

ORIGINAL ARTICLE

Analysis of Water Quality Parameters Using Environmetric Technique for Health Risk Assessment of Rivers Pollution Sources Pre COVID-19 Pandemic in Gebeng Industrial Area, Pahang, Malaysia

Siti Umi Kalthum Ab Wahab¹, Haikhal Fakhreez Hairuddin¹, Mohd Shukri Mohd Aris², Ahmed Jalal Khan Chowdhury³

¹ Department of Biotechnology, Kulliyah of Science, 017, 25200 Kuantan, Pahang, Malaysia

² Centre of Environmental Health & Safety, Faculty of Health Sciences, UiTM Puncak Alam Campus, Selangor, Malaysia

³ Department of Marine Science, Kulliyah of Science, International Islamic University Malaysia, 25200 Kuantan, Pahang, Malaysia

ABSTRACT

Introduction: Significance of rivers in Malaysia for drinking water, irrigation, aquaculture, fisheries, and tourism purpose is beyond argument. However, both natural processes and anthropogenic activities may result in the deprivation of river water quality. This state will spoil its use for different human purposes and also may affect the residents' health. **Methods:** A total of six sampling stations from both Balok and Tunggak Rivers were selected on a location basis and physicochemical parameters reading were measured monthly pre-COVID-19 Pandemic spread in Malaysia from September 2019 to February 2020. Later, the data was matched to the Malaysian National Water Quality Standard (MNWQS). Cluster and Principal Component Analysis (CA and PCA) were applied to determine the variations of pollution sources. **Results:** Temperature and pH were found to be within acceptable value throughout the period, while Dissolved Oxygen showed an abysmal reading for sampling point T1 and thus was classified as class V. CA grouped the six sampling stations into three clusters which are T1 was considered as the most polluted area. PCA yield only two significant components that represented 68.66% of the total variability. The first component (35.83%) has higher loadings on pH, specific conductivity, and salinity, while the second component (32.83%) has higher loadings on temperature and dissolved oxygen. **Conclusion:** The results from this study can provide significant information regarding the current status of rivers water quality in the Gebeng area, and the method proposed also can be used as a strategy for river water monitoring and ecological studies in the future.

Keywords: Water quality, Health Risk Assessment, Covid-19 pandemic, Environmetric Technique, Rivers

Corresponding Author:

Mohd Shukri Mohd Aris, PhD
Email: myshukri@uitm.edu.my
Tel: +603-32584525

INTRODUCTION

Gebeng Industrial Estate is a thriving hub for the nation's petrochemical industry. However, the surge in its industrial activities means more effluents are being produced. The effluents being discharged could be very detrimental to the environment especially the rivers nearby the area (1). Balok and Tunggak rivers are located in the vicinity of the Gebeng Industrial Estate. Once considered as the potential rivers for brackish

water fish sanctuary in Kuantan, both rivers since then have been tainted primarily by industrial activities from factories nearby especially before Covid-19 Lockdown in Malaysia, Movement Control Order as cited elsewhere (2, 3, 4), as both of this rivers became the easy route for effluents disposal.

In general, the river is a crucial source of water that is used for agricultural, domestic, and industrial needs (4, 5, 6). Maintaining good water quality is essential for many reasons, from improving human well-being by providing a clean water supply, to ensuring a healthy environment for aquatic life to thrive (7). Thus, ensuring that this valuable resource is well managed is necessary to continue enjoying the benefits that it provides.

In this case, constant water quality monitoring of Balok and Tunggak River is dire in ensuring the sustainability of aquatic organisms living in its water bodies. If not, humans will eat fish that catch from both of the rivers and use the polluted river water as alternative domestic water supply for agriculture planting without any treatments and this will affect the long-term resident's health (8) mentioned that water pollution ascends when unwanted substances flow into the water, and then the water physicochemical properties undergo changes. For instance, the abundance of pathogenic microbes, high level of ammonia nitrogen, inorganic and organic nitrogen compounds contained in wastewater of livestock from agricultural activities (9). Besides, other pollutants such as heavy metals, wastewater, crude oil, and plastic may affect the aquatic ecosystem (10). The condition of this polluted water will pose negative effects to humans, health, and the environment. One health may be jeopardized as exposure to heavy metals, and radioactive elements may cause cancer (11) and infectious pathogens in river water may result in the spreading of water-related diseases (12). Besides, the growing fear of emergence and resurface of other vector-borne diseases in Malaysia such as chikungunya, malaria, and yellow fever (13) also can be transmitted by the mosquito through the stagnant river's water.

To overcome this problem, a way to easier identify the most pollutant sources in both rivers is needed. Environmetrics techniques have been used to estimate and characterize environmental quality information such as water, land, air and it is important to notice spatial variations caused by natural or environmental factors (14, 15). Recently, environmetric techniques have become a momentous device in environmental health to evaluate and expose complex relationships of information in a varied discipline of community and environmental applications.

In this research, water quality physicochemical parameters were assessed based on the National Water Quality Standard (NWQS) adopted by the Department of Environment (DOE), Malaysia to determine the river status pre-Covid-19 Pandemic based on each water quality parameter class and further analysis by using environmetric techniques such as cluster and principal component analysis (CA and PCA) to pinpoint the source of contamination and find out the correlation between the parameters. This study focuses on two rivers located in Gebeng industrial area which are Balok and Tunggak Rivers.

METHODS AND MATERIALS

Sampling Area of Balok and Tunggak Rivers

The Balok and Tunggak rivers were selected due to their characteristics and geographical locations that are situated right next to an industrial area, in which all

discharges from factories will enter the river and flow towards the South China Sea (16). The site selections were chosen also justified based on previous studies (2,3,4) found that nearby industries activities such as waste residues from petrochemical effluent and local agriculture discharge gave initial pollution effects to both rivers. Thus, they provided the perfect landscape to study both natural processes and anthropogenic impacts on river water quality and its subsequent effects on aquatic organisms and also human health.

There were six sampling stations in total, namely B1, B2, B3, T1, T2, and T3. Three of them were located at Balok River and three were at Tunggak River as shown in Fig. 1. The sampling points for both rivers were chosen on a location basis, classified into three parts, which are the upstream, middle stream, and downstream (6). The map of the study area was prepared using QGIS 3.0 software. Sampling activities were carried out once a month, for six months.

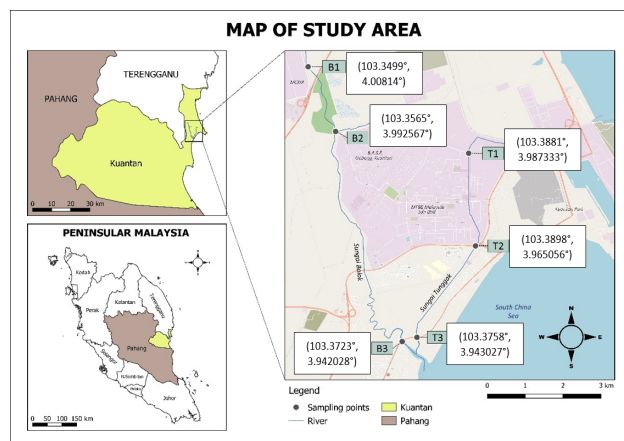


FIG. 1: Six sampling stations in total, namely B1, B2, B3, T1, T2, and T3. Three of them were located at Balok River and three were at Tunggak River

Water Quality Parameters Reading

The physicochemical parameters that were examined were temperature, turbidity, salinity, pH, dissolved oxygen (DO), and specific conductivity (SC). All readings were taken triplicate using Hydrolab Quanta Multi-Probe Meter (OTT Hydromet).

Environmetric Analysis

Environmetric analysis such as HCA and PCA were used to analyze water quality parameters (17) using IBM SPSS Statistics 26. For each sampling station, the mean was calculated to give the average result.

Cluster Analysis

Hierarchical cluster analysis (HCA) was used on the mean data of the parameters of all stations to determine their similarities. HCA was applied using Ward's method with squared Euclidean distances as the measure of similarity (18, 19, 20). All analyzed data were uniform to z scores to ensure normal distributions for the HCA.

Principal Component Analysis

Principal component analysis (PCA) was used in this study to trace the prominent pollutant sources in both rivers. Here, the full information was extracted instead of the mean to maximize the variability and the data size to yield a better result. Essentially, the PCA techniques had been applied to a wide range of environmental applications and this analysis was revered to be the paramount approach to periphraze the misinterpretation of large amounts of complex environmental monitoring data (4, 21). The data was standardized by using the correlation matrix. Varimax rotation was applied for a better interpretation of the data (22) since it can be extended to PCA in order to simplify the structure of the problem and to facilitate the interpretation. Here, a simple solution means that each factor has a small number of large loadings and a large number of zero or small loadings (23). This simplifies the interpretation because after this rotation was used, each original variable will tend to be associated with one or a small number of factors, and each factor represents only a small number of variables. Besides, Kaiser-Meyer-Olkin (KMO) and Bartlett's tests score were calculated in this study to find the fittingness of PCA on the data (24). Normally for KMO test, high values that close to 1.0 indicate that a factor analysis may be useful with the data. However, if the value is less than 0.50 which is the minimum value accounted, the results of the factor analysis for the study probably won't be very useful. For Bartlett's test, small values that less than 0.05 of the significance level indicate that a factor analysis may be useful with the data.

RESULTS

Classification of Water Quality Parameters (NWQS)

Fig. 2 showed the average reading of six physicochemical water quality parameters in Balok and Tunggak Rivers. The average temperature (T) for the whole six-month sampling period is shown in Fig. 2(a). All stations exhibited normal temperature (normal temperature range: 20°C - 30°C), with the lowest average was shown at T1 at 27.10°C. Referring to the National Water Quality Standards (NWQS) for Malaysia, the temperature for both rivers fell underclass IIA and III. However, both rivers showed little variations, with the highest standard deviation, 1.36 (Table I) was shown at B3. Even though the water temperature fluctuates between day/night and varies along the length of a river with latitude and elevation – human activities also affecting water temperature due to discharge of cooling water or heated industrial effluents.

Moreover, Fig. 2(b) represents the average turbidity from six sampling activities done for each sampling point from September 2019 until February 2020. Both rivers showed a very high turbidity value, with B1 recorded the highest value among all at 107.02 NTU, and went

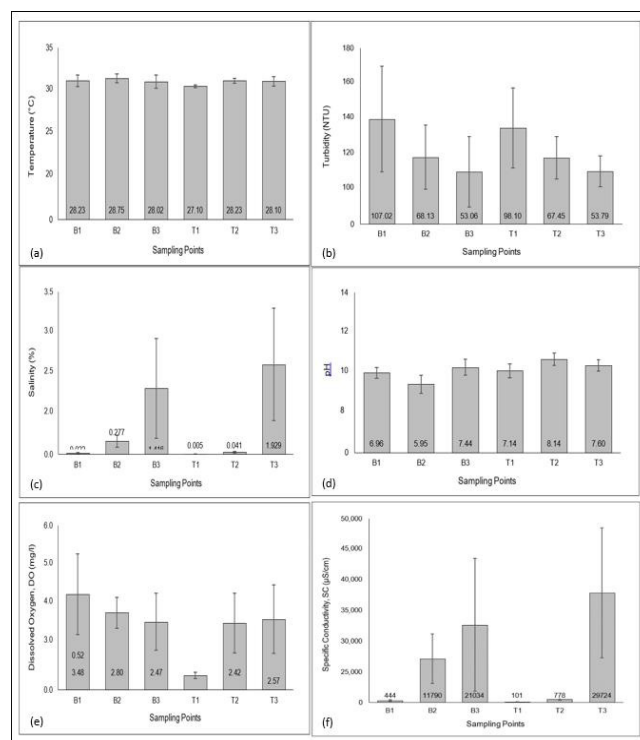


Fig. 2: The average reading of six physicochemical water quality parameters in Balok and Tunggak Rivers

beyond twice as much as the 50 NTU limit of class IIA and IIB stated in Malaysian standard. Overall, both rivers exceeded the limit set by the standard, although it was important to note that the standard variations (Table I) for each sampling site were quite high.

Furthermore, Fig. 2(c) shows the average salinity from six sampling activities done for each sampling point from September 2019 to February 2020. The same pattern was observed in terms of salinity values for both rivers, where B3 and T3, the downstream part of Balok River and Tunggak River respectively, exhibited a relatively high value when compared to the upstream and middle stream. Both B3 and T3 values, however, fell within the 1% to 2% range (lower is better), which classified them as class IV as per Malaysian NWQS. The remaining sampling sites were classified as class I.

The average pH for the six-month sampling period from September 2019 to February 2020 is shown in Fig. 2(d). Only B2 recorded a slightly acidic pH value (pH 5.95) while the others were within the normal pH range. The NWQS for Malaysia classified pH 6 – 9 as class IIA and IIB, and pH 5 - 9 as class III and IV (the closer to pH 7 the better). Thus, this classified B2 as class III and IV, whereas the other stations were within class IIA and IIB. While the average Dissolved Oxygen (DO) at each sampling site from September 2019 to February 2020 is shown in Fig. 2(e). Overall, all stations exhibited low DO values, especially T1 with 0.52 mg/l. When compared to the Malaysian standard, only point B1 managed to get an average DO above 3 (higher is better), which

Table I: Mean, minimum, and maximum values and standard deviation of water quality parameters of Balok (B1-B3) and Tunggak rivers (T1-T3).

Sam-pling point	Sta-tis-tical tools	T (°C)	Tur-bidity (NTU)	Salin-ity (%)	pH	DO (mg/l)	SC (µS/cm)
B1	Mean	28.23	107.02	0.022	6.96	3.48	444
	Std. Dev.	1.17	54.08	0.011	0.47	1.47	225
	Min	26.63	40.40	0.007	6.51	1.85	142
	Max	29.46	203.00	0.036	7.80	5.81	745
B2	Mean	28.75	68.13	0.277	5.95	2.80	11790
	Std. Dev.	0.89	32.62	0.130	0.77	0.57	6781
	Min	27.56	25.60	0.145	5.19	2.01	2803
	Max	29.61	107.43	0.479	7.07	3.51	19767
B3	Mean	28.02	53.06	1.416	7.44	2.47	21034
	Std. Dev.	1.36	35.83	1.077	0.70	1.04	18071
	Min	26.15	33.93	0.335	6.82	1.85	710
	Max	29.56	125.67	3.458	8.37	4.57	52167
T1	Mean	27.10	98.10	0.005	7.14	0.52	101
	Std. Dev.	0.26	40.86	0.001	0.62	0.12	17
	Min	26.66	50.53	0.004	6.22	0.36	69
	Max	27.48	169.00	0.006	8.08	0.64	120
T2	Mean	28.23	67.45	0.041	8.14	2.42	778
	Std. Dev.	0.54	21.50	0.014	0.54	1.09	291
	Min	27.33	40.57	0.020	7.55	0.37	412
	Max	28.80	90.37	0.060	9.12	3.42	1214
T3	Mean	28.10	53.79	1.929	7.60	2.57	29724
	Std. Dev.	1.01	15.78	1.210	0.48	1.25	17661
	Min	26.81	33.83	0.300	6.99	1.19	5523
	Max	29.29	75.57	3.437	8.22	4.25	51967

Note: These data are six-month mean values from September 2019 to February 2020. The full readings of each month can be viewed in Appendix B.

put it in class III. In contrast, the abysmal DO reading of T1 has classified it as class V. The rest, fell within class IV. Nonetheless, the overall spread of DO value for 6 months was quite big (excluding T1, which further showed that it was highly polluted), as was illustrated in the long standard deviation error bar.

Besides, Fig. 2(f) shows the average specific conductivity (SC) from six sampling activities done for each sampling point from September 2019 to February 2020. Based on Malaysian NWQS, the acceptable limit for SC under class IV is 6000 µS/cm. Referring to Figure 4.7, B2, B3, and T3 recorded SC values exceeding the limit set, while B1, T1, and T2 recorded low SC values of below 1000 µS/cm, thus classifying them underclass I or IIA. Overall, Balok River appeared to have a higher

SC value than that of Tunggak River and the variation of readings between sampling times was very significant, with higher standard deviations were shown for B2, B3, and T3.

Cluster Analysis

Hierarchical cluster analysis (HCA) was functional to group or cluster sampling stations according to their similarity based on the water quality parameters. Fig. 3 illustrates a dendrogram, combination all six different stations into three statistically meaningful groups. Cluster 1 includes B3 and T3, Cluster 2 includes B1, B2, and T2, and Cluster 3 includes T1. Inferred from the physicochemical parameters data, Cluster 1, Cluster 2, and Cluster 3 correspond to less polluted (LP), moderately polluted (MP), and highly polluted (HP), respectively (22).

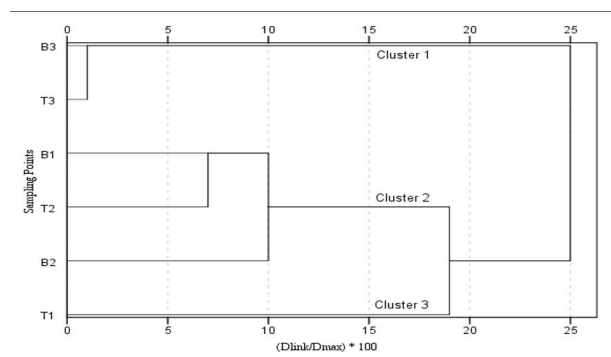


Fig. 3: A dendrogram, combination all six different stations into three statistically meaningful group

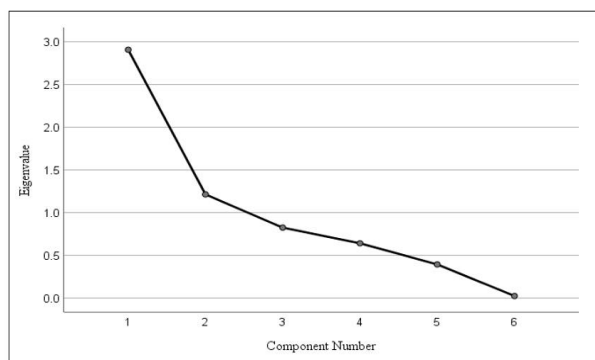
Principal Component Analysis

Principal component analysis (PCA) was used on the water quality information to find out which parameters were the most significant in contributing to the variations of the water quality (24) in both rivers, by reducing data but without losing the whole meaning as well as it could, and to summarize the statistical correlation between those parameters (19, 26). Two principal components with eigenvalues larger than one were obtained from the scree plot (Fig. 4), summing 68.66% of the total variance in the information. Corresponding components, variable loadings, and variance are presented in Table II. From Table III, the KMO measure of sampling adequacy test showed a score of 0.561, which is higher than the minimum value of 0.50 used by past literature (27, 28). The Bartlett’s test of sphericity showed a score of 0.007 significant level, which is less than the minimum value of 0.05. The difference in variability between both principal components was very small, about 3% only of variability difference. Overall, Component 1 explained 35.83% of the total variance and had robust positive loadings on pH, salinity, and specific conductivity meanwhile, Component 2 explained 32.83% of the total variance and had robust positive loadings on dissolved oxygen and temperature.

Table II: Total variance explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotated Sums of Squared Loadings		
	Total	Variance %	Cumulative %	Total	Variance %	Cumulative %	Total	Variance %	Cumulative %
1	2.907	48.447	48.447	2.907	48.447	48.447	2.149	35.825	35.825
2	1.212	20.207	68.655	1.212	20.207	68.655	1.970	32.830	68.655
3	0.824	13.730	82.384						
4	0.640	10.675	93.059						
5	0.393	6.550	99.610						
6	0.023	0.390	100.000						

Note: Eigenvalues > 1 is considered as significant. Extraction method: principal component analysis. Rotation method: Varimax with Kaiser normalization.



Note: Component with eigenvalues larger than one were considered as significant and selected for rotation.

Fig 4: Scree plot from principal component analysis

Table III: KMO test, rotated sums of squared loadings, and rotated component matrix of water quality parameters data extracted using principal component analysis (PCA).

Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy = 0.561		
Parameters	Component	
	1	2
Dissolved Oxygen	0.043	0.830
pH	0.738	-0.249
Salinity	0.774	0.455
Specific Conductivity	0.733	0.529
Temperature	0.167	0.839
Turbidity	-0.662	-0.167
Eigenvalue	2.149	1.970
Variability (%)	35.825	32.830
Cumulative (%)	35.825	68.655

Note: Items highlighted were found to be significant variables influencing principal component analysis. Only the first two components with eigenvalues > 1 were shown. Extraction method: PCA. Rotation method: Varimax with Kaiser Normalization. Rotation converged in 3 iterations.

DISCUSSION

Classification of Water Quality Parameters (NWQS)

The physicochemical parameters of a water body can determine its quality. For this research, there were six parameters monitored, which were temperature, turbidity, salinity, pH, dissolved oxygen (DO), and specific conductivity (SC). Since this research was focusing on field work or in-situ study only due to some limitation, therefore only selected water quality parameters were measured by using Hydrolab Quanta Multi-Probe Meter (OTT Hydromet). Hence, to know the “healthy” value of each individual parameter and its water class for both rivers, the National Water Quality Standards (NWQS) for Malaysia was referred instead of the whole rivers pollution status calculation using Water Quality Index (WQI). Each value was then compared and classified accordingly.

The temperature for both rivers fell underclass IIA and III because the sampling activities were carried out during sunny and rainy conditions. Since B3 is located downstream of Balok River, particularly near an estuarine region, the temperature from the ocean might have come into play. The high river water temperature also caused by sunlight’s intensity entering the water body without any obstructing plants in the rivers (29) since this study was carried out during day time.

Basically, temperature does not play a major role in the calculation of Water Quality Index (WQI) since it’s only focus on the six mandatory parameters only, but temperature also has its vital and essential value in river water quality status. It is act as a major factor for the river ecosystem since it can give a great influence on the health and life of an aquatic organism such as fish (29) which is one of the food supply for residents nearby the Balok and Tunggak Rivers. For example, the raised temperatures in river water can alter the biodiversity of river ecosystem by two ways, either by intolerable for aquatic biota or increase the microbial growth in

the rivers, which resulted in the decrease of dissolved oxygen level. A small increase in the temperature of a clean, fast-flowing stream may not affect the ecosystem adversely. However, if the heated effluents are usually discharged to river water that are already polluted to some degree such as in Balok and Tunggak River (2, 4, 11) plus its natural occurrence, so the polluting effects are compounded that may cause the aquatic organism to die. Therefore, this is one of the important parameters that need to be taken into consideration since both rivers were located in a swarming industrial area.

While, the turbidity of both Balok and Tunggak River were affected by some temporal factors, and not constantly appeared murky as the reading showed higher than the standard. However, it is important to note that for both rivers, the upstream part presented the highest turbidity. Location-wise, both B1 and T1 were located the nearest to Gebeng Industrial Area. This implied that the high turbidity values were likely to be caused by some anthropogenic activities. To be precise, there was a sign of land excavation for piping purposes done near B1 by an adjacent factory, predominantly on December 3, 2019 sampling time as evident by the high turbidity value recorded, which was 203 NTU. The pipes were likely installed to redirect an outflow from the factory. As a result of this, the soil being excavated was easily eroded, and the effect of rainfall, combined with the excavation process itself, caused a heavy amount of soil to flow into the river.

Besides, a high level of salinity reading indicates a high level of salt. According to (30), anthropogenic impacts such as irrigation, deicing road salt, domestic and industrial wastewater could affect salinity in water. However, a particular trend displayed in Fig. 2(c) suggested that the location of the sites played a vast role. Since both B3 and T3 were in the downstream part of both rivers, the high salinity values might be hugely influenced by the high salinity water of the estuarine or coastal system. The sampling point B2 showed a slightly acidic condition possibly due to its location, as it was located right next to a factory. Thus, the effluents that were discharged might have been affecting the pH of the river water near B2. Besides, the slight pH shift shown could also be due to many other reasons. According to (31), photosynthetic activity, microbial respiration, and decomposition all could be the factor affecting the pH value of water, especially when it comes to the river.

Moreover, the difference in dissolved oxygen (DO) showed by T1 and the rest of the sites can be attributed to the fact that the river water at T1 was stagnant and a clear sign of oil contamination can be seen. The low value of dissolved oxygen saw at T1 (0.52 mg/l) suggested that oxygen was being consumed by microorganisms faster than it was produced by plants from photosynthesis (32). It was perhaps due to the plants not receiving enough sunlight for photosynthesis to work, as evident from the

water turbidity. From turbidity data in Fig. 2(b), T1 has an especially high value of 98.10 NTU, showing that the accumulation of oil in the surface water has blocked sunlight from penetrating and reaching most plants, thus producing less oxygen.

Other than that, the highest SC value was recorded in the downstream part of both rivers. The same pattern can be observed in the salinity data in Fig. 2(c), implying that salinity and specific conductivity were related. As per prior discussion, the inflow of the water from the South China Sea has affected the readings, as the ocean and estuarine waters were of high salinity. High salinity can indirectly affect SC as more ions responsible for salination such as sodium, chlorine, potassium, and calcium were available to the water, and SC readings were determined by the ions contained in the water. Nonetheless, other anthropogenic factors should be considered as well. (11) reported in their research on similar sites that, ionic pollutants from industries, residential wastewater, and urban runoff might have some contribution to the high SC value. If the data for each sampling time is referred to, the third sampling activity is done on December 3, 2019, showed the lowest SC values for all B2, B3, and T3, thus proposing that there were smaller disturbances that occurred from the ocean during that time.

Cluster Analysis

Overall, cluster analysis assembled the sites based on the resemblance of readings of the water quality parameters. Based on the clusters formed, the readings were largely influenced by spatial variation among stations. B3 and T3 are both located at the downstream part of Balok and Tunggak rivers, B1, B2, and T2 roughly represent the middle stream, and T1 is on the upstream.

T1 was in the class (or in this case, cluster) of its own. Arguably, not only T1 showed the most undesirable water quality from the data collected, but also manifested it through the observations. Hence, it was classified as HP. T1 suffered heavily from oil pollution, which could be due to anthropogenic activities such as the disposal of industrial effluents, and the discharge from vehicle washing and workshops.

B1, B2, and T3 of Cluster 2 correspond to the MP classification. These locations were comparatively less polluted than T1, although, some of their water quality parameters still fell within the higher classes of Malaysian NWQS (note that lower is considered better). All three stations were connected to the industrial area, and away from domestic influence. Thus, most of the factors contributing to the pollution were the result of industrial activities such as effluents discharge, and human activities such as river banks excavation for piping purposes (11).

Cluster 1 comprises B3 and T3 and corresponds to LP sites. Although parameters such as salinity and specific

conductivity predominantly showed high values on these stations, this was largely due to the influx of seawater coming in and mixed with the river water. This indicates that the factors that contributed to the pollution in the downstream stations were mainly due to natural activities, and less influence from anthropogenic activities.

Principal Component Analysis

A low KMO test score in this study indicates that the result shown by PCA might not be as useful if we compared it to the past study (27, 28). However, according to IBM Knowledge Center, a score of higher than 0.5 for KMO test and a small values less than 0.05 of the significant level of Bartlett's test can still be considered as barely appropriate for PCA. On this basis, the result from PCA shown can still be used as a baseline to find out which parameters have the most significant contribution to the variations of water quality for the six-month sampling period and to see the correlations between all parameters. The small difference in variability between both principal components indicates that both components carried almost similar weight in contributing to the water quality variations for the whole six-month sampling period.

The pH, salinity, and specific conductivity parameters showed a strong correlation among themselves in Component 1. Briefly, the specific conductivity and salinity are affected by the number of ions available in the water. pH, on the other hand, is the measure of H⁺ ions concentration in the water. Thus, a high concentration of H⁺ ions would affect both pH and specific conductivity (33).

Theoretically, temperature and dissolved oxygen have an inverse correlation, where an increase in temperature would reduce the amount of dissolved oxygen available (34, 35). However, Table III exhibits a positive relationship among them as shown in Component 2. The reason for this was due to the little variation shown in the data, especially for temperature. Temperature for both rivers stayed relatively the same throughout the sampling period, thus an inference made from the temperature data on other parameters might be somewhat unreliable. To see significant changes temperature can have on other parameters, sampling can be done throughout different seasons (19).

CONCLUSION

The assessment of water quality was done based on the NWQS for Malaysia. All in all, temperature and pH were found to be within an acceptable value; dissolved oxygen showed an abysmal reading for T1 and thus was classified as class V, while the others were at acceptable levels; turbidity showed high values especially in the upstream; whereas salinity and specific conductivity showed high

values especially in the downstream. Cluster analysis and principal component analysis were applied to evaluate the variation shown by the water quality parameters. Cluster analysis assembled six sampling sites into three clusters based on the resemblance in water quality, in which; T1 was considered as the most polluted; B1, B2, and T2 as moderately polluted; B3 and T3 as less polluted. Principal component analysis yield only two significant components that represented 68.66% of the total variability. The first component (35.83%) has robust loadings on pH, specific conductivity, and salinity, while the second component (32.83%) has robust loadings on temperature and dissolved oxygen. The method that is discussed can serve as a potential tool that can benefit river monitoring programs in the future.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Ministry of Education (MOE) and Research Management Centre, International Islamic University Malaysia (IIUM) for funding this project through a grant (FRGS/1/2018/STG03/UIAM/01/1). The authors also indebted to Kulliyah of Science, IIUM for providing laboratory facilities along with laboratory staff for supporting this research project.

REFERENCES

1. Sujaul I, Hossain M, Nasly MA, Sobahan MA. Effect of industrial pollution on the spatial variation of surface water quality. *Journal of Environmental Science*. 2013; 9: 120-129.
2. Salah MG, Sujaul MI, Arafat MY, Abdulsyukor R, Idris A. Surface water quality assessment of the Gebeng industrial area using water quality standard and index. *Journal of Engineering and Science Research*. 2017; 1 (2): 117-125.
3. Wahab SUKA, Yunus K, Samah MAA, Aris MSM, Sabuti AA. Study of Relationship between Water Quality Parameters, Selective Heavy Metals, and Radioactive Elements Content in Rivers at Gebeng, Kuantan, Pahang, Malaysia. *Asian Journal of Chemistry*. 2019; 31: 433-437.
4. Wahab SUKA, Samah MAA, Sabuti AA, et al. Environmental Forensic Study: Tracing of Pollution Sources Using Environmetric Technique in Balok and Tunggak Rivers near Gebeng Industrial Area, Kuantan, Pahang, Malaysia. *Desalination and Water Treatment*. 2020; 191: 118-125.
5. Yeleliere E, Cobbina SJ, Duwiejuah AB. Review of Ghana's water resources: the quality and management with particular focus on freshwater resources. *Journal of Applied Water Science*. 2018; 8(93): 1-12.
6. Rahim AHA, Kasmuri N. Assessment of water quality index and heavy metals in Sungai Bunus,

- Malaysia. *Journal of Physics*. 2020; 1529: 1-10.
7. Wahab SUKA, Sabuti AA, Samah MAA, Yunus K. An Overview of Radioisotopes Study in Water Pollution. *International Journal of Engineering and Technology*. 2018; 7(2.29): 882-886.
 8. Haseena M, Malik MF, Javed A, et al. Water pollution and human health. *Environmental Risk Assessment and Remediation Journal*. 2017; 1(3): 16-19.
 9. Gasim MB, Jamil MM, Rahim SA, Toriman ME. Water-quality assessment of the Langat River at kilometre 7, Jalan Kajang-Bangi, Selangor, Malaysia. *The Arab World Geographer*. 2009; 12(3-4): 188–198.
 10. Hader DP, Banaszak AT, Villafane VE, Narvarte MA, Gonzalez RA, Helbling EW. Anthropogenic pollution of aquatic ecosystems: emerging problems with global implications. *Science Total Environment*. 2020; 713:136586.
 11. Wahab SUKA, Shaibullah SH, Samah MAA, Aris MSM. An Assessment of Surface Water Quality and Heavy Metals Involving the Radioactive Elements in Sungai Tunggak and Sungai Balok, Gebeng, Kuantan, Pahang: Comparison between The Year 2014 and 2015. *International Journal of Applied Chemistry*. 2016; 12: 146-151.
 12. Sherchand JB. Future emerging issues in waterborne diseases and microbial agents. *Journal of Institute of Medicine*. 2012; 34(3): 1-3.
 13. Hamid AAA, Isa MLM, Pien LS, Hassan NFN, Muhammad H, Hamid HA. Exposure Risk of Household Insecticide: Identification of the Knowledge, Attitude and Practice Levels among Kuantan Population, Pahang. *Malaysian Journal of Medicine and Health Sciences*. 2021; 17(1): 23-32.
 14. Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res*. 2000; 34(3):807–816.
 15. Singh KP, Malik A, Sinha S. Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques: a case study. *Anal Chim Acta*. 2005; 538(1–2):355–374.
 16. Wahab SUKA, Shaibullah SH, Samah MAA, Aris MSM. An Assessment of Surface Water Quality and Heavy Metals Involving the Rare Earth Elements in Sungai Tunggak and Sungai Balok, Gebeng, Kuantan, Pahang. *Journal of Malaysian Critical Metals*. 2016; 1: 11-23.
 17. Zhang Y, Guo F, Meng W, Wang X. Water Quality Assessment and Source Identification of Daliao River Basin Using Multivariate Statistical Methods. *Environmental Monitoring Assessment Journal*. 2009; 152: 105-121.
 18. Salam MA, Kabir MM, Yee LF, Khan MS. Water Quality Assessment of Perak River, Malaysia. *Pollution*. 2019; 5(3): 637-648.
 19. Pejman AH, Bidhendi GN, Karbassi AR, Mehrdadi N, Bidhendi ME. Evaluation of spatial and seasonal variations in surface water quality using multivariate statistical techniques. *International Journal of Environmental Science & Technology*. 2009; 6(3): 467-476.
 20. Taoufik G, Khouni I, Ghrabi A. Assessment of Physico-chemical and Microbiological Surface Water Quality Using Multivariate Statistical Techniques: A Case Study of the Wadi El-Bey River, Tunisia. *Arabian Journal of Geosciences*. 2017; 10 (181):1-19.
 21. Ouyang Y, Nkedi-Kizza P, Wu QT, Shinde D, Huang CH. Assessment of seasonal variations in surface water quality. *Water Res*. 2006; 40:3800–3810.
 22. Acal C, Aguilera AM, Escabias M. New modelling approaches based on varimax rotation of functional principal components. *Mathematics*. 2020; 8(2085): 1-15.
 23. Abdi H. Factor rotations in factor analyses. In: Lewis-Beck M, Bryman A, Futing T (Eds.). *Encyclopedia of Social Sciences Research Methods*. Thousand Oaks (CA). Sage. 2003; 1-4.
 24. Zheng L, Yu H, Wang Q. Application of multivariate statistical techniques in the assessment of surface water quality in Second Songhua River basin, China. *Journal of Central South University*. 2016; 23: 1040–1051.
 25. Kamble SR, Vijay R. Assessment of water quality using cluster analysis in the coastal region of Mumbai, India. *Environmental monitoring and assessment*. 2011; 178(1-4): 321-332.
 26. Syazwan A, Rafee BM, Juahir H, et al. Analysis of indoor air pollutants checklist using environmetric technique for health risk assessment of sick building complaint in nonindustrial workplace. *Drug, Healthcare and Patient Safety*. 2012; 4: 107–126.
 27. Olawale F, Garwe D. Obstacles to the growth of new SMEs in South Africa: A principal component analysis approach. *African Journal of Business management*. 2010; 4(5): 729-738.
 28. Parinet B, Lhote A, Legube B. Principal component analysis: an appropriate tool for water quality evaluation and management—application to a tropical lake system. *Ecological modelling*. 2004; 178(3-4): 295-311.
 29. Indriyani K, Hasibuan HS, Gozan M. Social conditions in wastewater processing to manage river water quality (study in Cirarab River, Tangerang District, Indonesia). *Earth and Environmental Science*. 2021; 623: 1-6.
 30. Florke M, Barlund I, Van Vliet MT, Bouwman AF, Wada Y. Analysing trade-offs between SDGs related to water quality using salinity as a marker. *Current Opinion in Environmental Sustainability*. 2019; 36: 96-104.
 31. Yap CK, Chee MW, Shamarina S, Edward FB, Chew

- W, Tan SG. Assessment of surface water quality in the Malaysian coastal waters by using multivariate analyses. *Sains Malaysiana*. 2011; 40(10): 1053–1064.
32. Eljaiek-Urzola M, Romero-Sierra N, Segre-Cabarcas L, Valdelamar-Martinez D, Quicones-Bolacos J. Oil and Grease as a Water Quality Index Parameter for the Conservation of Marine Biota. *Water*. 2019; 11(4): 856.
 33. Thomas M, Florion A, Chrétien D. The effect of pH and conductivity on the electric behaviour of *Apteronotus albifrons* (family apteronotidae) in a pollution monitoring system. *Environmental technology*. 1997; 18(11): 1069-1083.
 34. Said A, Stevens DK, Sehlke G. An innovative index for evaluating water quality in streams. *Environmental Management*. 2004; 34(3): 406–414.
 35. Verma AK, Singh TN. Prediction of water quality from simple field parameters. *Environmental earth sciences*. 2013; 69(3): 821-829.