JOURNAL OF MODERN MANUFACTURING SYSTEMS AND TECHNOLOGY (JMMST)
e-ISSN: 2636-9575
VOL. 7, ISSUE 2, 1 – 8
DOI: https://doi.org/10.15282/jmmst.v7i2.9582

ORIGINAL ARTICLE

Application of Mahalanobis-Taguchi system in Rainfall Distribution

N.H. Aris[1], M.Y. Abu[1*], M.A.M. Jamil[1], S.N.A.M. Zaini[1], N.S. Pinueh[1], W.Z.A.W. Muhamad[2], F. Ramlie[3], N. Harudin[4], and E. Sari[5]

[1]Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, 26600 Pekan, Pahang, Malaysia
[2]Institute of Engineering Mathematics, Universiti Malaysia Perlis, Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia
[3]Razak Faculty of Technology and Informatics, Department of Mechanical Engineering, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia
[4]Universiti Tenaga Nasional, 43000 Kajang, Selangor, Malaysia
[5]Universitas Trisakti, Faculty of Industrial Technology, Department of Industrial Engineering, 11440, Kyai Tapa No 1, West Jakarta, Indonesia

ABSTRACT – The rainfall time series is often nonlinear and multi-time scale because of hydrology, meteorological, and human activity. Weather stations gather information on a diverse set of parameters on order to monitor and analyses patterns of rainfall. Nevertheless, not all parameters are created equal in terms of its significance or effectiveness in carrying out classification and optimization actions. The objective is to classify rainfall occurrences by the RT method and optimize the parameter selection process by the T method using Mahalanobis-Taguchi system (MTS). The data was collected using Vantage Pro2 weather station at UMPSA Gambang campus and it consists of 16 various parameters. As a results, RT method can classify the data samples in terms of MD for the months of June, October and December by utilizing the, while simultaneously the number of parameters is reduced to only those that substantially contribute to the classification. This brings the total number of parameters decrease from 16 to 8 when compared to the T method. So, this research methods offer a simplified and effective way for analyzing rainfall patterns and optimizing the data gathering processes at weather stations.

ARTICLE HISTORY
Received: 28th July 2023
Revised: 29th Sept 2023
Accepted:30th Sept 2023

KEYWORDS
MD
MTS
Classification
Optimization
Rainfall distribution

INTRODUCTION

As global warming grows, numerous severe weather occurrences, including excessive rainfall, become increasingly common. The effects of rainfall on agriculture, hydrology, communications, and other fields are also obvious. The weight of the foregoing demands compel the researchers to put in additional effort to overcome the limitations of the current rainfall monitoring system, which has problems such as a lack of geographical representativeness and spatiotemporal precision [1]. The rainfall time series frequently demonstrates significant nonlinear and multi-time scale features. This is because the rainfall time series is influenced by a wide variety of elements, including hydrology, meteorology, and human activities. Multiple timescale rainfall systems occur when hydrological series change over time, resulting in a multi-level structure in the time domain [2].

The tropical location in Malaysia leads to a warm and humid climate throughout the year. The characteristics of the climate differ from region to region within Malaysia. There are two different monsoon seasons: southwest monsoon (SWM) and northeast monsoon (NEM). The (SWM) takes place from April to September while the NEM takes place from October to March [3]. The NEM provides heavy rain to Peninsular Malaysia's east coast and also in Sabah and Sarawak. While the SWM delivers rain to Peninsular Malaysia's west coast. The Peninsular Malaysian Titiwangsa mountain range controls rainfall distribution, reducing winter monsoons near the west coast [4]. Malaysia has six hours of direct sunshine a day and cloud cover for the afternoon or evening [3].

In order to conduct reliable analysis and predictions, it is essential to have a solid understanding of the major characteristics that have contributed to the classification of rainfall. According to a published paper, El Niño and La Niña are linked to severe precipitation in Malaysia. Researchers discovered that these occurrences have a major impact on the features of intense rainfall across the country [4]. Malaysian rainfall distribution and variability are also affected by long-term persistence and regional sea surface temperature anomalies [5].

The objective of this research is to classify and optimize the gathered data in order to locate the parameters that are significant by using Mahalanobis-Taguchi system (MTS). MTS is a strong approach for multivariate data analysis and quantitative decision-making that uses Mahalanobis distance (MD) measure and Taguchi's orthogonal array (OA) design [6]. MTS can handle non-linear patterns and analyse correlations between independent variables or components, making

*CORRESPONDING AUTHOR | M.Y. Abu | myazid@umpsa.edu.my

it better than linear discriminant analysis and logistic regression [7]. MTS is used in engineering, manufacturing, medical, and other areas [8].

LITERATURE REVIEW

Mahalanobis-Taguchi system (MTS) is a pattern recognition approach in multidimensional systems created by Taguchi and Jugulum that does not make any assumptions about statistical distribution. MTS is made up of the Mahalanobis distance (MD) and Taguchi's robuts engineering [9]. MD is used to develop a multidimensional measure scale and create the Mahalanobis space (MS), which is utilised as scale reference point with a set of normal data. MTS creates the multidimensional measure scale using single class samples rather than the whole training data set. So, it can be used to overcome class imbalance issues. Taguchi's robust engineering is used to choose useful characteristics gained through the usage of orthogonal arrays (OA) and signal-to-noise ratios (SNR) [10].

MD is a statistical measure designed by PC Mahalanobis, an Indian statistician, that considers the contribution of various variable to the distance value based on the variability of each variable. MD has several uses in multivariate statistics. It varies from Euclidean distance in that it considers the correlation between variables. Figure 1 shows a comparison of two distances, with the Euclidean and MD of points on circles and ellipses approximately 1 and 2 unit distant from the center of data. The apparent difference is due to the fact that the MD adjust for data correlation [11].

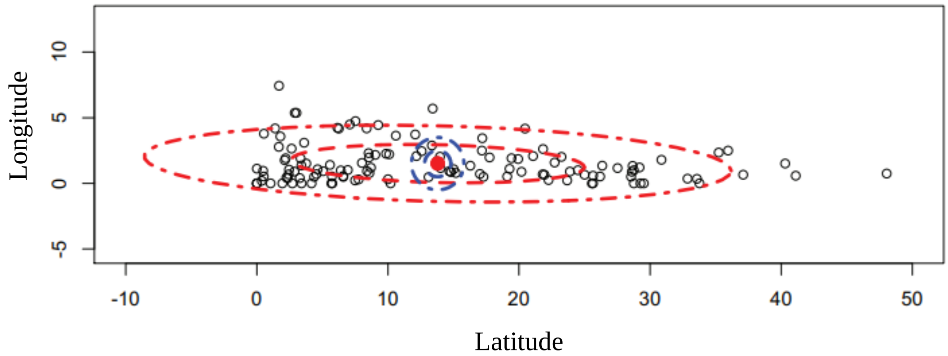


Figure 1. Schematic comparison of Mahalanobis (ellipse) and Euclidean (circle) distances

Additionally, MTS shows potential as a binary classification technique for unbalanced data. Its performance has been compared to other algorithms like Support Vector Machine (SVM) and Naïve Bayes. It has been found to be outstanding performance, especially when dealing with highly ratio imbalanced data sets [12]. Multidimensional data may be classified and predicted by the MTS using the measurement of MD [13]. Because of this, it is a useful tool for making decisions that need accurate and reliable predictions.

In general, the MTS provides a methodical and practical approach to quality improvement and decision-making, making it is useful in many sectors. MTS is widely used in several sectors because it can analyse complicated data, identify patterns, and optimize procedures [8].

RESEARCH METHODOLOGY

This research focused the rainfall distribution which is conducted in Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) of Gambang campus. The data was collected using Vantage Pro2 weather station that was set up at the field of Kolej Kediaman 2 (KK2) in UMPSA Gambang campus. The KK2 field is chosen because of its stability and compromise in achieving the optimal criterion for installing a rain gauge. The optimal site location should be free of obstacle, buildings, and steep slopes that will interfere with the weather station's data collecting [14]. There are 16 rainfall parameters that are emphasized for the MTS implementation, namely temperature outside, high temperature, low temperature, outside humidity, dew point, heat index, barometric pressure, rain, rain rate, cool degree-day, inside temperature, inside humidity, inside dew, inside heat index, inside EMC and inside air density. Thus, the data classification and optimisation parameter may be analysed. The rainfall distribution of 16 parameters for 2021 are used to determine the significant parameter. The 2021 data selection that are chosen is divided into three categories of monsoon phenomena that are southwest (June), northeast (December), and transition of southwest to northeast (October).

This research utilized two types of methods which is RT method and T method. The RT method is a detection method that can categories data into two different groups, those that are inside of unit space and those that are outside of unit space. RT method works effectively when a signal's real value is uncertain but its category is unambiguous and there are many unit spaces. On the other hand, T method establishes a unit space in which the outputs are moderate and equally distributed. The multivariate T method predicts and estimates output values like multiple regression analysis [15].

Firstly, classify the gathered data into two categories space by using RT method. Unlike T method, the RT method works when the real value of a measure output is unknown but its category is clear and various unit spaces are accessible. The average value for each parameter in normal group unit space can be found and may be written as equation (1).

$$\bar{x}_j = \frac{1}{n}(x_{1j} + x_{2j} + \cdots + x_{nj}) \qquad (j = 1, 2, \ldots, \mathrm{k}) \tag{1}$$

Next, finding the sensitivity, β for the initial samples of unit space as equation (2) required the linear formula, L and effective divider, r from equations (3) and (4).

Sensitivity $\beta_1 = \frac{L_1}{r}$ (2)

Linear equation $L_1 = \bar{x}_1 x_{11} + \bar{x}_2 x_{12} + \cdots + \vec{x}_k x_{1k}$ (3)

Effective divider $r = \bar{x}_1^2 + \bar{x}_2^2 + \cdots + \vec{x}_k^2$ (4)

After the initial discussion of total variation, S_T element variation is proportional to S_β, error variation, S_e and error variance, V_e that can be seen as equations (5), (6), (7), and (8) are computed before being substituted into standard SN ratio η.

Total variations $S_{T1} = {x_{11}}^2 + {x_{12}}^2 + \cdots + {x_{1k}}^2$ (f = k) (5)

Variation of proportional term $S_{\beta 1} = \frac{L_1^2}{r}$ (f = 1) (6)

Error variation $S_{e1} = S_{T1} - S_{\beta 1}$ (f = k – 1) (7)

Error variance $V_{e1} = \frac{S_{e1}}{k-1}$ (8)

Then, equation (9) calculates the standard SN ratio, η. The stronger correlation between the input and output is indicated by the higher of η values.

Standard SN ratio $\eta_1 = \frac{1}{V_{e1}}$ (9)

The next step is to compute the two variables Y_1 and Y_2 using the sensitivity, β and standard SN ratio, η in the normal group. A scatter diagram is created by calculating the Y_1 and Y_2 variables. The formula of Y_1 and Y_2 are shown in equation (10) and (11) correspondingly.

$Y_{i1} = \beta_i$ (i = 1, 2, …, n) (10)

$Y_{i2} = \frac{1}{\sqrt{\eta_i}} = \sqrt{V_{ei}}$ (i = 1, 2, …, n) (11)

The computation of the average for Y_1 and Y_2 are in equations (12) and (13) where the origin prediction is found.

$\bar{Y}_1 = \frac{1}{n}(Y_{11} + Y_{21} + \cdots + Y_{n1})$ (12)

$\bar{Y}_2 = \frac{1}{n}(Y_{12} + Y_{22} + \cdots + Y_{n2})$ (13)

The final step in normal group involves using equation (14) to determine the Mahalanobis distance (MD).

Mahalanobis distance $D^2 = \frac{YA^{-1}Y^T}{k}$ (14)

Again, in order to compute the abnormal group, the same equation as the normal group is repeated. However, the abnormal group should be normalised before being calculated. It is important to take note of the fact that the average values, $\bar{x}$ of the samples and parameters in equation (1), as well as the value of the effective divider, r' in equation (4) are equivalent to normal group value.

The significant parameter is established by T method after RT method defines the MD of normal and abnormal. T method creates a uniform unit space with average output values. The highest sample will be designated as a normal group, while the remaining samples will be designated as an abnormal group. In normal group unit space, the average value for each parameter and the output average value may be calculated as stated in equations (15) and (16).

$\bar{x}_j = \frac{1}{n}(x_{1j} + x_{2j} + \cdots + x_{nj})$ (j = 1, 2, …, k) (15)

$\bar{y} = M_0 = \frac{1}{n}(y_1 + y_2 + \cdots + y_n)$ (16)

The purpose of abnormal group normalization is to reduce redundant from data in order to make it more flexible. Equations (17) and (18) demonstrate the computation of normalised data for input and output.

$X_{ij} = x'_{ij} - \bar{x}_j$ (i = 1, 2, …, l) (j = 1, 2, …, k) (17)

$M_i = y'_i - M_0$ (i = 1, 2, …, l) (18)

The proportional coefficient, β and SN ratio, η for each parameter are computed by following the steps outlined in equations (19), (20), (21), (22), (23), (24), and (25).

Effective divider $r = M_1^2 + M_2^2 + \cdots + M_l^2$ (19)

Total variation $S_{T1} = {X_{11}}^2 + {X_{21}}^2 + \cdots + {X_{l1}}^2$ (f = l) (20)

Variation of Proportional term $S_{\beta 1} = \frac{(M_1X_{11} + M_2X_{21} + \cdots + M_1X_{11})^2}{r}$ (f = 1) (21)

Error variation $S_{e1} = S_{T1} - S_{\beta 1}$ (22)

Error variance $V_{e1} = \frac{S_{e1}}{l-1}$ (23)

Proportional coefficient $\beta_1 = \frac{M_1X_{11} + M_2X_{21} + \cdots + M_1X_{11}}{r}$ (24)

SN ratio $\eta_1 = \begin{cases} \frac{\frac{1}{r}(S_{\beta 1} - V_{e1})}{V_{e1}} & (\text{when } S_{\beta 1} > V_{e1}) \\ 0 & (\text{when } S_{\beta 1} \le V_{e1}) \end{cases}$ (25)

A positive value of β indicates that the slope is increasing to the right, whereas a negative value indicates that the slope is decreasing to the right. The value of η should be positive, but if it is negative, it is considered to be zero, indicating that there is no longer a significant relationship between input and output.

The proportional coefficient, β and SN ratio, η for each parameter are used to get the integrated estimate value of the abnormal group. Equation (26) shows how to calculate the integrated estimate value. The normalised values of each parameter are represented by $x_{i1}, x_{i2}, \ldots, x_{ik}$.

$$\widehat{M}_{i1} = \frac{\eta_1 \times \frac{x_{i1}}{\beta_1} + \eta_2 \times \frac{x_{i2}}{\beta_2} + \ldots + \eta_k \times \frac{x_{ik}}{\beta_k}}{\eta_1 + \eta_2 + \ldots + \eta_k} \qquad (i = 1, 2, \ldots, l) \qquad (26)$$

Utilising the following equations (27), (28), (29), (30), (31), (32), and (33) are the steps involved in the calculation fo the estimated SN ratio, η. In point of fact, the suitability of orthogonal array (OA) establishes a basis for the estimated SN ratio, η.

Linear equation $L = M_1\widehat{M}_1 + M_2\widehat{M}_2 + \cdots + M_l\widehat{M}_l$ (27)

Effective divider $r = M_1^2 + M_2^2 + \cdots + M_l^2$ (28)

Total variation $S_T = \widehat{M}_1{}^2 + \widehat{M}_2{}^2 + \cdots + \widehat{M}_l{}^2$ (f = l) (29)

Variation of Proportional term $S_\beta = \frac{L^2}{r}$ (f = 1) (30)

Error variation $S_e = S_T - S_\beta$ (31)

Error variance $V_e = \frac{S_e}{l-1}$ (32)

Integrated Estimate SN Ratio $\eta = 10\log\left(\frac{\frac{1}{r}(S_\beta - V_e)}{V_e}\right)$ (db) (33)

The relative importance of parameter is measured by how much the estimated SN ratio, η reduces when the parameter is not used. The evaluations employ two-level OA, namely level 1 and level 2. OA measures the expected SN ratio, η under different conditions. The two-level OA implies parameter level 1 will be used and parameter level 2 will not be used. For the estimated SN ratio, η, the difference between the averages of level 1 and level 2 for each parameter determines the relative relevance of the parameters. When utilised with larger SN ratio, η, the parameter of degree of contribution is positive. Otherwise, when not utilised, the parameter of degree of contribution is negative.

RESULTS AND DISCUSSION

The scatter diagram shows the normal and abnormal data produced from Mahalanobis distance (MD) values using the RT method for the distribution of rainfall. In order to produce the scatter diagram, the Y1 and Y2 data from both the normal and abnormal conditions were utilised. The Y1 value is presented along the x-axis and the Y2 value is presented along the y-axis. Orange and blue are the two categories that are used to separate the related patterns. On the graph, the normal group is represented by dots that are blue colour, while the abnormal group is represented by dots that are orange colour. There are 16 parameters that are involved in the process of creating the implementation method while using this approach.

The normal and abnormal findings for the month of June are displayed in Figure 2. Normal data have a maximum classification MD value of 6.2666 and a minimum value of 0.0001 for the MD classification. On the other hand, the MD value for abnormal classification starts at a minimum of 0.0234 and a maximum of 36.9001. There is overlap between normal and abnormal data in the scatter diagram. There are 30 samples amount of overlapping data that occurred. The average MD value for the month of June is 1 that belongs to the normal group, while 4.0729 belongs to the abnormal group.

The value of the correlation coefficient for normal data is then found to be -0.5406. This demonstrates that Y1 and Y2 have a weak negative correlation with one another. On the other hand, the abnormal correlation coefficient between Y1 and Y2 is -0.3857, which also indicates a weak correlation between the two variables.

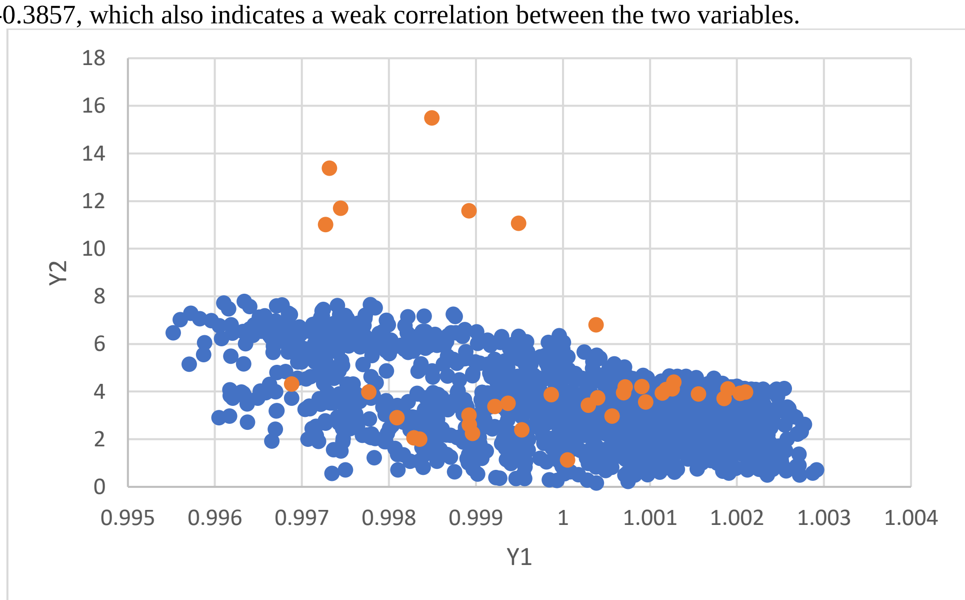


Figure 2. Scatter diagram of June

Figure 3 shows the normal and abnormal outcomes for October. The MD classification of normal data ranges from a maximum of 8.1492 to a minimum of 0.0007. For abnormal classification, the MD value ranges from a minimum of

0.0201 to a maximum of 45.7996 The scatter diagram shows overlap between normal and abnormal data. There are 91 occurrences of redundant data. For the month of October, the normal average MD value is 1, while the abnormal average MD value is 3.7726.

The normal data correlation coefficient is therefore calculated to be -0.5538. This proves that there is a weak negative relationship between Y1 and Y2. However, there is a weak negative relationship between Y1 and Y2 according to the abnormal correlation coefficient of -0.1994.

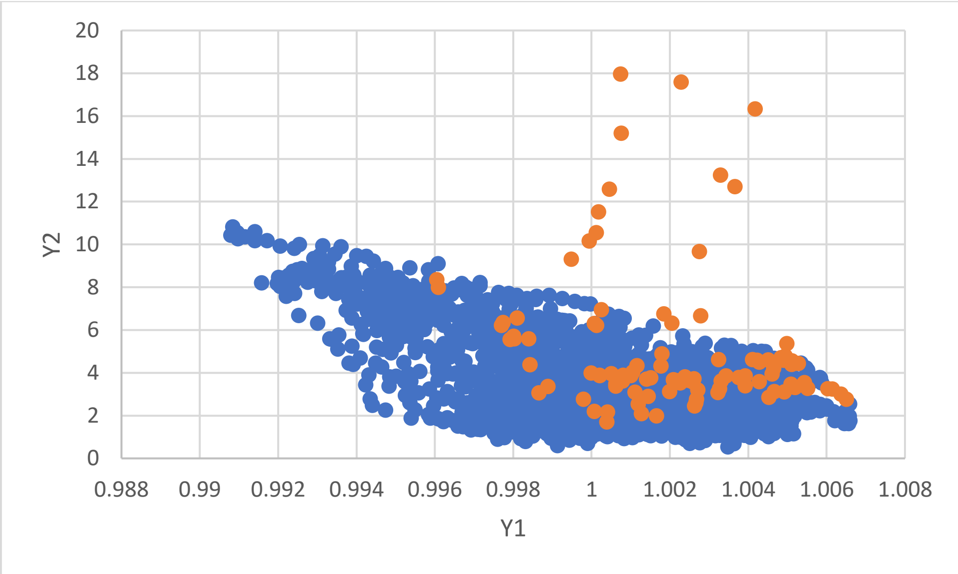


Figure 3. Scatter diagram of October

Figure 4 shows the normal and abnormal for the month of December. The maximum value for the MD classification of normal data is 9.7474, and the minimum value is 0.0015. The minimum MD value for abnormal classification is 0.0016 and the maximum is 112.6262. In the scatter diagram, there is some overlap between the normal and abnormal data. There are 262 samples of data overlapping each other. The average MD value for the month of December is 1, which is in the normal group. On the other hand, 8.0844 is in the abnormal group.

Then, for normal data, the correlation coefficient is found to be -0.2894. This shows that there is a weak negative correlation between Y1 and Y2. The abnormal correlation coefficient between Y1 and Y2 is -0.1532, which shows that two variables are weak negative correlation.

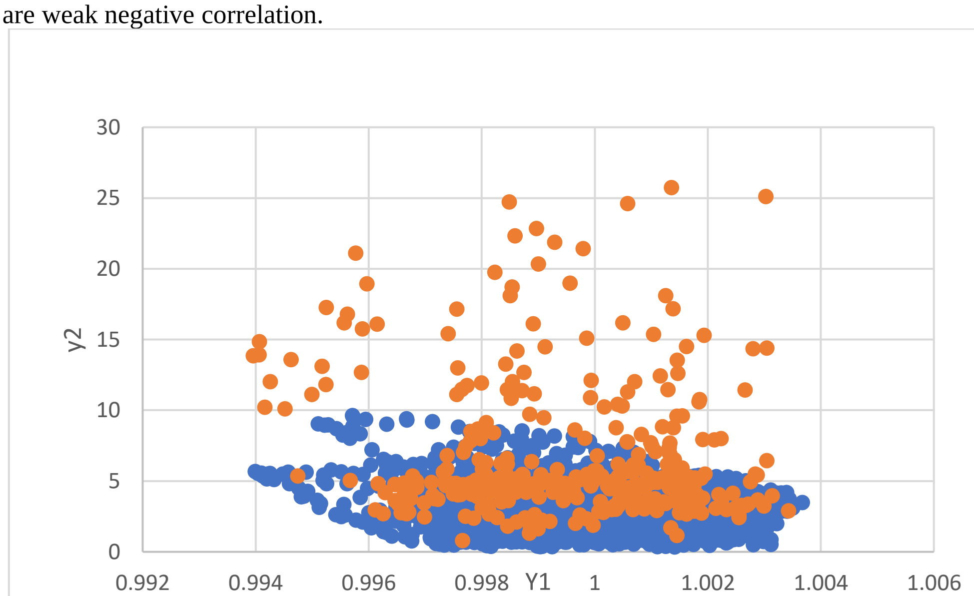


Figure 4. Scatter diagram of December

The number of normal group in the rainfall distribution is 2873, whereas the number of abnormal group is 103 and there are 17 total parameters. As can be seen in Figure 5, the data are arranged so that increasing output value corresponds to increasing order. The data collected is in December where the transition from southwest to northeast happened. The minimum MD value was 0.0007 for site number 1646, while the maximum MD value was 45.7996 for site number 3. Meanwhile, site number 2243 data was 1.0955, which is the closest to the average output value of all data (1.0960).

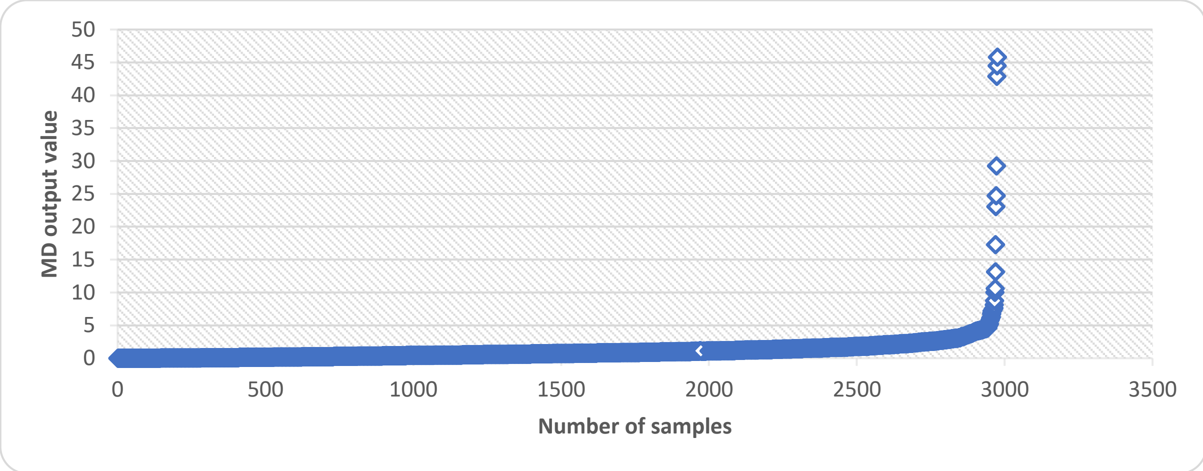
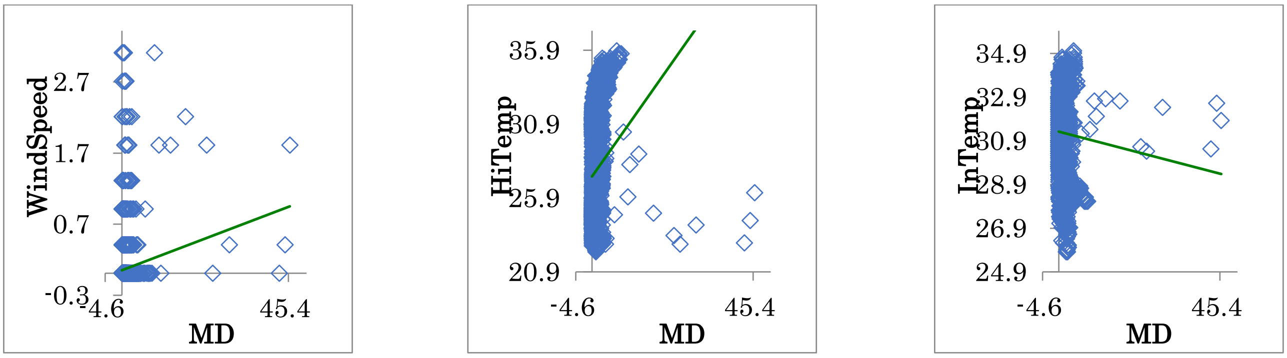


Figure 5. Rainfall distribution of output value

Then, the average values from the unit space samples are utilised to normalize the signal data. The signal data's parameter output values are normalized through subtracting the average of the parameter and unit space outputs. The output values are ordered ascending. The T method normalizes without dividing the parameter's standard deviation by what remains after removing the average values, unlike statically-based methods. Parameters with unit space standard deviation of zero that cannot be calculated are usually significant for prediction and estimation.

The prediction and estimation elements are found using proportional coefficient, β and SN ratio, η. Figure 6 shows how parameters are affected the output values. The output value is on the horizontal axis and the origin is on the vertical axis.

The proportional coefficient, β is positive in Figure 6 (a) and (b) because it is a straight line rising to the right. The plot of (a) demonstrates that the data are well distributed along the regression line, which goes near to zero and the outlier variance is low. The regression line plot for (b) is far from zero, but the data is well distributed. Both graphs assume that the SN ratio, η is big and that the wind speed and high temperature parameters are good for general yield estimation. Next, the proportional coefficient, β becomes negative when the line in (c) tilts to the right. The graph also demonstrates that the data are nicely distributed along the regression line with low variance. The inside temperature parameter and SN ratio, η imply output value estimations.



(a) (b) (c)

Figure 6. Scatter of output values and parameters (normalized data)

Figure 7 is a scatter diagram illustrating what happens when the real values are expressed in x-axis terms and estimated values in y-axis terms. Furthermore, the graph will provide extra information about an approximate straight line and its features. In general, the model provides 0.7314 of R^2 or -3.981 db of SN ratios. It indicates that the correlation is high and the distribution is approaching the green line. This demonstrates that the real value and estimated value for the signal data distribution have a strong relationship with one another. The estimated values that line up above a straight line demonstrates that a decent estimate was obtained. Equation (34) shows the line's equation.

Line equation $y = 1.1456x$ (34)

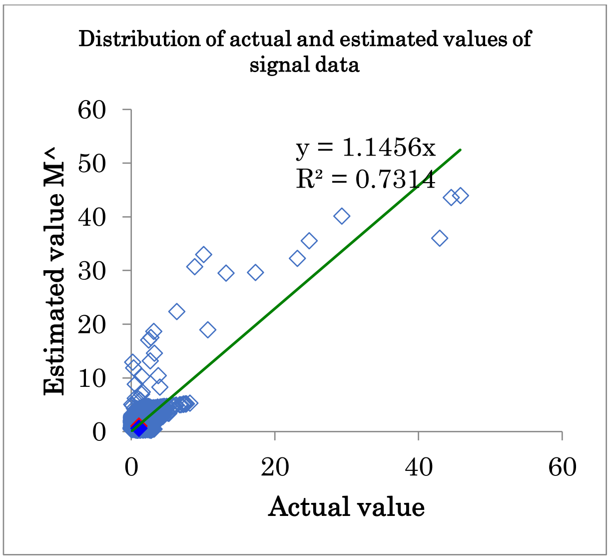


Figure 7. Distribution of actual and estimated values of signal data for rainfall distribution

However, just a selected number of those parameters can be utilised for integrated estimation, while the rest are unable to. As a result, the evaluation of the parameters is carried out by the utilisation of the L20 of OA, where level 1 indicates that the parameter will be utilised and level 2 indicates that the parameter will not be utilised. The value that was identified, which is -3.981 db of integrated estimate SN ratio was derived from the first run in L20. After that, the degree of contribution was converted into a bar graph, which can be seen in Figure 8.

The positive degree of contribution increases MD output, whereas negative degree of contribution decreases MD output. As a result, parameters 1, 2, 4, 9, 10, 12, 13, 14, and 16 have a positive degree of contribution. Meanwhile, parameters 3, 5, 6, 7, 11, 15, and 17 have a negative degree of contribution. According to the findings of this research, in order to achieve a lower MD, the positive degree of contribution should be reduced while the negative degree of contribution should be enhanced. As shown in Figure 8, this research shows that the transition from southwest to northeast has optimized 17 parameters into 9 parameters by lowering 8 parameters utilizing the MTS approach. It is important to note that the significant parameters are the decreased number of factors that MTS has advised for the future prediction and classification purposes.

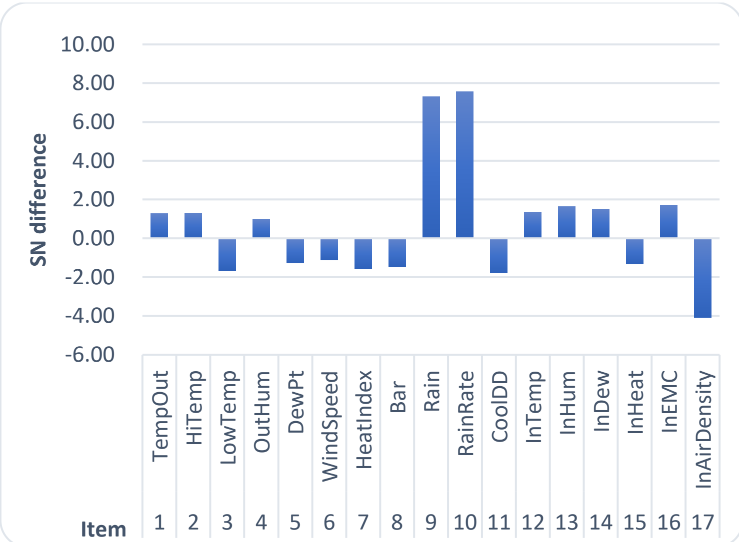


Figure 8. Degree of contribution for rainfall distribution

CONCLUSION

This research has been found that the MTS is capable of efficiently differentiating between the unit space data and the signal data within the context of rainfall data. Following completion of the computation, it was found that the MD values for unit space throughout the period of three months was equal to 1. Furthermore, it was discovered that in the month of June (southwest), the average MD value for the signal data is set as 10.7976. The average MD value in October (transition from southwest to northeast) is 13.4923 correspondingly. In comparison, in December (northeast) the average MD value is 30.5942. In addition to its capacity to determine the significant parameters for the rainfall patterns connected with monsoon phenomena, the MTS has been used to evaluate the data gathered in the UMPSA Campus Gambang, Malaysia.

In order to achieve optimal performance of the system, the total parameters were reduced from 16 parameters to 8 parameters. However, there were 17 parameters is reduced into 9 parameters that occurred in the month of October. Since this month have additional parameter of the windspeed because of it is the monsoon phenomena of transition from southwest to northeast. Lastly, according to MTS that these significant parameters should be taken into consideration while performing classification and prediction work in the future.

MTS is clearly employed in manufacturing industries including welding, mechanical materials, steel products, rotating equipment, and rolling bearing problem diagnostics. On the other hand, it might challenge to locate MTS in the weather application. It would be helpful if further research could be done to further expand on the study of the distribution of rainfall.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Higher Education for providing financial support under Fundamental Research Grant Scheme (FRGS) No. FRGS/1/2022/TK10/UMP/02/8 (University reference RDU220108).

REFERENCES

[1] K. Pu, X. Liu, X. Sun, and S. Li, “Error Analysis of Rainfall Inversion Based on Commercial Microwave Links With A-R Relationship Considering the Rainfall Features,” *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, doi: 10.1109/TGRS.2023.3253949.

[2] G. Zhang, B. Li, L. Li, X. Zhou, X. Xu, and L. He, “Multiple Time-scale Characteristics Analysis of Rainfall in Hunan Province Based on Ensemble Empirical Mode Decomposition,” *2019 3rd IEEE Conf. Energy Internet Energy Syst. Integr. Ubiquitous Energy Netw. Connect. Everything, EI2 2019*, pp. 851–855, 2019, doi: 10.1109/EI247390.2019.9061699.

[3] World Bank Group, “Climate Risk Country Profile: Uganda,” *World Bank Gr.*, p. 36, 2021, [Online]. Available: www.worldbank.org.

[4] F. Tangang *et al.*, “Characteristics of precipitation extremes in Malaysia associated with El Niño and La Niña events,” *Int. J. Climatol.*, vol. 37, pp. 696–716, 2017, doi: 10.1002/joc.5032.

[5] M. Niyongendako, A. E. Lawin, C. Manirakiza, and B. Lamboni, “Trend and Variability Analysis of Rainfall and Extreme Temperatures in Burundi,” *Int. J. Environ. Clim. Chang.*, vol. 10, no. 6, pp. 36–51, 2020, doi: 10.9734/ijecc/2020/v10i630203.

[6] Z. M. Marlan, F. Ramlie, K. R. Jamaludin, and N. Harudin, “Enhanced Taguchi’s T-method using angle modulated Bat algorithm for prediction,” *Bull. Electr. Eng. Informatics*, vol. 11, no. 5, pp. 2828–2835, 2022, doi: 10.11591/eei.v11i5.4350.

[7] B. John and R. S. Kadadevarmath, “A methodology for quantitatively managing the bug fixing process using mahalanobis taguchi system,” *Int. J. Ind. Eng. Comput.*, vol. 5, no. 12, pp. 1081–1090, 2015, doi: 10.5267/J.MSL.2015.10.006.

[8] 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, doi: 10.1108/IJQRM-10-2016-0174.

[9] Z. Chang, W. Chen, Y. Gu, and H. Xu, “Mahalanobis-taguchi system for symbolic interval data based on kernel mahalanobis distance,” *IEEE Access*, vol. 8, pp. 20428–20438, 2020, doi: 10.1109/ACCESS.2020.2967411.

[10] Z. P. Chang, Y. W. Li, and N. Fatima, “A theoretical survey on Mahalanobis-Taguchi system,” *Meas. J. Int. Meas. Confed.*, vol. 136, pp. 501–510, 2019, doi: 10.1016/j.measurement.2018.12.090.

[11] H. Ghorbani, “Mahalanobis Distance and Its Application for Detecting Multivariate Outliers,” *Facta Univ. Ser. Math. Informatics*, no. October 2019, p. 583, 2019, doi: 10.22190/fumi1903583g.

[12] M. El-Banna, “Modified Mahalanobis Taguchi System for Imbalance Data Classification,” *Comput. Intell. Neurosci.*, vol. 2017, 2017, doi: 10.1155/2017/5874896.

[13] N. Harudin *et al.*, “Binary Bitwise Artificial Bee Colony as Feature Selection Optimization Approach within Taguchi’s T-Method,” *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/5592132.

[14] A. A. G. Nadiatul Adilah, M. Mohamad Zarif, and A. Mohamad Idris, “Rainfall Trend Analysis using Box Plot Method: Case Study UMP Campus Gambang and Pekan,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 712, no. 1, 2020, doi: 10.1088/1757-899X/712/1/012021.

[15] S. Teshima, Y. Hasegawa, and K. Tatebayashi, “Quality_Recognition_and_Prediction: Smarter Pattern Technology with the Mahalanobis-Taguchi System,” *Momentum Press*, 2012, https://books.google.com.my/books?id=q1S4uAAACAAJ