Jurnal Kejuruteraan SI 5(2) 2022: 23-34
https://doi.org/10.17576/jkukm-2022-si5(2)-03

Spatial and Temporal Assessment of Marine Water Quality using Statistical Approaches

(Penilaian Kualiti Air Marin Mengikut Masa dan Ruang Menggunakan Pendekatan Statistik)

Firdaus Mohamad Hamzah [a*,b] , Mohd Nabil Fikri[c] , Izamarlina Asshaari [a,b] , Mohd Saifullah Rusiman [d], Mohd Khairul Amri Kamarudin [e] & Shamsul Rijal Muhammad Sabri[f]

[a]Department of Engineering Education, Faculty of Engineering & Built Environment, Universiti Kebangsaan, Malaysia
[b]Centre for Engineering Education Research, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia
[c]Department of Civil Engineering Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Malaysia
[d]Faculty of Applied Sciences and Technology, Universiti Tun Hussien Onn Malaysia
[e]Faculty of Applied Social Sciences, Universiti Sultan Zainal Abidin, Malaysia
[f]School of Mathematical Sciences, Universiti Sains Malaysia, Malaysia

**Corresponding author: fir@ukm.edu.my*

Received 15th June 2022, Received in revised form 22nd July 2022
Accepted 1st September 2022 , Available online 15th November 2022

ABSTRACT

The study was carried out to determine the classes and parameters that influence the water quality in the Straits of Johor using the Marine Water Quality Index (MWQI) method and the multi-variable statistical analysis, namely Principal Component Analysis (PCA). The MWQI method classifies the data sets according to four classes of contaminated, medium, good and very good classes. The data were studied from 2016 to 2017 according to two monsoon seasons - the northeast and southwest monsoons. The results for both seasons were compared to find out the difference. The results of the PCA analysis show the relationships between the parameters studied and determine which parameters are responsible for changing water quality standards. The results of the MWQI method reveals that all stations except WQ14 is in class 3 which is moderated while only WQ14 is in class 4 (contaminated). In terms of monsoon, the results of the analysis at the southwest monsoon station found that 25 out of 35 data were in class 3 while the other 10 were in class 4. In addition, MWQI analysis on the northeast monsoon found that 34 out of 35 data were in class 3 and only one was in class 4. These results show that the water quality in the northeast monsoon is better than in the southwest monsoon. The results from the PCA analysis indicates that fecal coliform, total suspended solids and phosphate influence the water quality at the southwest station while the dissolved oxygen and phosphate parameters influence the water quality at the northeast station due to high positive load values. In conclusion, MWQI was able to determine the class of water whilst PCA allowed the identification of types of parameters that affect the water quality. In conclusion, the two above methods can be used to determine pollution levels in water bodies.

Keywords: Marine Water quality; MWQI method; PCA analysis

ABSTRAK

Kajian ini dijalankan untuk menentukan kelas dan parameter-parameter yang mempengaruhi tahap kualiti air di kawasan Selat Johor mengunakan kaedah Indeks Kualiti Air Marin (IKAM) dan analisis statistik multi-pembolehubah iaitu Analisis Komponen Utama (PCA). Kaedah IKAM mengelaskan set data mengikut empat kelas iaiti tercemar, sederhana, baik dan terbaik. Data tersebut dikaji dari tahun 2016 hingga 2017 mengikut dua musim monsun iaitu monsun timur laut dan barat daya. Keputusan bagi kedua musim tersebut akan dibandingkan bagi memgetahui perbezaannnya. Hasil bagi analisis PCA menunjukkan hubungan antara parameter-parameter yang dikaji dan menentukan parameter manakah yang bertanggungjawab dalam perubahan piawaian kualiti air. Kaedah IKAM berdasarkan mean data mendapati kesemua stesen kajian kecuali WQ14 berada di kelas 3 iaitu sederhana manakala WQ14 berada di kelas 4 iaitu tercemar. Dari aspek monsun pula, 70 data dianalisis dan dibahagi kepada dua kumpulan iaitu timur laut dan barat daya. Keputusan analisis di stesen monsun barat daya mendapati 25 daripada 35 data berada di kelas 3 manakala 10 lagi berada di kelas 4. Selain itu, analisis IKAM pada monsun timur laut mendapati 34 daripada 35 data berada di kelas 3 dan hanya satu sahaja *berada di kelas 4. Hasil ini menunjukkan bahawa kualiti air pada monsun timur laut lebih baik berbanding monsun barat daya. Keputusan dari analisi PCA mendapati parameter koliform tinja, jumlah pepejal terampai dan fosfat mempengaruhi tahap kualiti air di stesen barat daya manakala parameter oksigen terlarut dan fosfat mempengaruhi kualiti air di stesen timur laut kerana mempunyai nilai pemberat positif tinggi. Secara kesimpulannya, kaedah IKAM dan PCA digunakan bagi menganalisis kompleks untuk kelas kualiti air bagi setiap stesen, parameter kualiti air dan menentukan parameter yang mempengaruhi tahap kualiti air. Akhir sekali, kedua-dua kaedah ini boleh digunakan bagi menentukan tahap pencemaran yang berlaku dalam perairan kita dan seterusnya mencari alternatif dalam meningkatkan kualiti air marin.*

Kata kunci: Kualiti air marin; Kaedah IKAM; Kaedah PCA

INTRODUCTION

Water is a natural resource. The origin of water has created many ecosystems that link between living things and non-living things that depend on one another. This study aims to determine the characteristics of the water supply system from a qualitative perspective. The unique physical properties of water have made water a universal solvent. The sea is important for its biodiversity. These include physical parameters (pH and dissolved oxygen), nutrients (ammonia and fostates) and metals (cadmium and ferment). However, the Department of Environment, Malaysia has used seven key parameters for assessing the quality of marine water, namely dissolved oxygen (DO), Coliform Tin (FC), oil and grease (OG), Nitrate (NH_3), suspended solids (TSS), Phosphate (PO_4) and ammonia (NO_3). All seven parameters were assigned a specific weight or value and then those values would be aggregated and represented by only one number, the Marine Water Quality Index (MWQI). This study was conducted with the help of computer software such as Microsoft Excel and Statistical Package for Social Science (SSPS) 22.0

Peninsular Malaysia lies between 1◦ and 7◦ north and 99◦ to 105◦ east, and covers an area of 131587 . It consists of highlands, flood plains and coastal zones. The Titiwangsa series forms the backbone of the Peninsula, from southern Thailand to the southeast, approximately 480 km and separates east from west (Subaila Jemain 2007.). Around the central highlands are the lowlands of the coast. The Peninsular Malaysia weather is hot and humid all year long with temperatures ranging from 21 ° C to 32 ° C, as are the features of moist tropical climates. The monsoon climate is characterized by two monsoon seasons associated with the Southwest monsoon from May to September and the Northeast Monsoon from October to March. Significant rainfall also occurred during the transition period (typically April and October) between monsoon seasons (Subaila Jemain 2007.). During the Northeast monsoon season, exposed areas such as the east coast of Peninsular Malaysia, West Sarawak and the northeastern coast of Sabah suffered heavy rainfall. In contrast, the interior or protected areas of the mountains are relatively free of its influence. It is best to describe the seasonal distribution of national rainfall.

The increase in human population has created a huge need for water resources, and a gradual decline in clean water resources will be a problem in the future (Mutlu 2019). This causes disruption of the hydrological cycle and a reduction in water resources (Rose et al. 2016). Deterioration of water quality has a negative impact on the lives of humans, animals, and plants in the short term, as well as many other negative effects in the long run (Uncumusaoğlu & Akkan 2017). Therefore, water quality management plays an important role in controlling water pollution (Sarkar & Pandey 2015). Population centers face escalating water difficulties such as water scarcity, water supply damage, and climate-related hazards as a result of fast urbanization. Population development and economic forces (Abshirini et al. 2016; Perera et al. 2018). Water comes from a variety of places, including the ocean, which makes up 97 percent of the earth's water, ice, which makes up 2% and snow which makes up 1% (Charles Ezugwua et al. 2021).

According to a study, water quality in Baram was found to be slightly contaminated where chemical oxygen demand and suspended solids are the most influential parameters of water quality (Ling et al. 2017). While the study by Martina Radjawane (2019) indicates that suspended solids volume, temperature, salinity and dissolved oxygen are the parameters that influence the water quality in Jakarta Bay. In addition, a study Teixeira de Souza et al. (2020) found that the Açude da Macela reservoir was classified as highly contaminated. All of these studies used the Water Quality Index (WQI) method and principal component analysis (PCA) to determine which classes and parameters affected the water quality in the study area.

Water Quality Index was used to determine the marine water quality class. Next, to compare the water quality level between southwest and northeast monsoons. Finally, to determine which parameters affect the water quality level.

STUDY AREA

The study was conducted in the Johor Strait located between southern Johor in Peninsular Malaysia and Singapore as in Figure 1. The Johor Strait has a width of six kilometers and has a depth of twenty-five meters. The main activity in the Straits of Johor is fishing and shipping activities (Wood et al. 1997). The coastline next to Malaysia and Singapore in the Straits of Johor is a popular tourist attraction. Add Johor for railways and highways built since 1924 to facilitate human and commercial goods. Rapid development between New Johor and Singapore has increased the level of water pollution throughout the Straits of Johor (Behera et al. 2013).

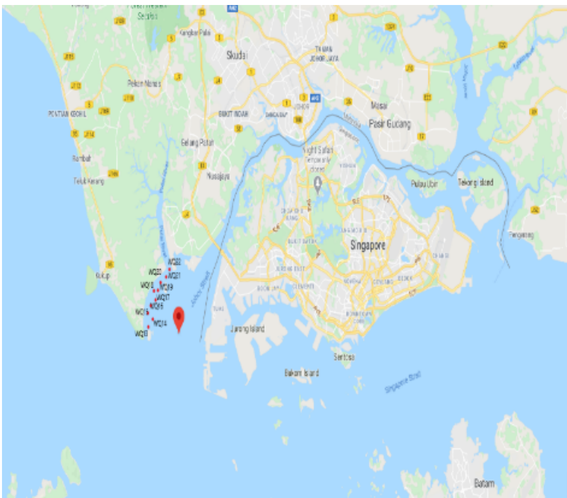

FIGURE 1. Study Area

METHODOLOGY

Water Quality Index (WQI is a numerical expression used to assess water quality to make it easier to understand (Oana 2010). To calculate WQI, in 1965 Horton proposed the first formula that took into account all the parameters needed to determine water surface quality and that reflected the composite influence of different parameters important for water quality assessment and management (Liou et al. 2004;Tyagi & Sharma 2014). Furthermore, WQIs which is similar to the Horton index were also developed by the Brown group in 1970 (Dawood & Gobreal 2010) which was based heavily on individual parameters. Currently, many modifications have been considered for the WQI concept through various scientists and experts (Akkaraboyina 2018; Sharma & Meena 2018). WQIs is used for the first time to focus on possible physical-chemical changes this year regarding the quality of flowing water (House & Ellis 1987; House 1990).

The southwest monsoon usually begins in late May or early June and continues until the end of August. In general, wind during this period is mild and fluctuating with some periodic changes in observed wind flow patterns (Daryabor et al. 2014). The dominant wind flow for the southwest monsoon is generally at speeds below 7 m / s (Camerlengo & Demmler 1997). Batmmetry, coastline and the complex existence of large islands, such as Natuna Island are also important in determining local circulation patterns (Akhir & Jawchuen 2011; Tangang et al. 2011).

PCA revealed that pollution levels in the three zones are mainly influenced by natural resources (temperature and river discharge) and anthropogenic resources (industrial, municipal and agricultural) (Onojake et al. 2011). The application of different pattern recognition techniques to reduce the complexity of large data sets has been shown to provide better interpretation and understanding of water as well as quality data (Qadir et al. 2008; Mu 2019). Based on previous studies, PCA-based statistical model for coastal water quality data have developed from Cochin coast in southwest India, which explains the relationship between the various physicochemical variables that have been monitored and the environmental conditions affecting the water quality of the coast (Jayakumar & Siraz 1997) Salman & Abu Ruka'h 1999; Praus 2005; Iyer et al. 2003).

Water quality classification was performed using Microsoft Excel and multi-variable analysis using Statistical Package for Social Science (SPSS) 22.0 computer software. Figure 2 shows the flow chart for the methodology.

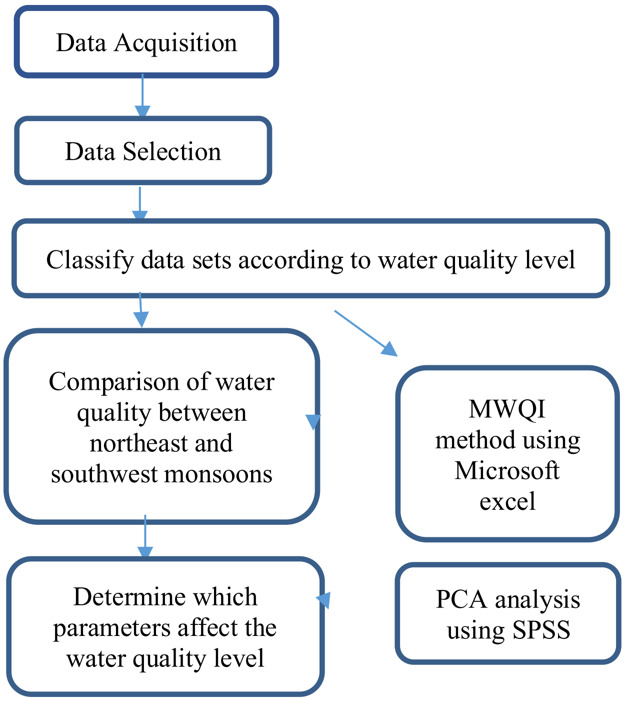

FIGURE 2. Methodology Flowchart

RESULTS AND DISCUSSION

TABLE 1. Station Coordinate

Station	Longitude (North)	Latitude (East)
WQ13	1.2786°N	103.5198°E
WQ14	1.2835°N	103.5266°E
WQ15	1.2886°N	103.5164°E
WQ16	1.2939°N	103.5193°E
WQ17	1.2981°N	103.5242°E
WQ18	1.3065°N	103.5203°E
WQ19	1.308°N	103.5297°E
WQ20	1.3189°N	103.5341°E
WQ21	1.3243°N	103.5462°E
WQ22	1.3306°N	103.5513°E

MARINE WATER QUALITY INDEX (MWQI)

In Malaysia, the standards used to control the quality of marine water are the Malaysian Department of Environment (DOE) Criteria and Standards for Marine Water. This standard is intended to protect aquatic life from extinction and to protect the marine environment from contamination. Department of Environment (DOE)

The most frequently cited local guidelines regarding water quality are the Environmental Quality Act (EQA), the National Water Quality Standards (NWQS) and the Malaysian Association of Water Quality (MWA) raw water quality criteria for consumption. The relevant parameter standards are given in Table 2 and Table 3 for reference. Class descriptions, in terms of categories and descriptions, are also provided in this section.

TABLE 2. Marine Water Quality Criteria and Standards

Parameters	Class 1	Class 2	Class 3	Class 4
Usage	Preservation, Marine Area protected, Marine Park	Marine Life, Fisheries,	Port, Oil & Gas Field	Mangrove Swamp & River estuary
Dissolved Oxygen (mg/L)	>80% saturated	5	3	4
TSS (mg/L)	25 mg / L or ≤ 10%	50mg / L (25 mg / L) or ≤ 10%	100 mg / L or ≤ 10%	100 mg / L or ≤ 30%
Oil and Grease (mg/L)	0.01	0.14	5	0.14
Ammonia (mg/L)	35	70	320	70
Nitrate (mg/L)	10	60	1,000	60
Phosphate (mg/L)	5	75	670	75
Fecal Coliform (mg/L)	70	100	200	100

TABLE 3. Classification of Marine Water Quality Index

Category	Index Value
Very Good	90-100
Good	80-<90
Moderate	50-<80
Contaminated	0-<50

Source: Jabatan Alam Sekitar, Bahagian Air dan Marin, Kementerian Sumber Asli dan Alam Sekitar, Tahun 2018-2019

MWQI is calculated using seven parameters which is DO, NO_{3-}, $PO_4{}^{3-}$, NH_3 fecal coliform, oil and grease also TSS. These parameters then summarized into; DO (0.2), NH_3 (0.16), FC (0.14), TSS (0.14), O&G (0.13), NH_3 (0.12), $PO_4{}^{3-}$ (0.11) where the figures in the bracket indicate the contribution of the parameter and at the same time reflect the importance of each parameter on the overall water quality. The formula for ADS is based on the weighted average Water Quality Index (WQI) formula:

$$WQI_M = SI\ DO^{0.2} \times SI\ NH_3{}^{0.16} \times SI\ FC^{0.14} \times SI\ TSS^{0.34} \times SA\ O\&G^{0.13} \times\ SI\ NO_3{}^{-0.12} \times SI\ PO_4^{3-0.11} \quad (1)$$

Source: Jabatan Alam Sekitar
Bahagian Air dan Marin
Kementerian Sumber Asli dan Alam Sekitar
Tahun 2018-2019

PRINCIPAL COMPONENT ANALYSIS (PCA)

Principal component analysis (PCA) was used to identify and classify each water quality parameter into an unrelated construct. The parameters that are found to overlap based on the relationships between constructs are derived from the advanced tests to produce one of their own constructs. The PCA test involves the phases of identifying factor correlations, extracting factors, and looking at the combined effects of factors to form distinct constructs. During the correlation identification process, only components with eigenvalues greater than 1 were subjected to extraction and rotation tests (Bengraïne & Marhaba 2003). Components with eigenvalues greater than 1 have clear variations in the study data. Sometimes the value of the weighting factor provided by the PCA is unclear and not available for interpretation, thus rotation is required according to the varimax rotation method (Juahir et al. 2011). In this study, only strong factor loadings (weight factor ≥0.7) were taken into account in interpreting results (Juahir et al. 2011). The PCA method is also a method of identifying potential factors affecting water quality status in the Straits of Johor. The basic concepts of factor analysis are represented by;

$$z_{ij} = a_{f1}f_{li} + a_{f2}f_{21} + \ldots + a_{fm}f_{mi} + e_{fi} \quad (2)$$

where, z is the measured variable value, a is the weighting factor, f is the score factor, e is the expected residual for error or variation from other sources, i is the number of samples, j is the number of variables and m is the sum of the number of factors. In this study, the PCA / FA used was from a set of pre-defined data sets (7 variables) for different sections of the study based on the results provided by the HCA. The matrix data input (variable × case) for PCA / FA was (7 × 10) covering all parts of the classification. 7 represent the 7 parameters while 10 represent 10 months, which is the time taken for data sampling. Northeast monsoon (November-Mac) and Southwest monsoon (May-September).

SCREE PLOT

The scree plot is a graph of eigenvalues. The vernacular definition of "scree" is a collection of loose rocks or rock fragments located on slopes or at the base of hills or cliffs. In a scree plot, it is advisable to find a sharp decrease in the size of the eigenvalues (such as a cliff), with the rest of the smaller eigenvalues being the ruins. As the eigenvalues drop dramatically, additional factors will add to the amount of information already taken. Because scree plots are subjective and arbitrary to interpret, their main utility is to provide two or three m values that are worth considering (Woods & Edwards 2011). The scree plot shows the variance explained for each data in descending order compared to the number of components (Liu et al. 2018).

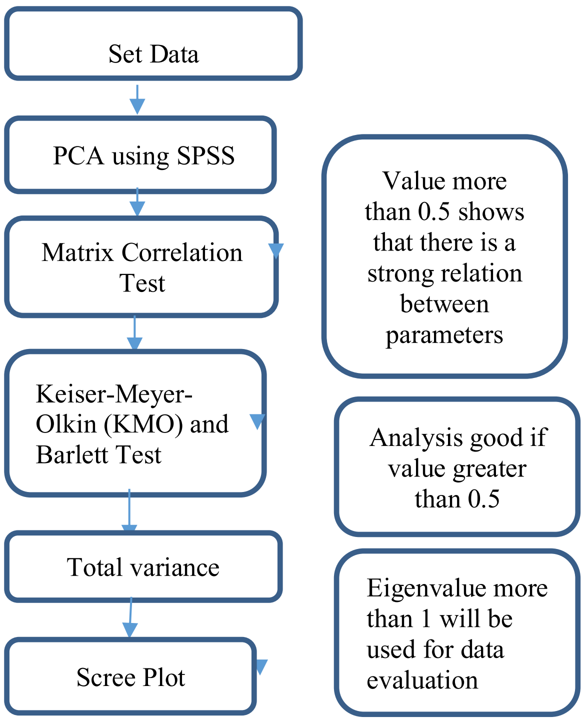

FIGURE 3. PCA Analysis Procedure

RESULTS AND DISCUSSION

According to Table 4 the values of DO parameters at stations WQ13, WQ14, WQ18, WQ19 and WQ20 are low due to an increase in the amount of organic matter in the seabed resulting in bacteria, fungi and microbes using oxygen to decompose these organic substances (Klabat et al. 2012). This can cause marine life to suffer from oxygen deprivation and cause death.

For the TSS parameter, all sampling stations showed values above 50 mg / L. This is due to the low water flow rate and transport and particles are also reduced and also influenced by low rainfall to transport sand, silt and organic matter from the mainland (Aminah et al. 2017).

At sampling stations in all of the Straits of Johor in the northeast and southwest, oil and grease readings showed 5.0 mg / L above the standard limit of 0.14 mg / L. According to a recent study by the Department of Town and Country Planning in the Selangor State Structure Plan Review Report 2035 B14: Environment (2010 - 2011), the factors leading to the increase in oil and grease were due to food outlets close to the coast where the residue of food or waste is not fully treated and has been discharged into the seawater of the coast.

Furthermore, for NH_3 parameters the readings at all five stations (WQ13, WQ14, WQ15, WQ16 and WQ17) exceeded the standard limit of 0.07 mg / L. The coastline is a tourist spot for recreation and water activities such as "banana boat" and bathing. Such activity would lead to an increase in NH_3 resulting from human waste. In addition, NH_3, which is a result of extensive combustion and absorption by clouds and is transported by rainwater to the seawater on the coast (Aminah et al. 2017)

Parameter NO_3^- shows readings exceeding the standards for all four stations at 0.06. Overall, the nitrate value of this sample is very high due to the waste generated from humans and animals, both marine and inland. The use of organic fertilizers can also contribute to the increase in nitrate content in marine water.

TABLE 4. shows parameter value for sampling of Johor Strait.

Parameters	Standard	Stations									
		WQ13	WQ14	WQ15	WQ16	WQ17	WQ18	WQ19	WQ20	WQ21	WQ22
DO (mg/L) TSS	5.00	4.22	2.97	6.32	5.53	6.04	4.48	4.81	4.85	5.05	5.05
TS (mg/L)	50.00	13.09	11.00	12.00	13.00	9.80	9.10	9.56	28.80	22.13	16.7
O&G (mg/L)	0.14	5.01	5.00	5.00	5.00	5.01	5.00	5.00	5.00	5.00	5.00
Ammonia (mg/L)	0.07	0.29	0.35	0.15	0.35	0.31	0.02	0.034	0.011	0.01	0.01
Nirate (mg/L)	0.06	1.17	1.78	1.10	1.58	1.50	0.31	0.58	0.22	0.20	0.20
Phosphate (mg/L)	0.75	0.22	0.20	0.20	0.26	0.20	0.20	0.20	0.20	0.20	0.20
Fecal Coliform (MPN/100mL)	100	4967	1600	2400	3500	3838	2083	1746	711	3556	1321

Parameter PO_4^{3-} the phosphate value of the four study stations is too low. This shows that the water at the beach has good oxidation but the nutrients needed by the marine life system are depleted. Phosphate production is due to decomposition of aquatic plants such as algae. Thus, the phosphate content is due to the decomposition that occurs in marine water.

Finally, the fecal coliform parameters for the coast of Johor Strait indicate values are above the standard limit of 100 MPN / 100mL due to sewage residues from humans and animals. In addition, the area is a recreational area where activities such as baths, picnics and the like have been done on the beach. Overall, the coastline of the Straits of Johor shows that all stations exceeded the standard. This is because the area around the beach is found to contain a lot of rubbish that can contribute to the increase in the number of fecal coliforms, mostly come from the fish. According to the Environmental Management and Resource Management Report (2011), fisheries activities, sewage sludge that are not adequately treated by shops or homes as well as domestic sewage found along the coast can lead to an increase in the amount of fungal colony. As a result, the content of fecal coliform in seawater at all stations increased.

MARINE WATER QUALITY INDEX (MWQI)

Water quality has become one of the major environmental problems worldwide (Gyawali et al. 2012). Water is needed for both; sustainable human development and healthy system function (Sabhapandit & Saikia 2010). With the growing problem of water environment pollution, research on comprehensive water quality assessment methods is becoming very important (Zahedi 2017; Noori et al. 2019).

Based on Figure 4, the overall value of MWQI for all stations is not in the very good category. In view of the MWQI values at WQ14, the value is 33.23 indicating that seawater at this station is contaminated. The station generally in Class 3 category but the MWQI values obtained indicate that this station does not meet the Class 3 classification.

The value of MWQI at stations WQ13, WQ15, WQ16, WQ17, WQ18, WQ19, WQ20, WQ21 and WQ22 are more than 50 which indicates that the seawater is moderate. The results shows that this area are for fishing and recreation. The stations should be in the Class 2 category but the value of the MWQI obtained indicates that the station is in Class 3.

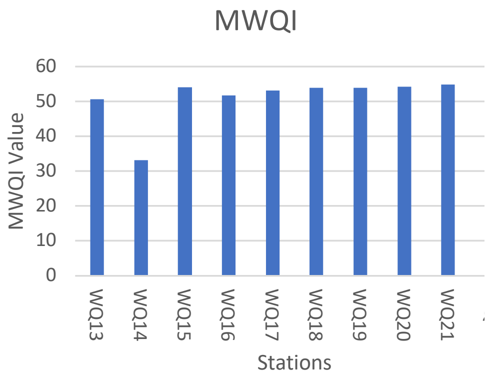

FIGURE 4. Marine Water Quality Index

MWQI COMPARISON ACCORDING TO MONSOON SEASONS

Figure 5 shows based on the thirty-five data extracted, twenty-five of them were in Class 3 which was moderate (50- <80) while the other ten were in Class 4 which was contaminated (0- <50).

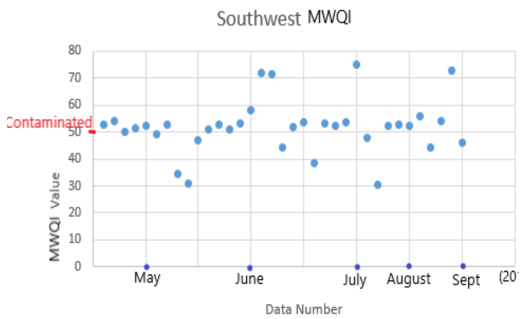

FIGURE 5. Marine Water Quality Index

While Figure 6 shows based on the thirty-five data extracted, thirty-four of them were in Class 3 that were moderate (50- <80) while only one were in Class 4 that were contaminated (0- <50). Hence, it was concluded that there was not much difference between the two monsoons as they both had almost identical classes. This may be due to data from all the stations taken to be located near each other.

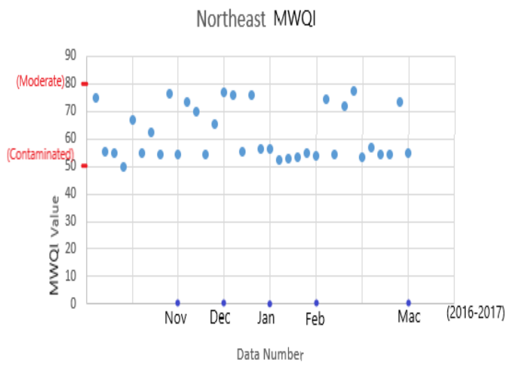

FIGURE 6. MWQI for northeast monsoon station

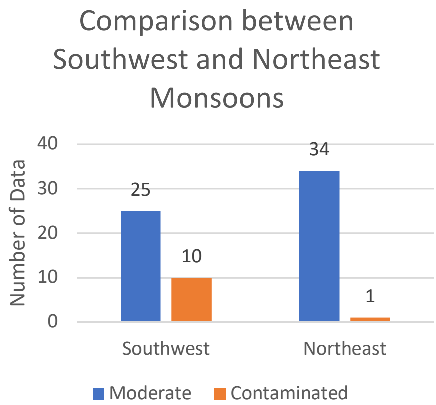

FIGURE 7. Comparison between Southwest and Northeast Monsoons

In conclusion, both monsoons season have the same classes which is moderate and contaminated. However, water quality in Northeast monsoon were better than Southwest monsoon due to the result, only one out of thirty-five data were in class 4.

PRINCIPAL COMPONENT ANALYSIS (PCA)

This study evaluates the ability of PCA techniques in determining the spatial and temporal changes of surface water quality in Shahr Chai River (Zeinalzadeh & Rezaei 2017). The main component is a linear combination of the original variables and does not correlate with each other (Zhang et al. 2019). PCA-based water quality classification method to classify water samples to gain a deeper understanding of water pollution levels (Tao et al. 2016). Using this technique and extracting the most important indicator parameters for water quality analysis are well estimated effects of rural activities, agriculture, and animal breeding on river water quality (Zeinalzadeh & Rezaei 2017).

Principal component analysis was performed to analyze the strength of the relationship between parameters and correlation values showing the relationship between x and y. The strength of the relationship between the parameters is evaluated by the correlation coefficient r. The correlation coefficient r is between -1 and +1. Values of ± 0.5 to ± 1 indicate strong relationships between parameters while values of 0 to ± 0.5 indicate weak relationships between the two parameters. The negative value can be ignored because there is no meaning in the correlation values.

Based on Table 5 there is no relation between parameters due to the r value lower than 0.5. However, there is a weak relation between dissolved oxygen and nitrate, TSS and dissolved oxygen, oil and grease and fecal coliform, oil and grease and TSS, nitrate and phosphate, nitrate and TSS and lastly phosphate and fecal coliform because the r value in range 0-0.5.

TABLE 5. Matrix Correlation for Southwest monsoon

Parameters	Dissolve oxygen (mg/L)	Oil and Grease (mg/L)	Nitrate (NO3) (mg/L)	Ammoniacal Nitrogen (mg/L)	Phosphate (mg/L)	Fecal Coliform MPN/100 ml	TSS (mg/L)
Dissolve oxygen (mg/L)	1.000						
Oil and Grease (mg/L)	-0.136	1.000					
Nitrate (NO3) (mg/L)	0.500	-0.070	1.00				
Ammoniacal Nitrogen (mg/L)	-0.039	-0.125	-0.136	1.000			
Phosphate (mg/L)	-0.424	-0.261	-0.176	0.249	1.000		
Fecal Coliform MPN/100 ml	-0.323	0.051	-0.127	-0.236	0.011	1.000	
TSS (mg/L)	0.089	0.076	-0.265	0.318	-0.033	-0.181	1.000

While based on Table 6 there is also no strong relation between parameters due to the r value lower than 0.5. However, there is a weak relation between parameters dissolved oxygen and oil and grease, dissolved oxygen and ammoniacal nitrogen, dissolved oxygen and phosphate, dissolved oxygen and TSS, oil and grease and TSS, nitrate and fecal coliform, ammoniacal nitrogen and phosphate, ammoniacal nitrogen and TSS and lastly fecal coliform and TSS due to r value varies in range 0-0.5.

These results may be due to varying pollution at each study station. In general, the causes of pollution that affect water quality can be attributed to the treatment of sewage plants, manufacturing industry, agriculture, urban and municipal areas, as well as indirect pollution such as plantation activities, construction or renovation activities, plantations, solid waste handling and others (Department of Irrigation and Drainage Malaysia, 2009).

However, at this stage it is very difficult to break down the parameters into components and give them a physical meaning. Therefore, in the next step, the main component analysis is implemented. The correlation matrix is subjected to analysis of the main components.

TABLE 6. Matrix Correlation for Northeast monsoon

Parameters	Dissolve oxygen (mg/L)	Oil and Grease (mg/L)	Nitrate (NO3) (mg/L)	Ammoniacal Nitrogen (mg/L)	Phosphate (mg/L)	Fecal Coliform MPN/100 ml	TSS (mg/L)
Dissolve oxygen (mg/L)	1.000						
Oil and Grease (mg/L)	0.074	1.000					
Nitrate (NO3) (mg/L)	-0.273	-0.108	1.000				
Ammoniacal Nitrogen (mg/L)	0.163	-0.011	-0.123	1.000			
Phosphate (mg/L)	0.028	-0.119	-0.118	0.023	1.000		
Fecal Coliform MPN/100 ml	-0.204	-0.156	0.053	-0.105	-0.082	1.000	
TSS (mg/L)	0.116	0.191	-0.054	0.088	-0.120	0.036	1.000

SCREE PLOT

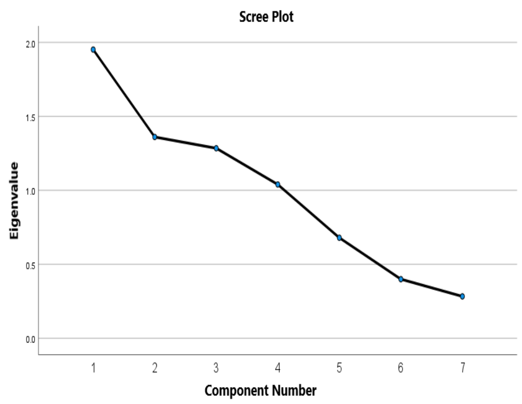

FIGURE 8. Scree Plot for Southwest monsoon

TABLE 7. Principal Component Loadings for seven Southwest parameters

Parameters	Principal Component		
	1	2	3
Dissolved Oxygen (mg/L)	-0.795	0.404	-0.0095
Oil and Grease (mg/L)	-0.703	-0.079	-0.372
Nitrate (NO3) (mg/L)	0.465	0.568	-0.139
Ammoniacal Nitrogen (mg/L)	0.613	0.068	-0.534
Phosphate (mg/L)	0.453	-0.398	0.006
Fecal Coliform MPN/100 ml	0.167	0.795	0.434
TSS (mg/L)	-0.026	-0.272	0.803
Eigenvalue	1.952	1.361	1.285
Percentage Variance (%)	27.89	19.44	18.35
Total percentage variance (%)	27.89	47.32	65.68

Based on Table 7, the analysis of the principal components includes the weights for the rotating component matrix, the eigenvalues for each component, the percent and cumulative percentages of variance explained by each component. It shows that the first tour components accounted for 80.53% of the variance in the data set. Where the first major component is 27.89%, the second main component is 47.32%, the third major component is 65.68% and the fourth component is 80.53% of the variance. The eigenvalues of the four main components (>1) can be used to evaluate the dominant hydro-chemical processes. Ammoniacal nitrogen concentration showed moderate positive loadings (0.61-0.65) while fecal coliform concentrations were at low positive loadings (0.16-0.34) for the first component. In the second component, only fecal coliform has high positive loadings (0.70-0.95). Meanwhile, the concentration of suspended solids has a high positive loading (0.70-0.95) on the third component. Finally, concentration of phosphate was at high positive loadings (0.70-0.95) while dissolved oxygen and oil and grease has low positive loadings (0.16-0.34) on the fourth component. PC1 is affected by physical properties while PC2, PC3 and PC4 affected by chemical properties which are fecal coliform, TSS and phosphate.

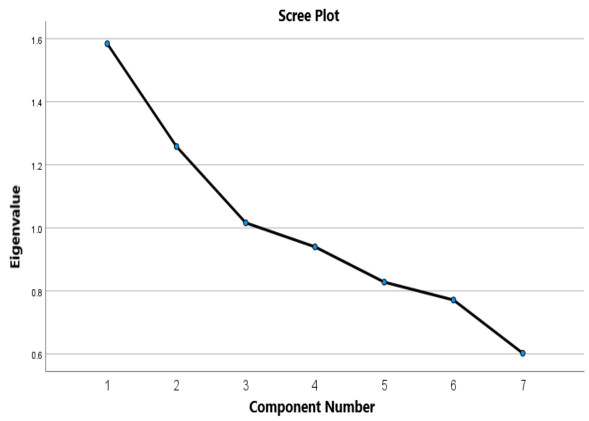

FIGURE 9. Scree Plot for Northeast monsoon

TABLE 8. Principal Component Loadings for seven Northeast parameters

Parameters	Principal Component		
	1	2	3
Dissolved Oxygen (mg/L)	0.699	-0.127	0.090
Oil and Grease (mg/L)	-0.593	0.218	-0.247
Nitrate (NO3) (mg/L)	0.455	-0.142	0.420
Ammoniacal Nitrogen (mg/L)	0.082	-0.708	-0.052
Phosphate (mg/L)	-0.482	0.190	0.675
Fecal Coliform MPN/100 ml	0.353	0.596	0.332
TSS (mg/L)	0.417	0.531	-0.449
Eigenvalue	1.584	1.258	1.016
Percentage Variance (%)	22.63	17.97	14.52
Total percentage variance (%)	22.63	40.60	55.12

Based on Table 8, the analysis of the main components includes the weights for the rotating component matrix, the eigenvalues for each component, the percent and cumulative percentages of variance explained by each component. It shows that the first three components accounted for 55.12% of the total variance in the data set. Where the first major component was 22.63%, the second major component was 40.60% and the third major component was 55.12% of the variance. The eigenvalues of the three main components (>1) can be used to evaluate the dominant hydro-chemical processes. The concentration of dissolved oxygen showed a high positive loading (0.70-0.95) for the first component. For second component, concentration of fecal coliform has a moderate positive loading (0.61-0.65) and concentration of phosphate and oil and grease has low positive loadings (0.16-0.34). Finally, concentration of phosphate has a high positive loading (0.70-0.95) on the third component. PC1 and PC3 are affected by chemical properties which are dissolved oxygen and phosphate while PC2 affected by physical properties.

Based on a study by Gajbhiye et al. (2015), he discovered that the Urdana Nala area is affected by the physical factors of PC1 while PC2 and PC3 are affected by chemical factors, specifically iron (Fe) and sulfate parameters (because it has a high positive loadings (0.70 -0.95). While the Moti Nala area is affected by chemical factors in PC1 namely copper (Cu) and chromium (Cr) while PC2 is affected by physical factors and PC3 is influenced by chemical factors namely thorium (TH).

CONCLUSION

Based on Figure 4, the overall value of MWQI for all stations is not in the very good category (90-100). Only station WQ14 has a value 33.23 which is below 50 indicating that seawater at this station is contaminated. The stations were in Class 3 but the value of MWQI obtained shows that the station do not meet Class 3 classification but were in Class 4. According to a study Aydin Uncumusaoğlu (2018), he also found that water quality in Lake Bektas, Turkey should most of its parameters are the quality of clean water according to WHO. However, some of the parameters found are contaminated water.

The value of MWQI at stations WQ13, WQ15, WQ16, WQ17, WQ18, WQ19, WQ20, WQ21 dan WQ22 are above 50 which indicates that the seawater is moderate. This area is a fishing and recreation area. The station should be in the Class 2 category but the value of the MWQI obtained indicates that the station is below Class 3. According to the Environmental Quality Report of Malaysia (2015) issued by the Department of Environment (DOE), marine water quality status for the Straits of Johor is predominantly in the good category of 80 to 90. The change in the quality of marine water quality in the Straits of Johor is due to factors such as industrial, fisheries and community activities in the area. The major factor in water quality pollution comes from the disposal of urban, indigenous and even agricultural areas (Alves et al. 2018). Although the method can evaluate the quality of water quality, it is not possible to determine the major factors affecting water quality (Avila et al. 2018).

Water quality in the southwest monsoon represented by stations WQ13, WQ14, WQ15, WQ16 and WQ17 while in the northeast monsoon season was represented by stations WQ18, WQ19, WQ20, WQ21 and WQ22. Based on Figure 6 and Figure 7, there is not much difference between the two seasons. This is because both monsoons have the same classes which is class 3 (moderate) and class 4 (contaminated). However, analysis at southwest stations indicated in 25 out of 35 data were is class 3 while the other 10 in class 4. Meanwhile, northeast stations have 34 out of 35 data were in class 3 and only one were in class 4. This shows that northeast monsoon has better water quality than southwest monsoon seasons.

ACKNOWLEDGEMENT

The authors would like to thank Universiti Kebangsaan Malaysia for their financial support under the research grant, GUP-2020-013.

REFERENCES

Abshirini, E. & Koch, D. 2016. Rivers as Integration Devices in Cities. *City, Territory and Architecture* 3(1): 1-21.

Akhir, M. F. & Jawchuen, Y. 2011. Seasonal variation of water characteristics during inter-monsoon along the East Coast of Johor (December).

Akkaraboyina, M. K. 2018. Assessment of water quality index of River Godavari at Rajahmundry Abstract : (January 2012).

Alves, J. do P. H., Fonseca, L. C., Chielle, R. de S. A. & Macedo, L. C. B. 2018. Monitoring water quality of the sergipe river basin: An evaluation using multivariate data analysis. *Revista Brasileira de Recursos Hidricos* 23: 1–12. doi:10.1590/2318-0331.231820170124

Aminah, S., Johari, M., Abdullah, P. & Aqma, W. S. 2017. The Effectiveness of the Use of New Marine Water Quality Index (MWQI) for Marine Water Quality Management Keberkesanan Penggunaan Indeks Kualiti Air Marin (IKAM) Baru bagi Pengurusan Kualiti Air Marin (The Effectiveness of the Use of New Marine Water (December).

Avila, R., Horn, B., Moriarty, E., Hodson, R. & Moltchanova, E. 2018. Evaluating statistical model performance in water quality prediction. *Journal of Environmental Management* 206: 910–919. doi:10.1016/j.jenvman.2017.11.049

Aydin Uncumusaoğlu, A. 2018. Statistical assessment of water quality parameters for pollution source identification in Bektaş pond (Sinop, Turkey). *Global Nest Journal* 20(1): 151–160. doi:10.30955/gnj.002369

Behera, M. R., Chun, C., Palani, S. & Tkalich, P. 2013. Temporal variability and climatology of hydrodynamic, water property and water quality parameters in the West Johor Strait of Singapore. *Marine Pollution Bulletin* 77(1–2): 380–395. doi:10.1016/j.marpolbul.2013.09.043

Bengraïne, K. & Marhaba, T. F. 2003. Using principal component analysis to monitor spatial and temporal changes in water quality. *Journal of Hazardous Materials* 100(1–3): 179–195. doi:10.1016/S0304-3894(03)00104-3

Camerlengo, A. & Demmler, M. I. 1997. Wind-driven circulation of Peninsular Malaysia's eastern continental shelf. *Scientia Marina* 61(2): 203–211.

Charles Ezugwua, K. O., Chisom Ezugwuc, , K. O., Adegboyega Odumadea, , O. O., , M. H. & , A. O., Wokocha Innocentc. 2021. Community Water Demand and Sustainable Water Supply Planning in Nigeria : A Review. *Jurnal Kejuruteraan* 33(3)(13): 517-530.

Daryabor, F., Tangang, F. & Juneng, L. 2014. Simulation of southwest monsoon current circulation and temperature in the east coast of peninsular Malaysia. *Sains Malaysiana* 43(3): 389–398.

Dawood, M. & Gobreal, S. 2010. Assessment of Trihalomethanes in Noubaria Water Purification Plant and its Distribution System اهب صاخلا عيزوتلا ةكبش و ه ايملا ةيقنتل ةيرابونلا ة طحم ب ناثيم ولاه لا يثلاث ريدقت.

Gajbhiye, S., Sharma, S. K. & Awasthi, M. K. 2015. Application of Principal Components Analysis for Interpretation and Grouping of Water Quality Parameters. *International Journal of Hybrid Information Technology* 8(4): 89–96. doi:10.14257/ijhit.2015.8.4.11

Gyawali, S., Techato, K., Yuangyai, C. & Monprapusson, S. 2012. Evaluation of Surface Water Quality using Multivariate Statistical Techniques : A Case Study of U-tapao River Basin , Thailand. *KMITL Sci. Tech. J.* 12(1): 7–20.

House, M. A. 1990. Water quality indices as indicators of ecosystem change. *Environmental Monitoring and Assessment* 15(3): 255–263. doi:10.1007/BF00394892

House, M. A. & Ellis, J. B. 1987. The development of water quality indices for operational management. *Water Science and Technology* 19(9): 145–154. doi:10.2166/wst.1987.0076

Iyer, C. S. P., Sindhu, M., Kulkarni, S. G., Tambe, S. S. & Kulkarni, B. D. 2003. Statistical analysis of the physico-chemical data on the coastal waters of Cochin. *Journal of Environmental Monitoring* 5(2): 324–327. doi:10.1039/b209219k

Jayakumar, R. & Siraz, L. 1997. Factor analysis in hydrogeochemistry of coastal aquifers - A preliminary study. *Environmental Geology* 31(3–4): 174–177. doi:10.1007/s002540050177

Juahir, H., Zain, S. M., Yusoff, M. K., Hanidza, T. I. T., Armi, A. S. M., Toriman, M. E. & Mokhtar, M. 2011. Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. *Environmental Monitoring and Assessment* 173(1–4): 625–641. doi:10.1007/s10661-010-1411-x

Klabat, T., Bangka, P. & Simanjuntak, M. 2012. Oksigen Terlarut dan Apparent Oxygen Utilizationdi Perairan Teluk Klabat, Pulau Bangka. *Ilmu Kelautan - Indonesian Journal of Marine Sciences* 12(2): 59–66. doi:10.14710/ik.ijms.12.2.59-66

Ling, T. Y., Soo, C. L., Liew, J. J., Nyanti, L., Sim, S. F. & Grinang, J. 2017. Application of Multivariate Statistical Analysis in Evaluation of Surface River Water Quality of a Tropical River. *Journal of Chemistry* 2017. doi:10.1155/2017/5737452

Liou, S. M., Lo, S. L. & Wang, S. H. 2004. A generalized water quality index for Taiwan. *Environmental Monitoring and Assessment* 96(1–3): 35–52. doi:10.1023/B:EMAS.0000031715.83752.a1

Liu, Y. J., Tran, T., Postma, G., Buydens, L. M. C. & Jansen, J. 2018. Estimating the number of components and detecting outliers using Angle Distribution of Loading Subspaces (ADLS) in PCA analysis. *Analytica Chimica Acta* 1020: 17–29. doi:10.1016/j.aca.2018.03.044

Martina, A. & Radjawane, I. M. 2019. Data analysis on sea water quality data in Jakarta Bay using Principal Components Analysis (PCA) method during transitional monsoon 2012. *IOP Conference Series: Earth and Environmental Science* 339(1). doi:10.1088/1755-1315/339/1/012023

Mu, A. 2019. 済無No Title No Title. *Journal of Chemical Information and Modeling* 53(9): 1689–1699. doi:10.1017/CBO9781107415324.004

Mutlu, E. 2019. Evaluation of spatio-temporal variations in water quality of Zerveli stream (northern Turkey) based on water quality index and multivariate statistical analyses. *Environmental Monitoring and Assessment* 191(6). doi:10.1007/s10661-019-7473-5

Noori, R., Berndtsson, R., Hosseinzadeh, M., Adamowski, J. F. & Abyaneh, M. R. 2019. A critical review on the application of the National Sanitation Foundation Water Quality Index. *Environmental Pollution* 244: 575–587. doi:10.1016/j.envpol.2018.10.076

Oana, I. 2010. Water Quality Index -Assessment Method of the Motru River Water Quality (Oltenia, Romania) Indicele De Calitate a Apei – Metodă De Determinare a Calităţii Apei Râului Motru (Oltenia, România). *Geography UNIVERSITATEA DIN CRAIOVA Seria: Geografie* 13(13): 74–83. Retrieved from http://analegeo.ro/wp-content/uploads/2010/12/5.-IONUS-Oana.pdf

Onojake, M. C., Ukerun, S. O. & Iwuoha, G. 2011. A Statistical Approach for Evaluation of the Effects of Industrial and Municipal Wastes on Warri Rivers, Niger Delta, Nigeria. *Water Quality, Exposure and Health* 3(2): 91–99. doi:10.1007/s12403-011-0046-7

Perera, M., Ng, A., Muthukumaran, S., O'connor, J. & Perera, B. 2018. Investigation of Water Quality in Combined Recycled Water and Stormwater Systems. *Urban Water Journal* 15(5): 478-487.

Praus, P. 2005. Water quality assessment using SVD-based principal component analysis of hydrological data. *Water SA* 31(4): 417–422. doi:10.4314/wsa.v31i4.5132

Qadir, A., Malik, R. N. & Husain, S. Z. 2008. Spatio-temporal variations in water quality of Nullah Aik-tributary of the river Chenab, Pakistan. *Environmental Monitoring and Assessment* 140(1–3): 43–59. doi:10.1007/s10661-007-9846-4

Rose, K. C., Winslow, L. A., Read, J. S. & Hansen, G. J. A. 2016. Climate-induced warming of lakes can be either amplified or suppressed by trends in water clarity. *Limnology and Oceanography Letters* 1(1): 44–53. doi:10.1002/lol2.10027

Sabhapandit, P. & Saikia, P. 2010. Interpretation of ground and surface water quality using principal component analysis in GOHPUR sub-division of SONITPUR DISTRICT , ASSAM , INIDA. *Int. J. Chem. Sci* 8(3): 1523–1536.

Salman, S. R. & Abu Ruka'h, Y. H. 1999. Multivariate and principal component statistical analysis of contamination in urban and agricultural soils from north Jordan. *Environmental Geology* 38(3): 265–270. doi:10.1007/s002540050424

Sarkar, A. & Pandey, P. 2015. River Water Quality Modelling Using Artificial Neural Network Technique. *Aquatic Procedia* 4(Icwrcoe): 1070–1077. doi:10.1016/j.aqpro.2015.02.135

Sharma, P. K. & Meena, R. S. 2018. PHYSICO-CHEMICAL ANALYSIS OF RIVER GANGA AT VARANASI CITY IN Indian Journal of Agriculture (March).

Subaila Jemain_2007.pdf. (n.d.).

Tangang, F. T., Xia, C. & Qiao, F. 2011. wave – tide – circulation coupled model Seasonal circulations in the Malay Peninsula Eastern continental shelf from a wave – tide – circulation coupled model (September). doi:10.1007/s10236-011-0432-5

Teixeira de Souza, A., Carneiro, L. A. T. X., da Silva Junior, O. P., de Carvalho, S. L. & Américo-Pinheiro, J. H. P. 2020. Assessment of water quality using principal component analysis: a case study of the Marrecas stream basin in Brazil. *Environmental Technology (United Kingdom)* 1–11. doi:10.1080/09593330.2020.1754922

Tyagi, S. & Sharma, B. 2014. Water Quality Assessment in Terms of Water Quality Index Water Quality Assessment in Terms of Water Quality Index Water Quality Assessment in Terms of Water Quality Index. *American Journal of Water Resources, 2013 1 (3), pp 34-38.* 1(3): 34–38. doi:10.12691/ajwr-1-3-3

Uncumusaoğlu, A. A. & Akkan, T. 2017. Assessment of water quality of yağlidere stream (Turkey) using multivariate statistical techniquess. *Polish Journal of Environmental Studies* 26(4): 1715–1723. doi:10.15244/pjoes/68952

Wood, A. K. H., Ahmad, Z., Shazili, N. A. M., Yaakob, R. & Carpenter, R. 1997. Geochemistry of sediments in Johor Strait between Malaysia and Singapore. *Continental Shelf Research* 17(10): 1207–1228. doi:10.1016/S0278-4343(97)00011-3

Woods, C. M. & Edwards, M. C. 2011. Factor Analysis and Related Methods. *Essential Statistical Methods for Medical Statistics*, hlm. Vol. 27. Elsevier B.V. doi:10.1016/B978-0-444-53737-9.50009-8

Zahedi, S. 2017. Modification of expected conflicts between Drinking Water Quality Index and Irrigation Water Quality Index in water quality ranking of shared extraction wells using Multi Criteria Decision Making techniques. *Ecological Indicators* 83(June): 368–379. doi:10.1016/j.ecolind.2017.08.017

Zeinalzadeh, K. & Rezaei, E. 2017. Determining spatial and temporal changes of surface water quality using principal component analysis. *Journal of Hydrology: Regional Studies* 13(August 2016): 1–10. doi:10.1016/j.ejrh.2017.07.002

Zhang, L., Zhong, M., Xu, Y., Wang, Z. & Huang, H. 2019. The Water Quality Evaluation in Balihe Lake Based on Principal Component Analysis. *Journal of Geoscience and Environment Protection* 07(08): 38–48. doi:10.4236/gep.2019.78003