelidikan ini. Aspek generalisasi adalah penambahan asas yang digabungkan dalam penyelidikan ini untuk mengawal kesan *overfitting* dalam analisis. Penganggar parameter optimum yang dicadangkan dengan pengoptimuman pemilihan ciri dalam analisis ini telah diuji pada sepuluh (10) kajian kes dan telah menunjukkan peningkatan ketepatan ramalan dengan purata 46.21% bergantung kepada kes. Sisihan piawai purata kepada MAE, yang menggambarkan kesan kebolehubahan kaedah yang dioptimumkan dalam semua 10 kajian kes, menunjukkan tren yang lebih baik berbanding dengan kaedah-T *Taguchi*. Keperluan untuk piawaian dan pendekatan yang lebih baik untuk titik terpencil disyorkan untuk penyelidikan masa depan. Kajian ini membuktikan bahawa seni bina *Hybrid Binary ABC-PSO* yang dibangunkan dengan *Bootstrap* dan kaedah ralat berdasarkan minimum bersama kaedah *leave-one-out* sebagai penganggar parameter optimum yang dicadangkan meningkatkan teknik dalam metodologi kaedah-T Taguchi dengan meningkatkan lebih ketepatan ramalan dengan berkesan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MTS	-	Mahalanobis Taguchi System
MD	-	Mahalanobis Distance
SNR	-	Signal to Noise Ratio
OA	-	Orthogonal Array
GIR	-	Generalized Inverse Regression Estimator
MML	-	Median Median Line
BABC	-	Binary Bitwise Artificial Bee Colony
BPSO	-	Binary Particle Swarm Optimization
LOO	-	Leave-One -Out
LOOB	-	Leave-One –Out Bootstrap
dB	-	Decibel
MAE	-	Mean Absolute Error
MLR	-	Multiple Linear Regression
SD	-	Standard Deviation
OLS	-	Ordinary Least Squares
UCI	-	University of California at Irvine
R-LCS	-	Rote Learning Classifier System
HL	-	Hodges Lehman
SB	-	Shamos Bickel
RMSE	-	Root Mean Squared Error
OOB	-	Out-of-Bag
MCN	-	Maximum Cycle Number
DBPSO	-	Discrete Binary Particle Swarm Optimization
MBPSO	-	Modified Discrete Binary Particle Swarm Optimization

LIST OF SYMBOLS

R^2	-	Coefficient of Determination
β	-	Proportional Coefficient
Z	-	Normalized Signal Data
σ^2	-	Variance
x	-	Output in (y-axis)
$\widehat{M}_{ii,j}$	-	Estimate of Output Variables
D	-	Number of of Features/Variables
$SS_{residual}$	-	Sum Squared Residual
M_i	-	Normalized Output Signal Data
$OA_N(S^m)$,	-	Orthogonal array consists of N x m array of multifactor combination having S test level, N number of run, and m number of variables.
$L_a(b^c)$	-	L for Latin squares, 'a' is the number of rows or run, 'b' is the level of variable and 'c' is the number of variables
l	-	Number of signal data
$Z_{ii,j}$	-	Normalized signal data
$SS_{residual}$	-	Sum Squared Residual
S_β	-	Variation of proportional term
S_T	-	Total variation
S_e	-	Error variation
V_e	-	Error variance
e_{Loocv}	-	Error prediction of leave-one-out cross validation
B	-	Maximum bootstrap size
NP	-	Size of employeed and onlooker bees (Colony size)
v	-	Food source position
P_i	-	Probability of solution x
N	-	Number of particles
p_i^t	-	Particle position at current iteration
p_g^t	-	Global particle position at current iteration

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Overview

The futures of industrial technologies look promising but somehow are always, to some extent, uncertain and unpredictable. Uncertainty and unpredictability are the focus of long-standing methodological concern since they are inherently related either to some form of modeling operation or to scientific risk considerations. Disruption tends to be a daily challenge to ensure a robust design, robust operating system as well as reliable performance and responsive risk control are all well responded and predicted. Much of our current issues of uncertainty have specific causes. How well we leverage it with the advanced cognitive algorithms, predictive modeling, and statistical analysis towards better industrial systems were aggressively discussed among analysts. Practitioners have used various prediction tools in numerous areas for the past few decades, and it is progressively enhanced and improved.

In 1930, Dr. Prasanta Chandra, a prominent Indian statistician, developed a concept called Mahalanobis Distance (MD), a distance-based statistical approach used to distinguish similarities of a population group from another. In the 1980s, Dr. Genichi Taguchi developed the Mahalanobis Taguchi System (MTS) as a pattern recognition technique that blends Mahalanobis Distance (MD) theory and Robust Engineering concept to systematically and effectively classify and predict data in a multidimensional environment (Taguchi, Chowdhury and Wu, 2001; Teshima, Hasegawa and Tatebayashi, 2012c). 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 MT-Method as a classification technique that has so far gained much attention among scholars, several methods have also been established since then, which have utilized the same integration principles. One of them is Taguchi's T-Method, which was developed specifically for predictive analysis.

The underlying principle of Taguchi's T-Method was based on a reference-point equation, which created and summed the linear regression for individual independent variables towards the dependent variables that pass through the zero-point (origin). The formulated model involved the integration of the inverse regression concept, unit space, dynamic SNR as a weighting factor, and OA, which used dynamic SNR (dB) as an objective function for the variable selection optimization mechanism. Figure 1.1 illustrates the formulation of the integrated estimated output model in Taguchi's T-method, which applying the classical least square theory for determining the proportional coefficient. Determination of unit space is done for the normalization phase, and it is said to be the population that is homogeneous concerning the target object. The selection of unit space is one of the most crucial phases to be identified in the early stages before any analysis has been carried out (Teshima, Hasegawa, and Tatebayashi, 2012b).

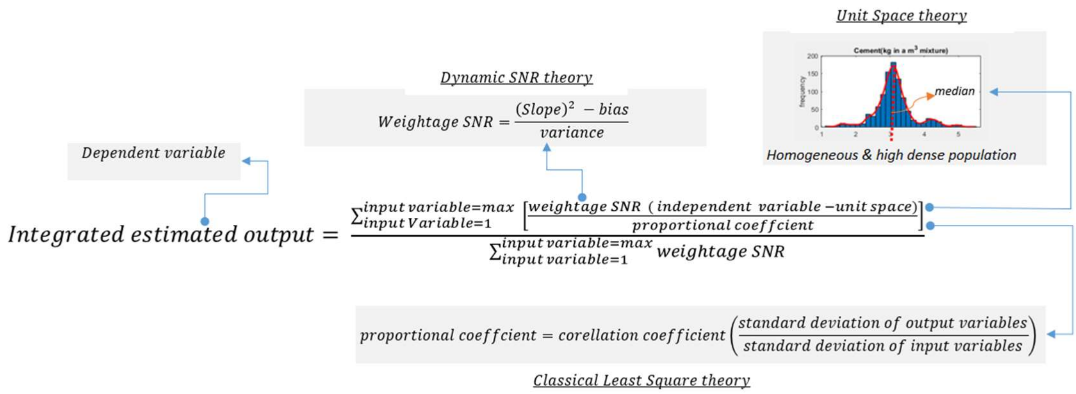

Figure 1.1 Simplified illustration of Taguchi's T-Method formulated model for integrated estimated output prediction

Taguchi's T-Method was proposed for multivariate estimation to predict the integrated estimated output value, and one of its significant advantages is its ability to predict even with limited sample data. With multiple regression analyses, there is a limitation in that the sample size has to be higher than the number of variables. On the other hand, the said limitation does not apply to the Taguchi's T-Method. Additionally, the Taguchi's T-Method too has no direct influence from multicollinearity due to consideration upon individual regression as illustrated in Figure 1.1 (Teshima, Hasegawa, and Tatebayashi, 2012a; Negishi *et al.*, 2017; Nishino and Suzuki, 2018). Unlike other optimization methods, the element of the dynamic SNR as the weightage element within Taguchi's T-Method are embedded to indirectly contribute to the robustness and sensitivity of the formulated model towards the variation effect as can be seen in Figure 1.1, in which the weighting SNR should increase as the variability decreases. Section 2.3 in Chapter 2 will explain the principle of Taguchi's T-Method in more depth.

Taguchi's T-Method has generally been practiced in Japan before and has only been practiced by non-Japanese researchers recently, owing to its simplicity and ease of interpretation as well as the advantages described earlier. Scholars using Taguchi's T-Method to conduct studies and solve various prediction analysis problems (Dasneogi and Cudney, 2009; Dasneogi, Cudney and Adekpedjou, 2009; Cudney, Shah, and Kestle, 2010). Figure 1.2 explicitly illustrates the pattern of studies performed since 2008 based on the number of papers published in the literature and how the progress is moving towards optimization of parameters and optimization of variable selection rather than just application purposes since the year 2012. From a different perspective, this increasing pattern has indirectly triggered that there are indeed a variety of enhanced approaches towards parameter and variable selection optimization available out there that can be further explored and incorporated into Taguchi's T-Method as a hybridization or integration element. However, studies related to the application of Taguchi's T-Method should be continually expanded in various fields of research in order to strengthen the theory and practicality of this method.

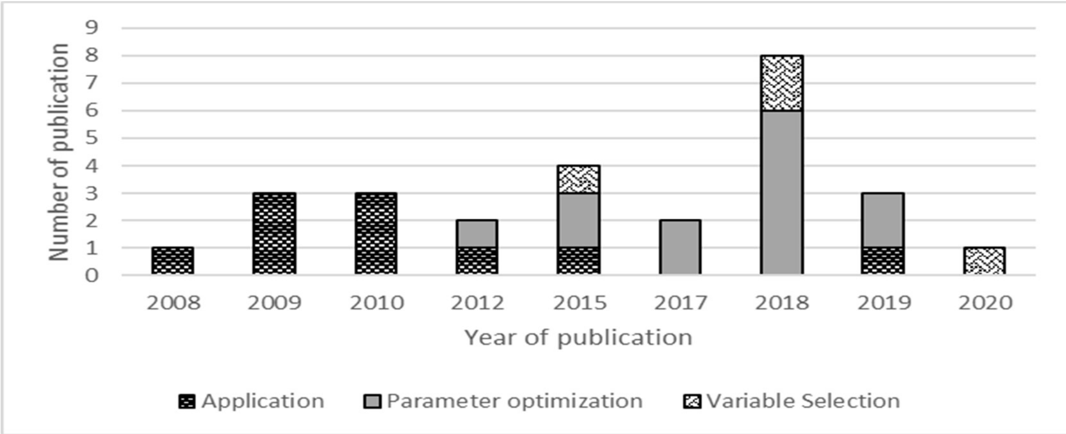

Figure 1.2 Trend of studies conducted in Taguchi's T-Method since 2008 clustered among areas of interest

1.2 Research/ Problem Background

The primary objective of developing a predictive model is to ensure that the results of previous learning observations are reliable and valid. Nonetheless, validity is typically restricted to many areas, including the need for a large dataset and risk of the small dataset, the bias in data, complexity of the model, and generalization aspect. Every limitation is, in reality, closely related to one another. In a situation involving limited sample data such as in the product development phase, there is a possibility of obtaining an inaccurate prediction model due to high variability and bias effect in the dataset (Wilcox, 2005). However, the ability to predict with a small number of sample data is a significant challenge for most researchers today. Therefore, in enhancing the model's accuracy, a large dataset is needed even though the risk for variation and predictive model bias to occurs still exist (Schmidt and Finan, 2017). A complex model that involves high dimensional variables is often a significant challenge as not all variables are significant for the model, some of which could even deteriorate the overall performance and contribute significantly to the predictive model's bias *(*Hu *et al.,* 2018). The generalization aspect defines the ability of the formulated prediction model to infer any future analysis precisely by depending on the formulated predictive model. The generalization in the context of this research is referring to the cross-validation approach, which is used to achieve satisfactory results for the training, testing, and validation phases despite limiting the impact of the overfitting issue. Undeniably, a good generalization is one of the keys to a successful future prediction (Xu and Goodacre, 2018).

Multiple regression is the most commonly employed methodological method for linear analysis and is well known for its weakness of dealing with the number of variables greater than the number of sample units, as the degree of freedom is equal or less than zero (Carrascal, Galván and Gordo, 2009). In the context of Taguchi's T-Method, limited sample data, including data set of samples that are less than the number of variables, is not a matter of concern, as the analysis carried out by Inoh et al. (2012) and Kawada and Nagata (2015) indicate that Taguchi's T-Method is performing better relative to multiple regressions up to twice the number of variables used. For instance, if the study considers six numbers of variables, the Taguchi's T-Method will estimate at least 12 numbers of samples with better execution than multiple linear regressions.

It is generally known that any prediction tools, including Taguchi's T-Method, are not exempted from the data bias issue. However, with significant data sampling size, the bias effect shall be able to cancel each other (Roberts and Russo, 1999). In overseeing data analysis such as Taguchi's T-Method, biases between actual and measured data might involve variation induced by combination features, other obscure properties, and the correlation between some parameters which are not completely optimized (Liu *et al.,* 2019). Although the SNR weightage was introduced in Taguchi's T-Method theory as an aspect of robustness and less sensitivity to variance, the possibility for variability still exists due to the sensitivity of noise variables as well as the lack of fit or non-linearity concern (Fowlkes and Creveling, 1995c). The formulated model for integrated estimated output presented in Figure 1.1 demonstrates how each of the parameter estimates *(unit space, SNR weightage, proportional coefficient)* is related to the dispersion and sensitivity to the impact of the variance. The presence of outliers could lead to an increased effect of bias and variability on each parameter estimates. This issue undoubtedly demands more research to ensure a higher degree of accuracy in the study.

Back then, it is a common practice among statisticians that the ordinary least square regression study is considered optimal with the condition that the errors are free from the effect of outliers. As a consequence, the normality assumption was believed to be randomized, independent, and equally distributed (Mazlina, Bakar, and Midi, 2017; Huber, 1973). In the context of Taguchi's T-Method, Nishino and Suzuki (2018) claimed that dealing with small samples, for example, in the case of early stages of mass production, Taguchi's T-Method, which uses the least-squares for proportional coefficient estimation, may fail to produce adequate models when the outliers are influenced. Not only in small sample cases but also large sample data, the effect of outliers was of great concern. Yoshimura and Nagata (2015), Negishi *et al.* (2017), Nakao and Nagata (2018), and Satoshi, Yasushi, and Nagata (2018) did highlight similar concern on the significant effect of outliers towards existing Taguchi's T-Method model in their research. Several efforts have been made to resolve the problem. The enhancement of least-square estimates as well implicitly leads to the improvement of SNR weightage since both estimators are relying on the accuracy of the fit-line model. If the model is fitted well, reductions on variance and bias are expected, which leads to a greater level of accuracy.

Kawada and Nagata (2015) were one of the early researchers who incorporated the Generalized Inverse Regression estimator (GIR) into Taguchi's T-Method, replacing the least-square approach for the proportional coefficient estimation. The approach was able to increase the accuracy of existing Taguchi's T-Method to a higher level. Negishi *et al.* (2017) proposed the Nonlinear Correction T-Method as an alternative way of improving the prediction of daily peak load, which involved a nonlinear regression. The results seem to provide higher accuracy, but it is only applicable for daily peak load prediction case. The study conducted to improve the parameter estimate of the proportional coefficient was performed with a variety of strategies incorporated into the existing Taguchi's T-Method which are; Circular least-squares T-Method (Satoshi, Yasushi, and Nagata, 2018) and Median-Median Line (MML) (Nishino and Suzuki, 2018). Both approaches have successfully shown a positive impact on the existing Taguchi's T-Method accuracy, but it was just tested with a very minimal case study.

Another point of interest relevant to the Taguchi's T-Method was the determination of unit space as a measuring scale in defining the target object. The unit space, which is homogeneous and originates from a highly dense population, shall be positioned in the vicinity of the average (Teshima, Hasegawa, and Tatebayashi, 2012a; Suguru and Yasushi, 2018). The current unit space theory noticed that there was so much dependency on the specified characteristics and patterns of the dependent variables without providing adequate procedures and including independent variables as part of the selection criteria. The small impact of the outliers could affect the accuracy of the optimal unit space selection. The Ta and Tb Method proposed by Inoh et al. (2012), and the embedded of homogeneity of output by Marlan *et al.* (2019) into the Taguchi's T-Method, substituting the existing unit space determination characteristics, offers another potential direction for unit space determination that is capable of having better accuracy. However, the findings are evaluated in a specific case analysis in which no particular method can be extended to all situations. Among all these proposed techniques, Ta Method is famously applied since it is done by just averaging every single variable data.

Among the shortcomings observed in all the preceding literature is the approach of the generalization that previous researchers have not thoroughly practiced. The term generalization is used to reflect on how well new data can characterize a predictive model. Over-fitting is a modeling error observed when a trained model performs exceptionally well during the training phase but perform poorly on new unknown samples resulting from the presence of certain noise levels (Xu and Goodacre, 2018). In determining the most optimum unit space, the Leave-one-out (LOO) approach is proposed in this study. The approach will test the individual sample separately. Known as a tool used to estimate errors and has been shown to provide a nearly unbiased estimate of the model's true generalization (Cawley and Talbot, 2004), LOO in this study will choose the most optimum samples that contribute to the minimum error of the training population.

However, the variability of LOO is considered to be relatively high (Kohavi, 1995; Cawley and Talbot, 2007; Elisseeff and Pontil, 2002). Concerning to the LOO risk towards high variation, a non-parametric resampling method which is called "bootstrapping," is proposed as another way to improve prediction performance due to small datasets, variation, and bias issues. Efron introduced the bootstrap method in 1979, based on the resampling approach with unknown distribution. The series of bootstrap applied in this study which are Leave-one-out bootstrap (LOOB) and approach that could help in reducing the effect of overfitting which is called 0.632 estimates capable in providing more reliable inferences when data is not statistically well-represented, or the sample size is small as well as reducing the potential of over-fitting issues to occur (Kisielinska, 2013; Samart, Jansakul, and Chongcheawchamnan, 2018). The integration of both the bootstrap series and the LOO helps to define a better accuracy of the unit space and parameter estimates *(proportional coefficient and weightage SNR)* within existing Taguchi's T-Method.

The complexity of linear regression described in this work is primarily linked to describing the set of variables used in model formulation. In literature, this process is also known as item selection, variable selection, feature selection, and dimensionality reduction, which carry the same purpose on identifying the most optimal features within the model formulation. In MTS, the orthogonal array (OA) is a feature selection search mechanism that has been established between a series of MTS methods that share standard procedures but vary in objective function determination. The element of the OA within MTS has been debated and is believed to be insufficient as it offers a sub-optimal solution (Woodall *et al.,* 2003; Pal and Maiti, 2010). Most of 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, as shown in the example of L12 array by Table 1.1 below. Abraham and Variyath (2003) argued that the fixed combination in OA is not optimal since the results may vary significantly if the column-to-column information is rearranged (Hawkins *et al.*, 2003). Foster, Jugulum, and Frey (2009) agreed with Abraham and Variyath after proving the fact with 1000 random variables to column assignment. Issues in OA have been highlighted as well by Hawkins (2003) and Tsui, Sukchotrat, and Chen (2009), 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. Therefore, developing a hybrid methodology for better accuracy is a preferred solution to these concerns.

Table 1.1 Orthogonal Array L12 and its fixed scheme assignment of variables

No	ITEMS										
	X1	**X2**	**X3**	**X4**	**X5**	**X6**	**X7**	**X8**	**X9**	**X10**	**X11**
1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	2	2	2	2	2	2
3	1	1	2	2	2	1	1	1	2	2	2
4	1	2	1	2	2	1	2	2	1	1	2
5	1	2	2	1	2	2	1	2	1	2	1
6	1	2	2	2	1	2	2	1	2	1	1
7	2	1	2	2	1	1	2	2	1	2	1
8	2	1	2	1	2	2	2	1	1	1	2
9	2	1	1	2	2	2	1	2	2	1	1
10	2	2	2	1	1	1	1	2	2	1	2
11	2	2	1	2	1	2	1	1	1	2	2
12	2	2	1	1	2	1	2	1	2	2	1

note: 1 implies the usage of the variables and 2 implies otherwise

Until recently, the OA element in the MTS classification approaches has been continuously improved by numerous machine-learning algorithms. However, the enhancement of the OA element within Taguchi's T-Method as a prediction tool is still at an initial stage and needs further attention. Kawada and Nagata (2015b) apply a stepwise forward and backward selection procedure for this purpose which, showed an increase in accuracy on many cases conducted. 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.

Hybrid Binary ABC-PSO algorithms, which have been found to provide more efficient results based on a literature study, have also been proposed. Motivated by their advantages in complementing each other's drawback, BPSO aims to allow greater use of the search mechanism in the hybrid algorithm, whereas the local ABC search mechanism helps prevent the particle (nectar) from being trapped at local optimum, which ideally contributes into an optimal amount of features selection combination.

1.3 Problem Statement

Although Taguchi's T-Method is still reasonably new to researchers outside Japan, it is now a prevalent strategy among Japanese researchers and their industry experts. It was evident in Section 1.2 that the Taguchi's T-Method prediction model had several limitations involving unit space determination procedure (Inoh *et al.*, 2012; Marlan *et al.*, 2019), parameter estimate accuracy (Negishi *et al.*, 2017; Nishino and Suzuki, 2018; Satoshi, Yasushi and Nagata, 2018), as well as the OA structure that was seen as inefficient in exploiting a higher number of interactions which prone to a sub-optimal solution (Ramlie *et al.*, 2016; Muhamad, 2019). Numerous studies are proposed to improve the accuracy of these limitations, which relate to the integrated estimate output model development.

In most cases, the assumption that receives much attention from many statisticians is that the ordinary least square-based proportional coefficient analysis must be free from the effect of outliers. The errors and distribution were assumed to be normally distributed, observations are random, independent, and identically distributed and equally reliable with no outlier in the data since even a single outlier can lead to a severe effect of bias and variation (Huber, 1973; Tarr, Weber and Muller, 2015; Mazlina, Bakar and Midi, 2017). Taguchi's T-Method is not fundamentally resistant to the effect of outliers (Kawada and Nagata, 2015a; Negishi *et al.*, 2017; Nishino and Suzuki, 2018), so dealing with high sample data or inconsistent small sample data will put the prediction accuracy at risk. Therefore, generalizing the procedure as early as determining the optimal unit space using a proposed method called T_{mbe} and further enhance it using a series of bootstrap approaches and extended it to increase the accuracy of the proportional coefficient and SNR weightage as parameter estimators are part of the main agenda of this research.

In the previous section, several OA arguments in MTS highlighted the need to enhance the features selection optimization by providing a new metaheuristic algorithm to substitute the OA. No attempt has been made to integrate the application of T_{mbe}, Bootstrap resampling, and Hybrid Binary ABC-PSO into Taguchi's T-Method framework so far. This attempt forms the basis of this research interest in using such approaches for greater predictive model accuracy and renders this research novel and stands at its level of notable achievement.

1.4 Research Objectives

The following research objectives have been developed to address the issues outlined in the previous section.

1. To design an approach in determining the optimum unit space for improving the accuracy of an integrated estimate output model.
2. To formulate mathematical models with optimum parameter estimators in improving the accuracy of an integrated estimate output model.
3. To develop an architecture of feature selection using Hybrid Binary Artificial Bee Colony and Particle Swarm Optimization *(Hybrid Binary ABC-PSO).*
4. To validate the accuracy and effectiveness of the algorithm.

1.5 Research Scope

The scope of this research focused on improving the accuracy of Taguchi's T-Method and how does it reduce the effect of variation towards optimality. The methods proposed in improving the integrated estimated output model were confined on; 1) LOO, 2) LOOB, and 3) 0.632 estimates, while for the features selection optimization; 1) BitABC, 2) PBPSO, and 3) Hybrid Binary ABC-PSO were bound to solve the weakness of OA's deployment on the existing Taguchi's T-Method. In this research, the performance of the proposed methods was assessed using several secondary datasets obtained from the UCI Machine Learning Repository (Lichman, 2013) and other sources that involved a different number of variables and sample data. The parameters setting applied were based on the compilation from the previous literature study.

The proposed technique and analysis carried out was driven by a linear relationship with no anticipated missing data and must involve multiple independent variables. The performance was measured mainly based on; 1) mean absolute error (MAE), 2) standard deviation (SD), and 3) SNR (dB) as the objective function. The robust element within SNR as the weightage parameter in the integrated estimated output model, and SNR (dB) as the objective function for feature optimization are the main reason to uphold the SNR element within this research.

1.6 Significance of the Study

In view of the current Taguchi's T-Method research domain, the integration of LOO, bootstrap resampling, and Hybrid Binary ABC-PSO into the Taguchi T-Method framework was the first to be introduced to date, which makes this research unique and stands at its level of significant accomplishment. The enhanced structure of unit space determination, as well as the generation of the optimal mathematical formulation by the integration between LOO and series of bootstrap algorithms, provide an easy, reliable, and systematic way to carry out the prediction analysis.

A variant of swarm intelligence algorithms such as the BitWise ABC, PBPSO, and Hybrid Binary ABC-PSO was introduced to improve the feature selection optimization within Taguchi's T-Method, replacing the OA and serve as the first introduced approach within the Taguchi's T-Method item selection domain. The PBPSO was known capable of providing better exploitation search mechanisms, while the binary ABC, such as BitWise ABC local search mechanism, helps to prevent the particle (nectar) been trapped in local optima. The respective advantage of these two algorithms leads to the development of an architecture for Hybrid Binary ABC-PSO that proved to strengthen the capability of the integration between LOO and series of the bootstrap algorithms in bringing the accuracy of the integrated estimated output model within Taguchi's T-Method to a much higher level of accuracy. By incorporating the generalization aspect into the proposed framework, the over-fitting, variability, and bias concerns were implicitly at low risk, depending on the complexity of some case study involved, as shown by the findings of the analysis in Chapter 4 of this thesis.

1.7 Thesis Outline

This thesis consists of five chapters. *Chapter 1* provides a summary of Taguchi's T-method as a modern approach to predictive research. The problem background, research questions, and objectives, as well as the significance of the research, are all well described for the reader's attention. *Chapter 2* explain the fundamental concept of Taguchi's T-Method and where it stands within the structure of the Mahalanobis Taguchi System (MTS). A brief and comprehensive overview of the utilization of Taguchi's T-Method in various engineering fields, as well as the enhancement research that has been adopted to the existing Taguchi's T-Method, is also being presented. An extended discussion on the drawback of Taguchi's T-Method concerning the limitation of the predictive model as a whole is well discussed. The discussion, therefore, leads to the reason for highlighting and introducing the proposed methods mentioned earlier.

Chapter 3 describes the overall methodology of this research, which was divided into stage I and stage II development, reflecting the objectives of improving the integrated estimated output model and optimizing the selection of features, respectively. The comprehensive framework, conceptual design, and pseudocode of the proposed algorithms are all well described in this chapter, which was the critical contribution of this research work. *Chapter 4* describes the implementation and theoretical review of the algorithms proposed. The discussions are described according to the development stage I and II and are focused on ten (10) separate case studies. The final results and findings of the specified case studies are further analyzed in terms of their overall feasibility and effectiveness against the research objectives and compared with a variety of research involving similar case studies as well as the existing Taguchi's T-Method and its variant from the literature. *Chapter 5* correlates the study results to the objectives, together with several recommendations for future research work.

REFERENCES

2-Day SBT (System Behavior Testing, MTS (Mahalanobis Taguchi System), T-Method Outline (2019) *ASI Consulting Group.* Available at: https://www.asiusa.com/what-we-do-2/workshops (Accessed: 20 January 2020).

Adler, W. and Lausen, B. (2009) 'Bootstrap estimated true and false positive rates and ROC curve', *Computational Statistics and Data Analysis*, 53(3), pp. 718–729.

Al-Salamah, M. (2015) 'Constrained binary artificial bee colony to minimize the makespan for single machine batch processing with non-identical job sizes', *Applied Soft Computing Journal*. Elsevier B.V., 29, pp. 379–385.

Alguraibawi, M., Habshah Midi, P. and Sohel Rana, L. (2015) 'Robust jackknife ridge regression to combat multicollinearity and high leverage points in multiple linear regressions', *Economic Computation and Economic Cybernetics Studies and Research*, (4), pp. 305–322.

Altun, O. and Korkmaz, T. (2014) 'Particle swarm optimization-artificial bee colony chain (PSOABCC): A hybrid metahuristic algorithm', in *Scientific Cooperations International Workshops on Electrical and Computer Engineering Subfields*, pp. 42–49.

Awang, N. A., Kamar, H. M. and Kamsah, N. (2017) 'Energy saving potential of an air-conditioner – Ice thermal storage (AC-ITS) system Akademia Baru', *Journal of advanced reseach in fluid mechanics and thermal sciences*, 31(1), pp. 1–10.

Berrar, D. and Werner, D. (2013) 'Bootstrapping', *Encyclopedia of Systems Biology*.

Billy Fang (2017) *STAT 151A (Linear Modeling), Fall 2017 [Blog]*. Available at: https://www.stat.berkeley.edu/~blfang/teaching/2017_stat151a/ (Accessed: 20 June 2018).

Booth, J. G. and Sarkar, S. (1998) 'Monte carlo approximation of bootstrap variances', *The American Statistician*, 52(4), pp. 354–357.

Borra, S. and Di Ciaccio, A. (2010) 'Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods', *Computational Statistics and Data Analysis*. Elsevier B.V., 54(12), pp. 2976–2989.

Bradley, E. and Robert, J. T. (1993a) *An introduction to the bootstrap*. 1st edn. United State of America: International Thomson Publishing.

Bradley, E. and Robert, J. T. (1993b) ‘Confidence interval based on bootstrap percentiles’, in *An introduction to the bootstrap*. 1st edn. United State of America: International Thomson Publishing, pp. 168–201.

Bradley, E. and Robert, J. T. (1993c) ‘Estimates of bias’, in *An introduction to the bootstrap*. 1st edn. United State of America: International Thomson Publishing, p. 124.

Bradley, E. and Robert, J. T. (1993d) ‘The .632 bootstrap estimator’, in *An introduction to the bootstrap*. 1st edn. United State of America: International Thomson Publishing, p. 253.

Bradly, E. and Robert, T. (1997) ‘Improvements on cross-validation : The 632 + bootstrap method’, *Journal of the American Statistical Association*, 92:438(June), pp. 548–560.

Cai, Z., Fan, J. and Yao, Q. (2000) ‘Functional-coefficient regression models for nonlinear time series’, *Journal of the American Statistical Association*, 95(451), pp. 941–956.

Cao, Y. *et al.* (2018) ‘Remote sensing of water quality based on HJ-1A HSI imagery with modified discrete binary particle swarm optimization-partial least squares (MDBPSO-PLS) in inland waters: A case in Weishan Lake’, *Ecological Informatics*. Elsevier, 44(January), pp. 21–32.

Carrascal, L. M., Galván, I. and Gordo, O. (2009) ‘Partial least squares regression as an alternative to current regression methods used in ecology’, *Oikos*, 118(5), pp. 681–690.

Cawley, G. C. and Talbot, N. L. C. (2004) ‘Fast exact leave-one-out cross-validation of sparse least-squares support vector machines’, *Neural Networks*, 17(10), pp. 1467–1475.

Cawley, G. C. and Talbot, N. L. C. (2007) ‘Preventing over-fitting during model selection via bayesian regularisation of the hyper-parameters’, *Journal of Machine Learning Research*, 8, pp. 841–861.

Chai, T. and Draxler, R. R. (2014) ‘Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature’, *Geoscientific Model Development*, 7(3), pp. 1247–1250.

Chan, Y. (2017) ‘Telecommunications- and Information Technology–Inspired Analyses: Review of an Intelligent Transportation Systems Experience’, *Transportation Research Record*, 2658(1), pp. 44–55.

Chen, Z., Li, B. and Han, B. (2018) ‘Improve regression accuracy by using an attribute weighted KNN approach’, *ICNC-FSKD 2017 - 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*. IEEE, (3), pp. 1838–1843.

Chowdhury, S. and Taguchi, S. (2016) ‘Future truck steering effort robustness’, in *Robust optimization-world best practices for developing winning vehicles*. 1st edn. John Wiley & Sons Ltd., pp. 147–172.

Chun-feng, W., Kui, L. and Pei-ping, S. (2014) ‘Hybrid Artificial Bee Colony Algorithm and Particle Swarm Search for Global Optimization’, *Hindawi-mathematical Problems in Engineering*, 2014, p. 8.

Cruz, R. M. O., Sabourin, R. and Cavalcanti, G. D. C. (2017) ‘META-DES.Oracle: Meta-learning and feature selection for dynamic ensemble selection’, *Information Fusion*. Elsevier B.V., 38, pp. 84–103.

Cudney, E. A. and Shah, P. A. (2010) ‘Predicting annual precipitation using the T-method’, in *Proceedings of the 2010 Industrial Engineering Research Conference A.*, pp. 1–6.

Cudney, E. A., Shah, P. A. and Kestle, R. (2010) ‘Predicting vehicle cost using the T-Method’, *International Journal of Product Development*, 12(3–4), pp. 311–323.

Daniels, B., Corns, S. and Cudney, E. (2012) ‘A comparison of representations for the prediction of ground-level ozone concentration’, *2012 IEEE Congress on Evolutionary Computation, CEC 2012*, pp. 10–15.

Dasneogi, P. *et al.* (2015) ‘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*, pp. 1–6.

Dasneogi, P. and Cudney, E. A. (2009) ‘Comparing the Predictive Ability of T-Method and Linear Regression Method’, in *Proceedings of the 2009 Industrial Engineering Research Conference Comparing*, pp. 2176–2182.

Dasneogi, P., Cudney, E. and Adekpedjou, A. (2009) ‘Comparing the predictive ability of t-method, linear regression method and cobb-douglas production function for warranty data’, in *30th Annual National Conference of the American Society for Engineering Management 2009, ASEM 2009*, pp. 113–119.

Dorigo, M. and Di Caro, G. (1999) ‘Ant colony optimization: A new meta-heuristic’, *Proceedings of the 1999 Congress on Evolutionary Computation, CEC 1999*, 2(February 2014), pp. 1470–1477.

Drgona, J. *et al.* (2018) 'Approximate model predictive building control via machine learning', *Applied Energy*. Elsevier, 218, pp. 199–216.

Eberhart, R. and Kennedy, J. (1995) 'A new optimizer using particle swarm theory', in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science (Nagoya, Japan), IEEE Service Center, Piscataway, NJ*, pp. 39–43.

Eck, D. J. (2018) 'Bootstrapping for multivariate linear regression models', *Statistics & Probability Letters*. North-Holland, 134, pp. 141–149.

El-Abd, M. (2011) 'A hybrid ABC-SPSO algorithm for continuous function optimization', in *IEEE SSCI 2011 - Symposium Series on Computational Intelligence - SIS 2011: 2011 IEEE Symposium on Swarm Intelligence*. IEEE, pp. 96–101.

Elisseeff, A. and Pontil, M. (2002) 'Leave-one-out error and stability of learning algorithms with applications', *NATO Science Series Sub Series III Comput. Syst. Sci.*, 190, pp. 111–130.

Elkhani, N. and Chandren Muniyandi, R. (2017) 'A Multiple Core Execution for Multiobjective Binary Particle Swarm Optimization Feature Selection Method with the Kernel P System Framework', *Journal of Optimization (Hindawi)*, p. 14.

Faizir, R. (2016) *Characterization of Bees Algorithm into the Mahalanobis-Taguchi System for classification*. Universiti Teknologi Malaysia.

Feng, S., Zhou, H. and Dong, H. (2019) 'Using deep neural network with small dataset to predict material defects', *Materials and Design*. Elsevier Ltd, 162, pp. 300–310.

Foster, C. R., Jugulum, R. and Frey, D. D. (2009) 'Evaluating an adaptive One-Factor-At-a-Time search procedure within the Mahalanobis-Taguchi System', *Internatinal Journal of Industrial and Systems Engineering*, 4(6), pp. 600–614.

Fowlkes, W. Y. and Creveling, C. M. (1995a) *Engineering Methods for Robust Design – Using Taguchi Methods in Technology and Product development*. 1st edn. United State of America: Addison-Wesley Publishing Company.

Fowlkes, W. Y. and Creveling, C. M. (1995b) 'The dynamic signal-to-noise method and metrics', in *Engineering Methods for Robust Product Design—Using Taguchi Methods in Technology and Product Development*. 1st edn. United State of America: Addison Wesley Publishing Company, p. 403.

Fowlkes, W. Y. and Creveling, C. M. (1995c) 'The dynamic signal -to-noise methods and metrics', in *Engineering method for robust product design using Taguchi Methods in technology and product development*. 1st edn. United State of America: Addison-Wesley Publishing Company, p. 104.

Freedman, D. A. (1981) 'Bootstrapping regression models', *The Annals of Statistics*, 6(9), pp. 1218–1228.

Freedman, D. A. and Peters, S. C. (1984) 'Bootstrapping a Regression Equation : Some Empirical Results', *Journal of the American Statistical Association*, 79(February 2015), pp. 37–41.

Fukunaga, K. and Hummels, D. M. (1989) 'Procedures for Nonparametric Error Estimates', *EEE Transactions on Pattern Analysis and Machine Intelligence*, 11(4), pp. 421–423.

Gharleghi, B., Nor, A. hassan shaari M. and Sarmidi, T. (2014) 'Application of the Threshold Model for Modelling and Forecasting of Exchange Rate in Selected ASEAN Countries', *Sains Malaysiana*, 43(10), pp. 1609–1622.

Gronau, Q. F. and Wagenmakers, E.-J. (2019) 'Limitations of Bayesian Leave-One-Out Cross-Validation for model selection', *Computational Brain & Behavior*, 2(1), pp. 1–11.

Guangyou, Y. (2007) 'A modified particle swarm optimizer algorithm', in *8th International Conference on Electronic Measurement and Instruments, ICEMI*. Xian, pp. 2675–2679.

Guo, W. *et al.* (2019) 'Method for Mid-Long-Term Prediction of Landslides Movements Based on Optimized Apriori Algorithm', *Applied Sciences (MDPI)*, 9(3819), pp. 1–11.

Hamadu, D. (2012) 'A Bootstrap Approach to Robust Regression', *International journal of applied science and technology*, 2(9), pp. 114–118.

Hamdi, A. *et al.* (2010) 'Bee-based algorithms : a review A quick Bee based algorithms overview', in *International Conference on Metaheuristics and Nature Inspired Computing*, pp. 9–10.

Han, S., Hwang, S. and Lee, S. (2016) 'Ensemble approach for improving prediction in kernel regression and classification', *Communications for Statistical Applications and Methods*, 23(4), pp. 355–362.

Han, Z., Li, Y. and Liang, J. (2018) 'Numerical improvement for the mechanical performance of bikes based on an Intelligent PSO-ABC Algorithm and WSN Technology', *IEEE Access*. IEEE, 6, pp. 32890–32898.

Hancer, E. *et al.* (2015) 'A binary ABC algorithm based on advanced similarity scheme for feature selection', *Applied Soft Computing Journal*. Elsevier B.V., 36, pp. 334–348.

Hao, Z. *et al.* (2016) 'Random Forest Regression based on Partial Least Squares', in *International Conference on Artificial Intelligence: Technologies and Applications (ICAITA 2016) Random*. Atlantis Press, pp. 191–196.

Harudin, N., Jamaludin, K. R., Nabil Muhtazaruddin, M., Ramlie, F. and Muhamad, W. Z. A. W. (2018) 'A feasibility study in adapting Shamos Bickel and Hodges Lehman estimator into T-Method for normalization', *IOP Conference Series: Materials Science and Engineering*, 319(1), p. 6.

Harudin, N *et al.* (2018) 'Artificial Bee Colony for Features Selection Optimization in Increasing T-Method Accuracy', *International Journal of Engineering & Technology*, 7(4.35), pp. 885–891.

Harudin, N., Jamaludin, K. R., Nabil Muhtazaruddin, M., Ramlie, F., Ismail, S., *et al.* (2018) 'Increasing T-Method Accuracy through Application of Robust M-estimator', *International Journal of Engineering & Technology*, 7(3.25)(09-special issue), pp. 44–48.

Hassanien, aboul ella and Emary, E. (2015) 'Artificial Bee Colony Optimization', in *Swarm Intelligence: Principles, Advances, and Applications*. 1st edn. United Kingdom: CRC press Taylor & Francis Group, pp. 135–170.

Hawkins, D. M. *et al.* (2003) 'Discussion', *Technometrics*, 45(1), pp. 25–29.

He, X. *et al.* (2009) 'Feature selection with discrete binary differential evolution', *2009 International Conference on Artificial Intelligence and Computational Intelligence, AICI 2009*, 4, pp. 327–330.

He, X., Sun, Y. and Li, Y. (2002) 'Stability and Generalization Olivier', *Journal of Machine Learning Research*, 2, pp. 499–526.

Hedayat, A. S., Sloane, N. J. A. and Stufken, J. (1999) *Orthogonal Arrays: Theory and Applications*. 1st edn. New York: Springer-Verlag.

Hu, J. *et al.* (2018) 'Comparative studies of error metrics in variable fidelity model uncertainty quantification', *Journal of Engineering Design*. Taylor & Francis, 29(8–9), pp. 512–538.

Huber, P. J. (1973) 'The 1972 wald memorial lecture-Robust regression : asymptotics , conjectures and Monte Carlo', *The Annals of Statistics*, 1(5), pp. 799–821.

Inoh, J. *et al.* (2012) 'Prediction Accuracies of Improved Taguchi's T Methods Compared to those of Multiple Regression Analysis', *Journal of the Japanese Society for Quality Control*, 42(2), pp. 103–115.

Jahan, A. and Edwards, K. L. (2015) 'A state-of-the-art survey on the influence of normalization techniques in ranking: Improving the materials selection process in engineering design', *Materials and Design*. Elsevier Ltd, 65, pp. 335–342.

Jahedpari, F. *et al.* (2017) 'Online Prediction via Continuous Artificial Prediction Markets', *IEEE Intelligent Systems*, 32(1), pp. 61–68.

Jain, A. (1997) 'Feature selection: evaluation, application, and small sample performance', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2), pp. 153–158.

Jain, A. K., Duin, R. P. W. and Mao, J. (2000) 'Statistical pattern recognition: A review', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), pp. 4–37.

Jia, D., Duan, X. and Khan, M. K. (2013) 'An Efficient Binary Differential Evolution with Parameter Adaptation', *International Journal of Computational Intelligence Systems*, 6(2), pp. 328–336.

Jia, D., Duan, X. and Khan, M. K. (2014) 'Binary Artificial Bee Colony optimization using bitwise operation', *Computers and Industrial Engineering*. Elsevier Ltd, 76, pp. 360–365.

Jia, Q. and Guo, Y. (2016) 'Hybridization of ABC and PSO algorithms for improved solutions of RCPSP', *Journal of the Chinese Institute of Engineers, Transactions of the Chinese Institute of Engineers,Series A/Chung-kuo Kung Ch'eng Hsuch K'an*, 39(6), pp. 727–734.

Kacker, R. N., Lagergren, E. S. and Filliben, J. J. (1991) 'Taguchi's orthogonal arrays are classical designs of experiments', *Journal of Research of the National Institute of Standards and Technology*, 96(5), pp. 577–591.

Kambekar, A. R. and Deo, M. C. (2012) 'Wave prediction using genetic programming and model trees', *Journal of Coastal Research*, 28(1), pp. 43–50.

Karaboga, D. (2005) *An Idea Based on Honey Bee Swarm for Numerical Optimization, Technical Report - TR06*.

Karaboga, D. *et al.* (2014) 'A comprehensive survey: Artificial bee colony (ABC) algorithm and applications', *Artificial Intelligence Review*, 42(1), pp. 21–57.

Karaboga, D. and Akay, B. (2009) 'A comparative study of Artificial Bee Colony algorithm', *Applied Mathematics and Computation*. Elsevier Inc., 214(1), pp. 108–132.

Karaboga, D. and Basturk, B. (2007) 'A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm', *Journal of Global Optimization*, 39(3), pp. 459–471.

Karoui, N. El and Purdom, E. (2018) 'Can we trust the bootstrap in high-dimensions? The case of linear models', *Journal of Machine Learning Research*, 19, pp. 1–66.

Kashan, M. H., Nahavandi, N. and Kashan, A. H. (2012) 'DisABC: A new artificial bee colony algorithm for binary optimization', *Applied Soft Computing Journal*. Elsevier B.V., 12(1), pp. 342–352.

Kasule, A. and Ayan, K. (2019) 'Forecasting Uganda's Net Electricity Consumption Using a Hybrid PSO-ABC Algorithm', *Arabian Journal for Science and Engineering*, 44(4), pp. 3021–3031.

Kawada, H. and Nagata, Y. (2015a) 'An application of a generalized inverse regression estimator to Taguchi ' s T-Method', *Journal of the Japanese Society for Quality Control*, 1, pp. 12–21.

Kawada, H. and Nagata, Y. (2015b) 'Studies on the Item Selection in Taguchi's T-Method', *Journal of The Japanese Society for Quality Control*, 45(2), pp. 179–193.

Kennedy, J. and Eberhart, R. C. (1997) 'A discrete binary version of the particle swarm algorithm', *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, 5, pp. 4104–4108.

Khalaf, A. *et al.* (2017) 'Students' Success Prediction based on Bayes Algorithms', *International Journal of Computer Applications*, 178(7), pp. 6–12.

Khuat, T. T. and Le, M. H. (2018) 'A Novel Hybrid ABC-PSO Algorithm for Effort Estimation of Software Projects Using Agile Methodologies', *Journal of Intelligent Systems*, 27(3), pp. 489–506.

Kim, S.-K. and Oh, J. (2018) 'Information science techniques for investigating research areas: a case study in telecommunications policy', *The Journal of Supercomputing*, 74(12), pp. 6691–6718.

Kiran, M. S. (2015) ‘The continuous artificial bee colony algorithm for binary optimization’, *Applied Soft Computing Journal*. Elsevier B.V., 33, pp. 15–23.

Kisielinska, J. (2013) ‘The exact bootstrap method shown on the example of the mean and variance estimation’, *Computational Statistics*, 28(3), pp. 1061–1077.

Kohavi, R. (1995) ‘Cross validation number.pdf’, in *International Joint Conference on Artificial Intelligence*, p. 7.

Kovács, G. (2019) ‘An empirical comparison and evaluation of minority oversampling techniques on a large number of imbalanced datasets’, *Applied Soft Computing Journal*. Elsevier B.V., 83, p. 105662.

Koyuncu, H. and Ceylan, R. (2019) ‘A PSO based approach: Scout particle swarm algorithm for continuous global optimization problems’, *Journal of Computational Design and Engineering*. Society for Computational Design and Engineering, 6(2), pp. 129–142.

Kramar, V. A. *et al.* (2018) ‘Application of support vector machine for prediction and classification’, *Journal of Physics: Conference Series*, 1015(3).

Krutchkoff, R. G. (1969) ‘Classical and Inverse Regression Methods of Calibration in Extrapolation’, *Technometrics*, 11(3), pp. 605–608.

Kumar, N. *et al.* (2019) ‘Hyperspectral tissue image segmentation using semi-supervised NMF and hierarchical clustering’, *IEEE Transactions on Medical Imaging*. IEEE, 38(5), pp. 1304–1313.

Kuo, C. Y. *et al.* (2018) ‘Comparison of models for the prediction of medical costs of spinal fusion in Taiwan diagnosis-related groups by machine learning algorithms’, *Healthcare Informatics Research*, 24(1), pp. 29–37.

Lau, K., López, R. and Oñate, E. (2009) ‘A Neural Networks Approach to Aerofoil Noise Prediction’.

Li, D. C. *et al.* (2019) ‘Using virtual samples to improve learning performance for small datasets with multimodal distributions’, *Soft Computing*. Springer Berlin Heidelberg, 23(22), pp. 11883–11900.

Li, Y. *et al.* (2016) ‘Intercrossed access controls for secure financial services on multimedia big data in cloud systems’, *ACM Transactions on Multimedia Computing, Communications and Applications*, 12(4s), pp. 1–18.

Li, Zhiyong *et al.* (2015) ‘PS-ABC: A hybrid algorithm based on particle swarm and artificial bee colony for high-dimensional optimization problems’, *Expert Systems with Applications*. Elsevier Ltd., 42(22), pp. 8881–8895.

Lichman, M. (2013) *UCI Machine Learning Repository*, *Irvine, CA: University of California, School of Information and Computer Science.* Available at: http://archive.ics.uci.edu/ml (Accessed: 23 February 2018).

Liu, K. Y. *et al.* (2013) 'Evaluation and amelioration of computer-aided diagnosis with artificial neural networks utilizing small-sized sample sets', *Biomedical Signal Processing and Control*. Elsevier Ltd, 8(3), pp. 255–262.

Liu, T., Zhang, L. and Zhang, J. (2013) 'Study of binary artificial bee colony algorithm based on particle swarm optimization', *Journal of Computational Information Systems*, 9(16), pp. 6459–6466.

Liu, Z. *et al.* (2019) 'A novel support vector regression algorithm incorporated with prior knowledge and error compensation for small datasets', *Neural Computing and Applications*. Springer London, 31(9), pp. 4849–4864.

Marlan, Z. M. *et al.* (2019) 'Determination of Optimal Unit Space Data for Taguchi ' s T-method based on Homogeneity of Output', *Open International Journal of Informatics (OIJI)*, 7(special issue), pp. 167–179.

Mazlina, N., Bakar, A. and Midi, H. (2017) 'Robust Weights of Generalized M-Estimator for Panel Data', in *AIP Conference Proceedings*, p. 6.

McCullough, B. D. and Vinod, H. (1993) 'Implementing the single bootstrap: Some computational considerations', *Computational Economics*, 6(1), pp. 1–15.

Mehmed, K. (2011) *Data Mining:Concepts , Models , Methods , and Algorithms*. 2nd edn. New Jersey: John Wiley & Sons, Inc.

Mehta, K. (2019) 'Abalone Age Prediction Problem: A Review', *International Journal of Computer Applications*, 178(50), pp. 43–49.

Menhas, M. I. *et al.* (2011) 'Coordinated controller tuning of a boiler turbine unit with new binary particle swarm optimization algorithm', *International Journal of Automation and Computing*, 8(2), pp. 185–192.

Menhas, M. I. *et al.* (2012) 'Comparative performance analysis of various binary coded PSO algorithms in multivariable PID controller design', *Expert Systems with Applications*. Pergamon, 39(4), pp. 4390–4401.

Michael, C. (2008) *Bootstrap Methods: A Guide for Practitioners and Researchers by CHERNICK, M. R.* 2nd edn, *Biometrics*. 2nd edn.

Michael R. Chernick (2008) *Bootstrap methods. A Guide for practitioners and researchers*. 2nd edn. United State of America: John Wiley & Sons Ltd.

Miwa, T. (1985) ‘Comparison among point estimators in linear calibration in terms of mean squared error’, *Japanese Journal of Applied Statistics*, 14(2), pp. 83–93.

Moayedi, H., Mu’azu, M. A. and Foong, L. K. (2020) ‘Novel swarm-based approach for predicting the cooling load of residential buildings based on social behavior of elephant herds’, *Energy and Buildings*. Elsevier B.V., 206.

Mohd Noor, C. W. *et al.* (2015) ‘Application of artificial neural network for prediction of marine diesel engine performance’, *IOP Conference Series: Materials Science and Engineering*, 100(1).

Molinaro, A. M., Simon, R. and Pfeiffer, R. M. (2005) ‘Prediction error estimation: A comparison of resampling methods’, *Bioinformatics*, 21(15), pp. 3301–3307.

Muhamad, W. Z. A. W. (2019) *Hybrid Mahalanobis Taguchi System with Randomized Algorithms For simultaneous feature selection and binary classification*. Universiti Teknologi Malaysia.

Muhtazaruddin, M. N. *et al.* (2013) ‘Optimal capacitor placement and sizing via artificial bee colony’, *International Journal of Smart Grid and Clean Energy*, 3(2), pp. 200–206.

Muhtazaruddin, M. N. Bin and Fujita, G. (2013) ‘Distribution network power loss by using Artificial Bee Colony’, in *Proceedings of the Universities Power Engineering Conference*, p. 5.

Muhtazaruddin, M. N. Bin, Jamian, J. J. Bin and Fujita, G. (2014) ‘Determination of optimal output power and location for multiple distributed generation sources simultaneously by using artificial bee colony’, *IEEJ Transactions on Electrical and Electronic Engineering*, 9(4), pp. 351–359.

Nakao, Y. and Nagata, Y. (2018) ‘Analysis of data including missing values in the Taguchi’s T method’, *Total Quality Science*, 4(2), pp. 53–64.

Negishi, S. *et al.* (2017) ‘Daily Peak Load Forecasting by Taguchi’s T Method’, *Electrical Engineering in Japan (English translation of Denki Gakkai Ronbunshi)*, 201(1), pp. 57–65.

Nishino, K. and Suzuki, A. (2018) ‘Taguchi’s T-method using median-median line for small sample with outliers’, *Electronics and Communications in Japan*, 138(7), pp. 598–604.

Ozger, Z. B., Bolat, B. and Diri, B. (2017) ‘Acomparative study on binary artificial bee colony optimization methods for feature selection’, *Focus on Swarm Intelligence Research and Applications*, (2), pp. 109–128.

Pal, A. and Maiti, J. (2010) ‘Development of a hybrid methodology for dimensionality reduction in Mahalanobis–Taguchi system using Mahalanobis distance and binary particle swarm optimization’, *Expert Systems with Applications*, 37(2), pp. 1286–1293.

Pasini, A. (2015) ‘Artificial neural networks for small dataset analysis’, *Journal of Thoracic Disease*, 7(5), pp. 953–960.

Peng, C. Y. J., Lee, K. L. and Ingersoll, G. M. (2002) ‘An introduction to logistic regression analysis and reporting’, *Journal of Educational Research*, 96(1), pp. 3–14.

Penny, W. *et al.* (2004) ‘Sources of Variation and Bias in Studies of Diagnostic Accuracy’, *annals of internal medicine*, 140(3), pp. 189–203.

Peprah, M. S. and Mensah, I. O. (2017) ‘Performance evaluation of the Ordinary Least Square (OLS) and Total Least Square (TLS) in adjusting field data: an empirical study on a DGPS data’, *South African Journal of Geomatics*, 6(1), p. 73.

Peter J, B. and Freedman, D. A. (1981) ‘Some asymptotic theory for the bootstrap’, *The Annals of Statistics*, 9(6), pp. 1196–1217.

Phadke, M. . (1989) *Quality Engineering Using Robust Design*. 1st edn. United State of America: PTR Prentice-Hall, Inc.

Pham, D. T. *et al.* (2006) ‘The Bees Algorithm - A Novel Tool for Complex Optimisation Problems’, in *Intelligent Production Machines and Systems - 2nd I*PROMS Virtual International Conference 3-14 July 2006*, pp. 454–459.

Pons, E. *et al.* (2019) ‘Impact of guidelines for the management of minor head injury on the utilization and diagnostic yield of CT over two decades, using natural language processing in a large dataset’, *European Radiology*. European Radiology, 29(5), pp. 2632–2640.

Qing, X. and Niu, Y. (2018) ‘Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM’, *Energy*. Pergamon, 148, pp. 461–468.

Ramlie, F. *et al.* (2016) ‘Optimal feature selection of Taguchi character recognition in the Mahalanobis-Taguchi system using bees algorithm’, *Global Journal of Pure and Applied Mathematics*, 12(3), pp. 2651–2671.

Rao, C. R. and Wu, Y. (2005) ‘Linear model selection by cross-validation’, *Journal of Statistical Planning and Inference*, 128(1), pp. 231–240.

Rho, Y. and Shao, X. (2015) ‘Inference for Time Series Regression Models With Weakly Dependent and Heteroscedastic Errors’, *Journal of Business & Economic Statistics*. Taylor & Francis, 33(3), pp. 444–457.

Roberts, M. J. and Russo, R. (1999) ‘Average, measures of dispersal and the t-test’, in *A student’s guide to Analysis of Variance*, p. 4,13.

Rousseeuw, peter j. and Leroy, A. M. (2003) ‘Simple regression’, in *Robust Regression and Outliers Detection*, p. 70.

Rushing, C. *et al.* (2015) ‘A leave-one-out cross-validation SAS macro for the identification of markers associated with survival’, *Computers in Biology and Medicine*, 57, pp. 123–129.

Saka, M. ., Dogan, E. and Ibrahim Aydogdu (2013) ‘Analysis of Swarm Intelligence-Based Algorithms for Constraint Optimization’, in *Swarm Intelligence and Bio-Inspired Computation Theory and Applications*. 1st edn. Elsevier Inc., pp. 25–48.

Samart, K., Jansakul, N. and Chongcheawchamnan, M. (2018) ‘Exact bootstrap confidence intervals for regression coefficients in small samples’, *Communications in Statistics: Simulation and Computation*, 47(10), pp. 2953–2959.

Santana, C. J. *et al.* (2019) ‘A novel binary artificial bee colony algorithm’, *Future Generation Computer Systems*. Elsevier B.V., 98, pp. 180–196.

Sapra, R. L. (2014) ‘Using R-squared with caution’, *Current Medicine Research and Practice*. Elsevier Ltd, 4(3), pp. 130–134.

Satoshi, K., Yasushi, N. and Nagata, Y. (2018) ‘Application of Taguchi ’ s T-method to Circumferential Data Considering’, in *16th ANQ Congress Future Quality Leadership in IT-Age*, pp. 238–249.

Schmidt, A. F. and Finan, C. (2017) ‘Linear regression and the normality assumption’, *Journal of Clinical Epidemiology*. Pergamon.

Schulerud, H. and Albregtsen, F. (2004) ‘Many are called, but few are chosen. Feature selection and error estimation in high dimensional spaces’, *Computer Methods and Programs in Biomedicine*, 73(2), pp. 91–99.

Seber, G. A. F. and Lee, A. J. (2003) *Linear regression Analysis*. 2nd edn. New Jersey: Wiley series in probability and statistics.

Shao, Z., Er, M. J. and Wang, N. (2015) ‘An Efficient Leave-One-Out Cross-Validation-Based Minimal User Intervention’, pp. 1–13.

Shin, T. *et al.* (2003) 'Discussion', *Technometrics*, 45(1), pp. 16–21. Available at: http://www.tandfonline.com/doi/abs/10.1198/004017002188618635.

Smith, G. C. S. *et al.* (2014) 'Correcting for optimistic prediction in small data sets', *American Journal of Epidemiology*, 180(3), pp. 318–324.

Steyerberg E.W. (2019) 'Overfitting and Optimism in Prediction Models', in *Clinical Prediction Models. Statistics for Biology and Health. Springer, Cham*, pp. 95–112.

Su, C. T. and Hsiao, Y. H. (2007) 'An evaluation of the robustness of MTS for imbalanced data', *IEEE Transactions on Knowledge and Data Engineering*, 19(10), pp. 1321–1332.

Suguru, S. and Yasushi, N. (2018) 'A Study of Multivariate Taguchi ' s T Method', in *16th ANQ Congress<<Future Quality Leadership in IT-Age>>*, pp. 250–262.

Supreetha, B. S., Prabhakara Nayak, K. and Narayan Shenoy, K. (2019) 'Hybrid artificial intelligence based abc-pso system for ground water level forecasting in Udupi region', *Journal of Engineering Science and Technology*, 14(2), pp. 797–809.

Suzuki, A. (2015) 'Inverse-model-based cuffless blood pressure estimation using a single photoplethysmography sensor', *Journal of Engineering in Medicine*, 229(7), pp. 499–505.

Suzuki, A. and Sugimoto, K. (2010) 'Robust estimation of heat capacity in cooking devices', in *2010 IEEE International Conference on Control Applications*, pp. 1374–1378.

Taguchi, G. *et al.* (2005) *Taguchi's Quality Engineering Handbook*. 1st edn. New Jersey: John Wiley & Sons Ltd.

Taguchi, G., Chowdhury, S. and Taguchi, S. (2000) *Robust Engineering – Learn how to boost quality while reducing costs and time to market*. 1st edn. New York: McGraw-Hill.

Taguchi, G., Chowdhury, S. and Wu, Y. (2001) *The Mahalanobis-Taguchi System*. 1st edn. New York: McGraw-Hill.

Taguchi, G. and Jugulum, R. (2002) *The Mahalanobis-Taguchi Strategy : A Patter Technology System*. 1st edn. United State of America: John Wiley & Sons, Inc.

Tang, K. *et al.* (2019) 'Superpixels for large dataset subspace clustering', *Neural Computing and Applications*. Springer London, 31(12), pp. 8727–8736.

Tao, F. *et al.* (2018) 'Digital twin-driven product design, manufacturing and service with big data', *International Journal of Advanced Manufacturing Technology*. The International Journal of Advanced Manufacturing Technology, 94(9–12), pp. 3563–3576.

Tarr, G., Weber, N. C. and Muller, S. (2015) 'The difference of symmetric quantiles under long range dependence', *Statistics and Probability Letters*. Elsevier B.V., 98, pp. 144–150.

Teshima, S., Hasegawa, Y. and Tatebayashi, K. (2012a) 'Examples of actual application', in *Quality recognition and prediction: smarter pattern technology with the Mahalanobis-Taguchi System*. 1st edn. New York: Momentum Press, pp. 145–156.

Teshima, S., Hasegawa, Y. and Tatebayashi, K. (2012b) 'Merits of the MT system and its computation methods', in *Quality recognition and prediction: smarter pattern technology with the Mahalanobis-Taguchi System*. 1st edn. New York: Momentum Press, pp. 15–54.

Teshima, S., Hasegawa, Y. and Tatebayashi, K. (2012c) 'Pattern Recognition and the MT System', in *Quality recognition and prediction: smarter pattern technology with the Mahalanobis-Taguchi System*. 1st edn. New York: Momentum Press, pp. 1–13.

Teshima, S., Hasegawa, Y. and Tatebayashi, K. (2012d) *Quality Recognition and Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System*. 1st edn. New York: Momentum Press.

Teshima, S., Hasegawa, Y. and Tatebayashi, K. (2012e) 'T-Method application procedures and Key points', in *Quality recognition and prediction: smarter pattern technology with the Mahalanobis-Taguchi System*. 1st edn. New York: Momentum Press, pp. 87–104.

Tsanas, A. and Xifara, A. (2012) 'Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools', *Energy and Buildings*. Elsevier B.V., 49, pp. 560–567.

Tsui, K., Sukchotrat, T. and Chen, V. C. P. (2009) 'A comparison study and discussion of the Mahalanobis-Taguchi System Seoung Bum Kim', *Int. J. Industrial and Systems Engineering*, 4(6), pp. 631–644.

Turabieh, H., Mafarja, M. and Li, X. (2019) 'Iterated feature selection algorithms with layered recurrent neural network for software fault prediction', *Expert Systems with Applications*. Elsevier Ltd, 122, pp. 27–42.

Vicente, T. F. Y., Hoai, M. and Samaras, D. (2018) 'Leave-One-Out Kernel Optimization for Shadow Detection and Removal', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3), pp. 682–695.

Vieira, S. M. *et al.* (2013) 'Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients', *Applied Soft Computing*. Elsevier, 13(8), pp. 3494–3504.

Wang, L. *et al.* (2008) 'A novel probability binary particle swarm optimization algorithm and its application', *Journal of Software*, 3(9), pp. 28–35.

Waqar, M. *et al.* (2018) 'Prediction of Stock Market by Principal Component Analysis', in *Proceedings - 13th International Conference on Computational Intelligence and Security, CIS 2017*, pp. 599–602.

Whiting, P. F. *et al.* (2013) 'A systematic review classifies sources of bias and variation in diagnostic test accuracy studies', *Journal of Clinical Epidemiology*, 66(10), pp. 1093–1104.

Wilcox, R. R. (2005) *Introduction to Robust Estimation and Hypothesis Testing*.

Willmott, C. J. and Matsuura, K. (2005) 'Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance', *Climate Research*, 30(1), pp. 79–82.

Woodall, W., Koudelik, R, *et al.* (2003) 'A review and analysis of the Mahalanobis-Taguchi system', *Technometrics*, 45(1), pp. 1–15.

Woodall, W., Koudelik, Rachelle, *et al.* (2003) 'Response', *Technometrics*, 45(1), pp. 29–30.

Wu, X., Zhang, Q. and Hu, J. (2016) 'QSAR study of the acute toxicity to fathead minnow based on a large dataset', *SAR and QSAR in Environmental Research*, 27(2), pp. 147–164.

Xu, Y. and Goodacre, R. (2018) 'On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning', *Journal of Analysis and Testing*. Springer Singapore, 2(3), pp. 249–262.

Yan, X. and Su, X. G. (2009) *Linear Regression Analysis*. 1st edn. Singapore: World Scientific Publishing Co. Pte. Ltd.

Yang, X.-S. and Deb, S. (2009) ‘Cuckoo Search via Levy Flights’, in *Proc. of World Congress on Nature & Biologically Inspired Computing, IEEE publication*, pp. 210–214.

Yang, X.-S. and Slowik, A. (2010) ‘Firefly Algorithm’, in *Engineering Optimization: An Introduction with Metaheuristic Applications*, pp. 221–230.

Yang, X. and Deb, S. (2010) ‘A new methaheuristic Bat-Inspired algorithm’, *Nature Inspired Cooperative Strategies for Optimization*, 284, pp. 65–74.

Yeh, I. C. (1998) ‘Modeling of strength of high-performance concrete using artificial neural networks’, *Cement and Concrete Research*, 28(12), pp. 1797–1808.

Yoshimura, A. and Nagata, Y. (2015) ‘An Application of Directional Data Analysis to Taguchi’s T Method’, in *The ANQ Congress 2015 Taipei*.

Zabidi, A. *et al.* (2018) ‘Comparison between Binary Particles Swarm Optimization (BPSO) and Binary Artificial Bee Colony (BABC) for nonlinear autoregressive model structure selection of chaotic data’, *Journal of Fundamental and Applied Sciences*, 9(3S), p. 730.

Zhang, Z. and Gui, W. (2019) ‘Statistical inference of reliability of Generalized Rayleigh distribution under progressively type-II censoring’, *Journal of Computational and Applied Mathematics*. Elsevier B.V., 361, pp. 295–312.

Zhao, P., Sun, X. and Chen, P. (2018) ‘An Improved Binary Bee Colony Algorithm for Satellite Resource Scheduling Method’, *Lecture Notes in Electrical Engineering*, 473(10), pp. 181–188.

Zhu, X. and Gao, Z. (2018) ‘An efficient gradient-based model selection algorithm for multi-output least-squares support vector regression machines’, *Pattern Recognition Letters*. Elsevier B.V., 111, pp. 16–22.

Zuo, J. *et al.* (2014) ‘Time series prediction based on Gene Expression Programming’, in *International Conference on Web-Age Information Management WAIM*, pp. 55–64.

LIST OF PUBLICATIONS

International Journal Papers

1. **Harudin, N.**, Jamaludin, K. R., Ramlie, F., Muhtazaruddin, M. N., Munira, C., Razali, C., & Muhamad, W. Z. A. W. (2020). Binary Particle Swarm Optimization for Variables Selection Optimization in Taguchi ' s T-Method. *Matematika: Mjiam*, *36*(1), 69–84. **(Indexed by ISI-**ESCI**)**

2. **Harudin, N.**, Jamaludin, K. R., Nabil Muhtazaruddin, M., Ramlie, F., Ismail, S., Muhamad, W. Z. A. W., & Jaafar, N. (2018). Increasing T-Method Accuracy through Application of Robust M-estimator. *International Journal of Engineering & Technology*, *7(3.25)*(09-special issue), 44–48. **(Indexed by Scopus)**

3. Razali, C. M. C., Hussein, S. F. M., **Harudin, N.** , & Abdullah, S. S. (2018). Estimation of building energy efficiency performance using Radial Basis Function Neural Network. *International Journal of Engineering and Technology(UAE)*, *7*(4), 755–759. **(Indexed by Scopus)**

4. **Harudin, N**., Jamaludin, K. R., Muhtazaruddin, M. N., Ramlie, F., Muhamad, W. Z. A. W., Muhtazaruddin, M. N., … Muhamad, W. Z. A. W. (2018). Artificial Bee Colony for Features Selection Optimization in Increasing T-Method Accuracy. *International Journal of Engineering & Technology*, *7*(4.35), 885–891. **(Discontinue Scopus at time of publication)**

5. **Harudin, N.**, KR, J., F, R., Muhtazaruddin, M. N., Marlan, Z., Muhamad, W., & Jaafar, N. (2019). An Overview of Taguchi' S T-Method as A Prediction Tool for Multivariate Analysis. *Open International Journal of Informatics (OIJI)*, *7*(special issue), 158–166.

6. Marlan, Z. M., Jamaludin, K. R., Ramlie, F., **Harudin, N.**, & Jaafar, N. N. (2019). Determination of Optimal Unit Space Data for Taguchi ' s T-method based on Homogeneity of Output. *Open International Journal of Informatics (OIJI)*, *7*(special issue), 167–179.

International Conference Proceedings

1. **Harudin, N.**, Jamaludin, K. R., Nabil Muhtazaruddin, M., Ramlie, F., & Muhamad, W. Z. A. W. (2018). A feasibility study in adapting Shamos Bickel and Hodges Lehman estimator into T-Method for normalization. *IOP Conference Series: Materials Science and Engineering*, *319*(1), 6. **(Indexed by SCOPUS)**

2. Muhamad, W. Z. A. W., Jamaludin, K. R., Ramlie, F., **Harudin, N.**, & Jaafar, N. N. (2018). Criteria selection for an mba programme based on the mahalanobis taguchi system and the kanri distance calculator. In *IEEE Student Conference on Research and Development: Inspiring Technology for Humanity, SCOReD 2017 - Proceedings*. **(Indexed by SCOPUS)**

3. Wan Zuki Azman, W. M., Khairur Rijal, Jamaludin Mohd Nabil, M., Zainor Ridzuan, Yahya Faizir, R., & **Harudin, N.** (2018). Optimal variable screening in automotive crankshaft remanufacturing process using statistical pattern recognition approach in the Mahalanobis-Taguchi system. *AIP Conference Proceedings*, *2013*(1), 6. **(Indexed by SCOPUS)**

International Conference Presentations

1. **Presenter. Harudin, N.**, Jamaludin, K. R., Ramlie, F., Muhtazaruddin, M. N., Munira, C., Razali, C., & Muhamad, W. Z. A. W. (2020). Binary Particle Swarm Optimization for Variables Selection Optimization in Taguchi ' s T-Method. 4th International Conference on Robust Quality Engineering (ICRQE 2018). Kuala Lumpur, Malaysia. 4-5 August 2018.

2. **Presenter. Harudin, N.**, Jamaludin, K. R., Nabil Muhtazaruddin, M., Ramlie, F., Ismail, S., Muhamad, W. Z. A. W., & Jaafar, N. (2017). Increasing T-Method Accuracy through Application of Robust M-estimator. 3rd ASIA International Conference (AIC 2017). Kuala Lumpur, Malaysia. 9-10 December 2017. **(Best Poster Presentation Award)**