Applied Mathematics and Computational Intelligence
Volume 11, No.2, Dec 2022 [492-504]

Multivariate Methods Approach for Water Quality Index Estimation

O. Kairan[1], N.F. Roslan [2], N.A. Samsudin[3], N.S. Shaharani[4], F.F. Suhami[5], A.A. Rahman[6], M.N.A. Nafi[7]

[1,2,3,4,5,6,7]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Kelantan

Corresponding author: omarkr@uitm.edu.my

Received: 28 September 2022; Accepted: 7 November 2022; Available online (In press): 5 December 2022

ABSTRACT

Water Quality Monitoring is crucial in determining the healthy level of a river. The variety of methods for monitoring water quality is essential to cover mistakes or damage in the future. For this reason, the comparison between the multivariate approach and the existing water quality index formulation was carried out in this study. The secondary data on the water quality of the Selangor river basin used in this study was obtained from the Department of Environment. The multivariate methods, specifically Factor Analysis (FA) and Principal Component Analysis (PCA) were used to generate Water Quality Index (FA-WQI and PCA-WQI). The result shows that FA-WQI is closer to DOE-WQI compared to PCA-WQI.

Keywords: Department of Environmental Water Quality Index (DOE-WQI), Factor Analysis Water Quality Index (FA-WQI), Principal Component Analysis Water Quality Index (PCA-WQI).

1 INTRODUCTION

Increasing growth, expanding economic activity, and urban sprawl contribute to increased water demand. Overuse of groundwater and surface water places a range of services at risk as a consequence of the decline of available supplies and the loss of their importance [1]. In order to track source pollution and improve water quality, it is necessary to know the spatial temporal characteristics of organic contaminant concentrations in the affected receiving waters [2]. River water quality can be calculated either with a few important parameters of water quality or with individual parameters [3]. The Water Quality Index (WQI) has been regarded as one of the possible river water classification standards, focusing on the use of common water classification parameters. The WQI reflects the standard of water quality, a numerical expression used to convert vast quantities of water categorization data into a single number (Edge). There are two key indices used to evaluate the consistency of river water monitored; the Water Quality Index (WQI) which, in turn, is focused on the Interim National Water Quality Guidelines (INWQS), a set of parameters extracted from the beneficial application of water.

Many water quality indices are developed by individual scientists by modifying previous indices, all of which leave the decision as to how and how to implement all the parameters in the professional judgement of the consumer. Through analyzing 36 indices, [4] noticed that there were substantial differences between the water quality ratings given by various indices in the same water sample. Many other prior studies have shown that various estimation methods have produced specific results [5][4][6].

For a multitude of indices that have been developed and used, it is difficult to say the world whether the index is the highest or even the list of 'ten best' or 'twenty best' indices. A number of indexes are more general than some others. For example, the WQI of the US National Sanitation Foundation, widely referred to as the NSF-WQI [7], is used not only in its country of origin but also in many other countries that cover many continents (Brazil, Mexico, Guinea-Bissau, Poland, Egypt, Portugal, Italy and India, among others). WQI of the Canadian Council of Ministers of the Environment, called CCMEWQI [8], which was formed 31 years after the NSF-WQI. In Malaysia the water quality status of rivers has always been a cause concern for different local authorities, government agencies and the open to the general public. In Malaysia, rivers are usually considered polluted with consistent examples such as Sg. Klang in Selangor. We can deduce from physical observation alone that something is wrong with the present state of these rivers in terms of water quality. On a technical point of view, however, the degree of emissions also needs to be quantified to address pollution problems in a comprehensive and coordinated manner. The reason why river in Selangor was chosen because it considered one of the most polluted river in Malaysia. This study estimates the water quality index using multivariate methods and compare the index with available water quality index applied by Department of Environment Malaysia.

In Malaysia, the classification of the river by the Department of the Environment (DOE) is based on the WQI. The WQI compares the water quality determinants on a category to a specific scale and integrates them into a single number, in order to comply with the selected measurement process or model. The WQI was developed based on the results of the panel of experts' opinion poll, who decided on the option of the determinants and weights allocated to each of the determinants of water quality chosen. DOEWQI consists of six determinants: dissolved oxygen (0.22), biological oxygen requirement (0.19), chemical oxygen requirement (0.16), ammonia nitrogen (0.15), suspended solids (0.16) and pH (0.12). Weight factors are specified in brackets. The WQI has been studied in Malaysia for around 34 years since it first established in [9]. A few shortcomings of the new Malaysian WQI have been established by [10], mainly linked to the choice of parameters, sub-index equations and weighting factors.

Previously, there are numerous multivariate methods specifically Factor Analysis (FA) and Principal Component Analysis (PCA) are applied mostly on characterizing or categorizing the water quality parameter into groups that forming the type of sources of pollution. The latest five years articles include of [11], [12], [13], [14] and [15] use similar application of these methods. However, there are limited application of PCA and FA in order to estimate water quality index. Principal component score was transformed to an index by [16]. Later [17] and [18] used the same method of [16] to obtain indices. By referring the previous study, this research tries to compare the consistency of PCA and FA generated index towards the DOE water quality index.

2 MATERIAL AND METHODS

2.1 Sources of Data

A secondary data collection on the water quality of the Selangor River ranges from 2008 to 2018 were be obtained from the Department of the Environment. This covers four stations that are selected in Selangor River Basin which are ISR01, ISR03, ISR05 and ISR07, from highly polluted to less contaminated or unpolluted rivers. The data should be used to achieve the indices for each of the following listed methodology.

2.1.1 Study Area

This study is located in Selangor area. Only four stations that are selected in Selangor River Basin which are 1SR01, 1SR03, 1SR05 and 1SR07. Station 1SR01 is located to Sungai Selangor at Kampung Kelip-Kelip and nearest to kuala. For station 1SR03 is located to Sungai Batang Kali at Jambatan Pekan Batang Kali and nearest to Batang Kali's town. Next, for station 1SR05 is located to Sungai Selangor and placed in Empangan K.Kubu. This location is nearest to Sungai Selangor dam. Lastly, for station 1SR07 is located at Sungai Kanching and placed in Taman Templer Park Komanwel that nearest to industrial area. All this station selected because the most polluted area to calculate and identify the water quality index to make the comparison. The Selangor Basin's Station are shown as figure below:

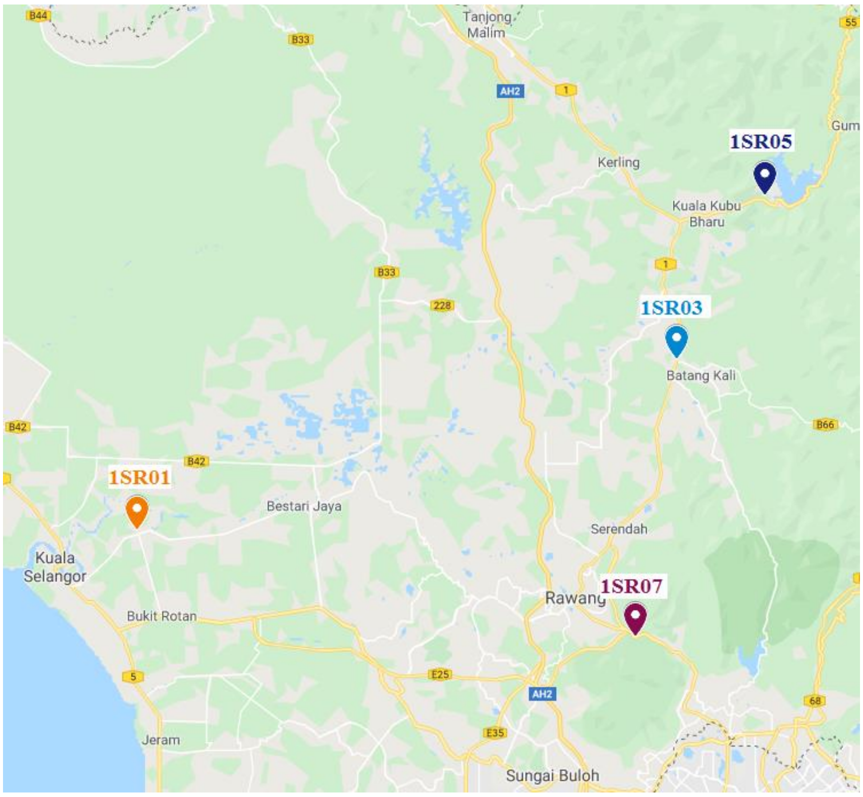

Figure 1: Maps of Selangor Basin's Station

2.1.2 Data Description

For DOEWQI, in this analysis, six parameters of Chemical Oxygen Demand (COD), Dissolved Oxygen (DO), Ammonia Nitrogen (AN), Biochemical Oxygen Demand (BOD), Suspended Solids (SS) and pH were collected for the WQI. The measurement was based on the value of the sub-indices called Sub-Indices Dissolved Oxygen (SIDO), Sub-Indices Biochemical Oxygen Demand (SIBOD), Sub-Indices Chemical Oxygen Demand (SICOD), Sub-Indices Suspended Solids (SISS), Sub-Indices Ammoniacal Nitrogen (SIAN) and Sub-Indices pH (SIpH). PCA-WQI and FA-WQI are used same six parameters were chosen in this study which are Sub-Indices Dissolved Oxygen (SIDO), Sub-Indices Biochemical Oxygen Demand (SIBOD), Sub-Indices Chemical Oxygen Demand (SICOD), Sub-Indices Suspended Solids (SISS), Sub-Indices Ammoniacal Nitrogen (SIAN) and Sub-Indices pH (SIpH).

2.2 Method of Analysis

2.2.1 Department of Environment Water Quality Index (DOEWQI)

As a matter of fact, more than 120 Biological and physico-chemical requirements have been reviewed using the INWQS. The WQI indexing method has also been implemented to improve the vast volume of data gathered to assess the level of water quality. Six parameters were chosen in this study for the WQI which are Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Suspended Solids (SS), Ammoniacal Nitrogen (AN) and pH. The measurement was being determined on their sub-indices value which being called as Sub-Indices Dissolved Oxygen (SIDO), Sub-Indices Biochemical Oxygen Demand (SIBOD), Sub-Indices Chemical Oxygen Demand (SICOD), Sub-Indices Suspended Solids (SISS), Sub-Indices Ammoniacal Nitrogen (SIAN) and Sub-Indices pH (SIpH). Therefore, the best fit equations used for estimating the six sub-index values are shown as Table 1 based on DOE's Environmental Quality Report, 2006.

Table 1: Sub-Index and Formula

Sub-Index	Formula
Sub-Index Dissolved Oxygen (SIDO)	SIDO = 0 for DO ≤ 8 % = 100 for DO ≥ 92 % = -0.395 + 0.030DO^2 -0.00020DO^3 for 8 % < DO< 92 % Dissolved Oxygen in % saturation (DO)
Sub-Index Biochemical Oxygen Demand (SIBOD)	SIBOD= 100.4 - 4.23BOD for BOD ≤ 5 = 108$e^{-0.005BOD}$ - 0.1 for BOD > 5 Biochemical Oxygen Demand in mg/l concentration (BOD)
Sub-Index Chemical Oxygen Demand (SICOD)	SICOD= -1.33COD + 99.1 for COD ≤ 20 = 103$e^{-0.0157COD}$- 0.04x for COD > 20 Chemical Oxygen Demand (COD), in mg/l concentration
Sub-Index Suspended Solids (SISS)	SISS = 97.5$e^{-0.00676SS}$ = 0.05SS for SS ≤ 100 = 71$e^{-0.0016SS}$ – 0.015SS for 100 <SS < 1000 = 0 for SS ≥ 1000 Suspended Solids (SS), in mg/l concentration
Sub-Index Ammoniacal Nitrogen (SIAN)	SIAN = 100.5 - 105AN for AN ≤ 0.3 = 94$e^{-0.573AN}$ – 5AN – 2l for 0.3 < AN < 4 = 0 for AN ≥ 4 Ammoniacal Nitrogen (AN) in mg/l concentration
Sub-Index pH (SIPH)	SIPH = 17.2 -17.2pH + 5.02pH^2 for pH<5.5 = -242+95.5pH - 6.67pH^2 for 5.5 ≤ pH< 7 = -181+82.4pH - 6.05pH^2 for 7 ≤ pH< 8.75 = 536 - 77.0pH + 2.76pH^2 for pH ≥ 8.75

Source [19]

WQI can be calculated as:

$$DOEWQI = 0.22(SIDO) + 0.19(SIBOD) + 0.16\ (SICOD) + 0.15\ (SIAN) + 0.16(SISS) + 0.12(SIPH) \quad (1)$$

2.2.2 Principal Component Method Water Quality Index

The basic principle of the main component analysis (PCA) is to minimize the dimensionality of a data set composed of a large number of interrelated variables while preserving as much variance as possible in the data set. All new variables are separated, i.e. they are not correlated with each other (where the previous, untransformed variables might have been correlated to a lesser or greater extent). In principle, each of the key component is a linear combination of the initial X values for the p variables given as:

$$PCA_1 = e_{11}ZSICOD + e_{12}ZSIDO + e_{13}ZSIAN + e_{14}ZSIBOD + e_{14}ZSISS + e_{15}ZSIPh \quad (2)$$

In this analysis, the correlation structure between the sub-indices was analyzed using a parametric correlation scale, namely the pearson correlation coefficient. All sub-indices, that is SIDO, SIBOD, SICOD, SIAN, SISS and SIPH from the research data collection (2008-2018) were configured to represent variables on a single unified scale. The eigenvectors and eigenvalues were calculated based on a basic formula as shown in Equation 3.

$$Ae = \lambda e \quad (3)$$

Where **A** is the covariance matrix and **e** is an eigenvector of **A** corresponding to the eigenvalue λ if and only if **e** and λ satisfy$(A - \lambda I)e = 0$. The solution of Equation 3 can be easily done using PCA procedures in R. Finally, the new variables and main component PC scores were computed as shown in Equation 4.

$$PCA_1 = \sum_{j=1}^{p} e_{j1}Z_j \quad (4)$$

Where e_{j1} is eigenvector or weights of standardized sub-indices in the first PCA, Z_j = standardized form of subindices and PCA_1 is first PCA scores. The first PCA scores were chosen to illustrate the greatest data variability. The score value can be negative or positive. Subsequently, the scores were adopted as a water quality measure, PCAWQI, and standardized to a scale of 0 to 100 for ease of comparison with DOE.

The water quality scores calculated in Equation 5 can be either negative or positive. In order to ease the comparison, the scores were standardized on the based on Equation 5 [16].

$$PCA_2 = \frac{PCA_1 - Minimum}{Range} \quad (5)$$

Where the first principal component (PCA_2) score transformed to a 0 – 1 scale and then times with 100 to obtained index values. For minimum is minimum value of PCA_1 and Range is range value (maximum value minus minimum value) of PCA_1. The assumption of multivariate normality in PCA is relaxed due to the descriptive approach used in the development of WQI. Following the identification and estimation phase, the next step in the validation phase has been taken. At the validation point, the test data for 2008-2018 were used to verify the robustness of the index. Assuming that the test data share the distribution of the training data with the same mean μ and covariance matrix, the PCAWQI was re-calculated accordingly.

2.2.3 Factor Analysis Method Water Quality Index

Building on this, in FAWQI, Factor Analysis was applied to produce Water Quality Index. Similar to Standard Water Quality produce by DOE its used water quality data of six parameters were chosen in this study which are Sub-Indices Dissolved Oxygen (SIDO), Sub-Indices Biochemical Oxygen Demand (SIBOD), Sub-Indices Chemical Oxygen Demand (SICOD), Sub-Indices Suspended Solids (SISS), Sub-Indices Ammoniacal Nitrogen (SIAN) and Sub-Indices pH (SipH).The WQI was built in three steps:

a) preparing a matrix of correlations

b) abstraction of specific factors and potential spatial decrease

c) rotation of axes in relation to specific variables, looking for an easy and convenient solution to represent. For this research, varimax rotation was applied.

d) The factor score of the first factor is estimated using the regression [20] as stated below:

$$Factor\_1 = \sum_{p=1}^{m} a_{jp} F_{pi} + u_j Y_{ji} \text{ (i=1,2, K, N; j=1, 2, K, n)} \quad (6)$$

Where a_{jp} F_{pi} = contribution of the common p factor to the linear combination and $u_j Y_{ji}$ = residual error in the representation of the observed Factor_1 measurement. In order to ease the comparison, the scores were standardized on the based on Equation 6.

e) In order to estimate the FAWQI, similar method as in equation (5) were used;

$$FAWQI = \frac{Factor\_1 - Minimum}{Range} \times 100 \quad (7)$$

Where Factor_1 is the factor score for the first Factor. While minimum is minimum values of first Factor and Range is range values (maximum value minus minimum value) of first Factor. In this study, the assumption of multivariate normality in FAWQI is relaxed due to the descriptive approach used in the development of WQI. Following the identification and estimation phase, the next step in the validation phase has been taken. At the validation point, the test data for 2008-2018 were used to verify the index.

3 RESULT AND DISCUSSION

3.1 Descriptive statistics

Table 2: Descriptive Statistics for 6 Sub-index

Station		N	Mean	Std. Deviation	Variance
1SR01	SIDO	66	1.89918	12.263703	150.398
	SIBOD		73.35492	18.071726	326.587
	SICOD		66.66235	19.499852	380.244

	SIAN		71.44408	14.637259	214.249
	SISS		54.24933	24.115411	581.553
	SIPH		81.14644	25.199145	634.997
	SIDO		88.011076	32.5285412	1058.106
	SIBOD	66	85.401665	8.5038172	72.315
	SICOD		80.386314	8.6396813	74.644
	SIAN		79.578822	13.6057148	185.115
	SISS		84.119824	10.5821286	111.981
1SR03	SIPH		97.717975	2.7854207	7.759
	SIDO		83.506418	37.1634107	1381.119
	SIBOD	66	84.808070	10.0999798	102.010
	SICOD		80.227419	10.6115966	112.606
	SIAN		88.903728	11.8080638	139.430
	SISS		91.679159	7.5399403	56.851
1SR05	SIPH		97.519015	2.0145181	4.058
	SIDO		89.502713	30.7092933	943.061
	SIBOD	66	83.987930	9.9132956	98.273
	SICOD		80.194697	9.4680953	89.645
	SIAN		83.376059	10.0883223	101.774
	SISS		85.880838	13.9493473	194.584
1SR07	SIPH		96.193362	2.8003476	7.842

Descriptive analysis of subindices for the four stations produced as in table 2. Station 1SR03, 1SR05 and 1SR07 produce common results of variation compare to station 1SR01. Excluding the SIDO subindices on the station 1SR03, 1SR05 and 1SR07, variation of subindinces of these stations are all lower than variation as in station 1SR01.

3.2 Comparison DOEWQI between PCAWQI and FAWQI

Figure 2 shows that the graph between DOEWQI and PCAWQI for four stations. For station 1SR01, the graph shows a little gap movement between these two methods. For station 1SR03, 1SR05 and 1SR07, the graphs show a long gap movement between PCA-WQI and DOE-WQI. Generally, PCAWQI are over expected the river pollution rather than DOEWQI since range index for polluted in 0 to 59. Figure 3 shows that the graph between DOE-WQI and FA-WQI for station 1SR01, 1SR03, 1SR05 and 1SR07. The graphs in station 1SR01 and 1SR05 shows that there is a constant movement along the DOE-WQI's movement. For station 1SR03 and 1SR07, the graphs show that a little gap movement between these two methods. But, the overall graphs shows that FAWQI more closer to DOE-WQI compare to PCA-WQI.

3.3 Validation of PCA-WQI and FA-WQI

Table 3: Correlation Coefficient

STATION	Pearson Correlation	DOE-WQI vs FA-WQI	DOE-WQI vs PCA-WQI
1SR01	Coefficient	0.857**	-0.661**
	p-value	0.00	0.00
1SR03	Coefficient	0.410**	-0.197
	p-value	0.001	0.113
1SR05	Coefficient	0.414**	-0.276*
	p-value	0.001	0.025
1SR07	Coefficient	0.318**	-0.293*
	p-value	0.009	0.017

In order to find the best approach, multivariate method for estimating the water quality index correlation analysis was conducted. As depicted in the table 3 high correlation is detected between DOE-WQI and FA-WQI at station 1SR01. Whereas, the lowest correlation is occur between DOE-WQI and PCA-WQI is at station 1SR03. Comparing the correlation coefficient produce by both multivariate methods, FA-WQI is always high compare to PCA-WQI. However, FA-WQI not consistent with DOE-WQI for other stations. However, lower correlation might be due to outlier as depicted in the matrix plot in figure 4. FA-WQI for 1SR01 obviously not affected by outlier compare to the other location. Furthermore, PCA-DOE might be affected by outliers but the absent of the outlier still not makes the method of estimation better because of contradiction of trend values to DOE-WQI.

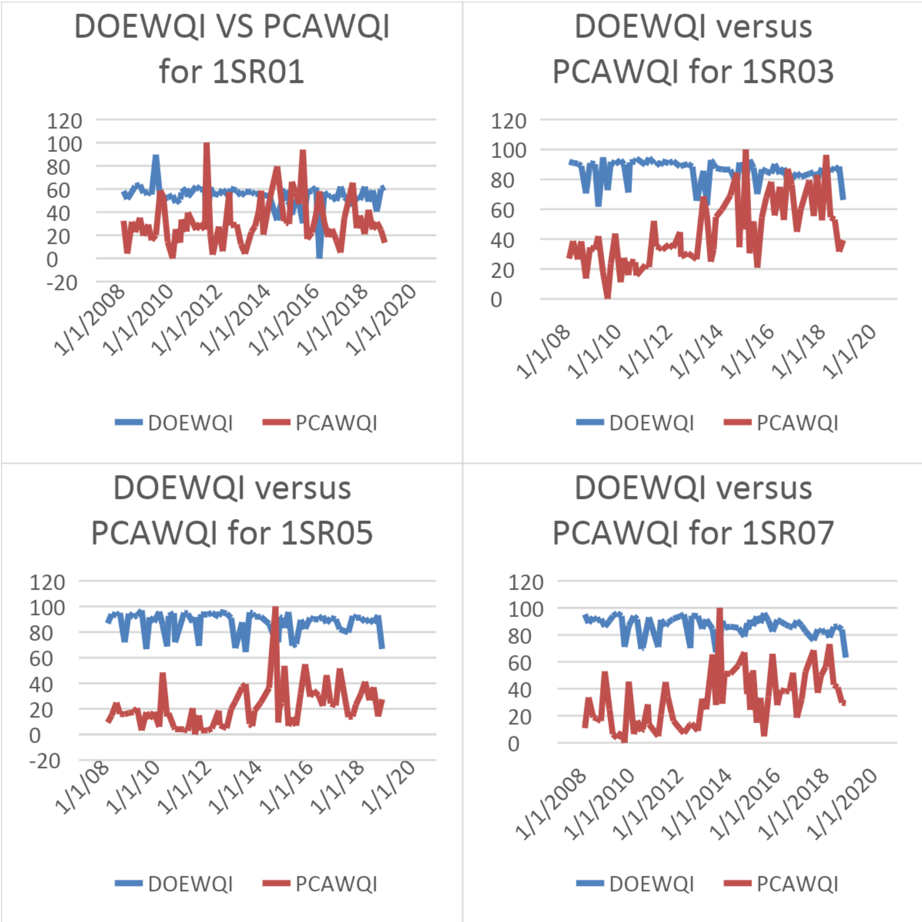

Figure 2: DOEWQI versus PCAWQI for four stations in Selangor River Basin

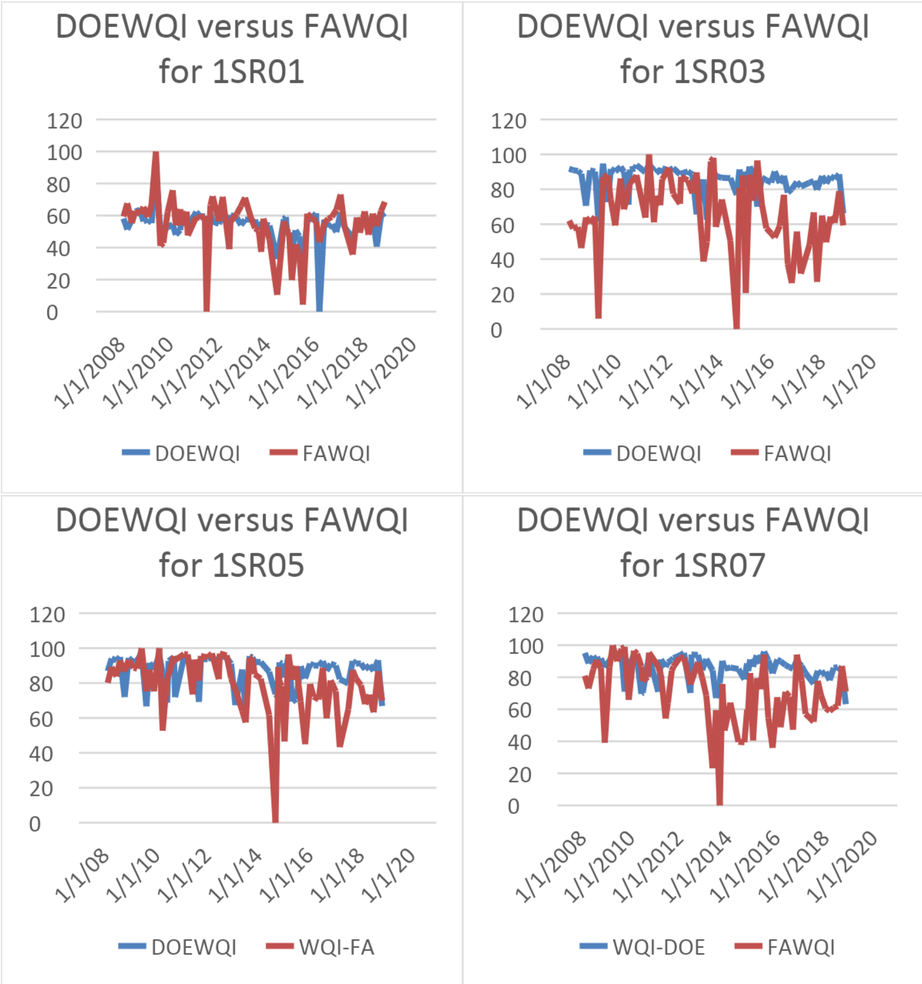

Figure 3: DOEWQI versus FAWQI for four stations in Selangor River Basin

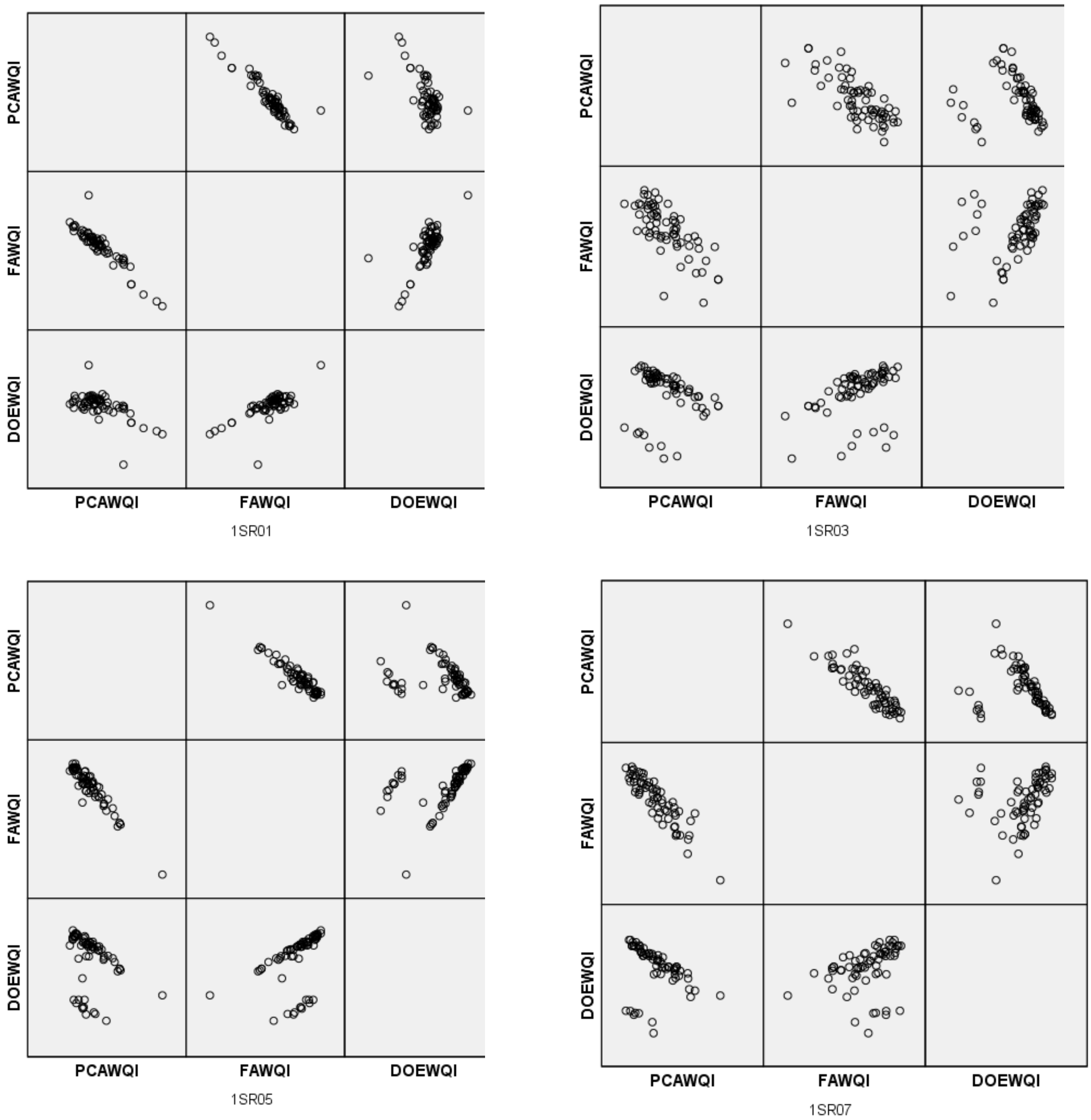

Figure 4 Matrix Scatter Plot of the three Water Quality index

4 CONCLUSION AND RECOMMENDATION

The results indicates that FAWQI more closer to DOEWQI rather than PCA-WQI. But it depends on where the river station is located and how the data are fluctuated or varied. Principle component method and Factor analysis method are performed better if the data are highly varied and the absent of outliers. Finally, it is known that it is impossible to establish a standard water quality metric with a conclusive approach. In the potential evaluation of the water quality index, it is recommended that statistics will be collected with long-term data along with some supplementary details in order to achieve a more precise figure of the results.

ACKNOWLEDGEMENT

This paper and the research behind it would not have been possible without the exceptional support of Universiti Teknologi Mara Kota Bharu Campus for the facilities provided for the completion of the study. A thousand appreciation to the Department of Environmental Malaysia (DOE) for providing environmental data.

REFERENCES

[1] M. A. Massoud,"Assessment of water quality along a recreational section of the Damour River in Lebanon using the water quality index," *Environ. Monit. Assess.*, vol. 184, pp. 4151–4160, 2012, doi: 10.1007/s10661-011-2251-z.

[2] Y. Lu, S. Song, R. Wang, Z. Liu, J. Meng, A. J. Sweetman, and T. Wang, "Impacts of soil and water pollution on food safety and health risks in China," *Environment international*, vol. 77, pp. 5-15, 2015.

[3] I. Mohamed, F. Othman, A. I. Ibrahim, M. E. Alaa-Eldin, and R. M. Yunus, "Assessment of water quality parameters using multivariate analysis for Klang River basin, Malaysia," *Environmental monitoring and assessment*, vol. 187, no. 1, p. 4182, 2015.

[4] N. Fernández, A. Ramírez, and F. Solano, "Physico-chemical water quality indices—A comparative review," *Bistua Revista de la Facultad de Ciencias Basicas*, vol. 2, pp. 19–30, 2004

[5] A. K. Yadav, P. Khan, and S. K. Sharma, "Water Quality Index Assessment of Groundwater in Todaraisingh Tehsil of Rajasthan State, India-A Greener Approach," *J. Chem.*, vol. 7, pp. 428–432, 2010.

[6] A. Akkoyounlu and M. E. Akiner, "Pollution evaluation in stream using water quality indices: a case study from Turkey's Sapanca lake basin," *Ecol Ind*, vol. 18, pp. 501–511, 2012.

[7] R. M. Brown, N. I. McClelland, R. A. Deininger, and R. G. Tozer, "A water quality index: do we dare?," *Water Sew. Works*, vol. 117, pp. 339-343, 1970.

[8] CCME. "Canadian Water Quality Guidelines for the Protection of Aquatic Life: CCME Water Quality Index 1.0," Technical Report, Canadian Council of Ministers of the Environment: Quebec city, QC, Canada, 2001.

[9] I. Naubi, N. H. Zardari, S. M. Shirazi, N. F. Ibrahim, and L. Baloo, "Effectiveness of Water Quality for Monitoring Malaysian River Water Quality," *Pol. J. Environ. Stud*, vol. 25, no. 1, pp. 231-239, 2016.

[10] A. Idris, A. A. Mamun, A. M. S. Mohd, and N. A. S. Wan, "Review of water quality standards and practices in Malaysia," *Pollution Research*, vol. 22, no. 1, pp. 145–155, 2003.

[11] S. N. F. Abdullah, A. Ismail, H. Juahir, R. B. Ahmad, F. Lananan, N. M. Hashim, and S. M. Desa, "Chemical composition of rainwater harvested in East Malaysia," *Environmental Engineering Research*, vol. 27, no. 2, 2022.

[12] A. Masthurah, H. Juahir, and N. B. Mohd Zanuri, "Case study Malaysia: Spatial water quality assessment of Juru, Kuantan and Johor River Basins using environmetric techniques," *Survey in Fisheries Sciences*, vol. 7, no. 2, pp. 19-40, 2021.

[13] I. B. Kokia, K. H. Lowa, S. M. Zaina, H. Juahirc, A. S. Bayerob, A. Azidd, and M. A. Zalie, "Spatial variability in surface water quality of lakes and ex-mining ponds in Malacca, Malaysia: the geochemical influence," *Desalination and Water Treatment*, vol. 197, pp. 319-327, 2020.

[14] H. Aydin, F. Ustaoğlu, Y. Tepe, and E. N. Soylu, "Assessment of water quality of streams in northeast Turkey by water quality index and multiple statistical methods," *Environmental forensics*, vol. 22, no. 1-2, pp. 270-287, 2021.

[15] M. F. Howladar, E. Chakma, N. J. Koley, S. Islam, M. A. Al Numanbakth, Z. Ahmed, and S. Akter, "The water quality and pollution sources assessment of Surma river, Bangladesh using, hydrochemical, multivariate statistical and water quality index methods," *Groundwater for sustainable development*, vol. 12, p. 100523, 2021.

[16] M. C. Smith, D. Fillis, M. Charpentier, R. McKinney, R. Johnson, and J. Heltshe, "Development and validation of rapid assessment indices of condition for coastal tidal wetlands in southern New England, USA Cathleen Wigand, Bruce Carlisle, Jan," *Environ Monit Assess*, vol. 182, pp. 31-46, 2011.

[17] L. A. Staszak and A. R. Armitage, "Evaluating salt marsh restoration success with an index of ecosystem integrity," *Journal of Coastal Research*, vol. 29, no. 2, pp. 410-418, 2013.

[18] Z. M. Ali, N. A. Ibrahim, K. Mengersen, M. Shitan, and H. Juahir, "Robust principal component analysis in water quality index development," in *AIP Conference Proceedings*, 2014, vol. 1602, no. 1, pp. 1091-1097.

[19] Department of Environment, "Malaysia Environmental Quality Report 2006," Ministry of Natural Resources and Environment Malaysia, 2006. [Online]. Available: https://enviro2.doe.gov.my/ekmc/wp-content/uploads/2016/08/1403318138-Environmental_Quality_Report_(EQR)_2006.pdf

[20] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 6th ed. Upper Saddle River, N. J.: Prentice Hall, 2007, p. 516.