Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2022.28677

269 (2022) 106–122
September

Spatial and temporal variations of river water quality using multivariate statistical techniques

Hassan M.A. Alssgeer[a,b], Mohd Khairul Amri Kamarudin[a,c,*], Mohd Armi Abu Samah[d], Mohd Ekhwan Toriman[e], Muhammad Barzani Gasim[a], Marlia M. Hanafiah[f], Laila O.M. Alubyad[a], Ahmad Shakir Mohd Saudi[g], Khairul Nizam Maulud[h], Noorjima Abd Wahab[a], Siti Nor Aisyah Bati[a], Mohamed Erhayem[b]

[a]*East Coast Environmental Research Institute (ESERI), University Sultan Zainal Abidin, Gong Badak Campus, 21300 Kuala Nerus, Terengganu, Malaysia, email: mkhairulamri@unisza.edu.my (M.K.A. Kamarudin)*
[b]*Chemistry Department, Faculty of Science, Sebha University, Sebha, Libya*
[c]*Faculty of Applied Social Science, Universiti Sultan Zainal Abidin, Gong Badak Campus, 21300 Terengganu, Malaysia*
[d]*Kulliyyah of Science, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia*
[e]*Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
[f]*Center for Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*
[g]*Institute of Medical Science Technology, Universiti Kuala Lumpur, 43000 Kajang, Selangor, Malaysia*
[h]*Earth Observation Center, Institute of Climate Change, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia*

Received 21 May 2022; Accepted 22 July 2022

A B S T R A C T

The assessment of temporal/spatial variability and the interpretation of large and complex datasets of water quality were performed using multivariate statistical techniques such as cluster analysis (CA) and factor analysis (FA). Water quality of the Nerus River for 27 parameters was monitored at eight sampling stations. Three different similarity groups between sampling sites that reflected different water quality parameters were identified by the CA, while the FA/principal component analysis has determined nine factors responsible for the data structure that account for 82.24% of the total variance of the dataset. 14 parameters are needed to explain 82.24% of water quality changes for both temporal and spatial, hence the significant data reduction was not achieved. The findings suggested the compulsion and effectiveness of environmental techniques for interpretation of large datasets are targeting to gain information about water quality using temporal and spatial characterizations at the designated water monitoring stations in the river.

Keywords: Spatial and temporal assessment; Water quality; River water; Nerus River basin; Multivariate analysis

1. Introduction

Water is one of the rich natural resources used by human beings, animals, plants and other living organisms [1–3]. The welfare of humanity depends on the quality and quantity of clean water supply [3]. Natural processes including rainfall intensity, climatology, sediment transport and anthropogenic activities including urbanization, industrial and agricultural practices have greatly affecting the quality of water resources [4]. Rivers play a major

* Corresponding author.

role in the adaptation or transport of urban and industrial effluent and the outflow from agricultural land. Municipal and industrial spills are a constant source of pollution, and surface run off is a seasonal phenomenon that is heavily influenced by the climate in the basin. [5,6].

Pollutants often enter into river systems and are ultimately transported to the marine environment, disturbing the water quality of the river in the form of organic or inorganic [7]. It was reported that one of the worldwide issue related to the environmental concern was the surface water pollution with toxic chemicals and excess nutrients, caused by the groundwater discharges, storm water runoff and vadose zone leaching [8]. River discharge and concentration of pollutants in river water were greatly affected by the seasonal variations in precipitation, surface runoff, interflow, groundwater flow and pumped in and outflows [5,9]. Hence, fundamental understanding of hydromorphological, biological and chemical characteristics is needed to develop an effective surface water management. However, a monitoring program providing a demonstrative and trustworthy prediction of surface water quality is necessary due to spatial and temporal variations in water quality [4,8,10]. As a result of the acknowledgement of these environmental issues, various water quality research had been carried out in Nerus River basin, especially with respect to identify the thresholds that have significant impact on the environment. Furthermore, the effects of spatial and temporal differences and the effect of rain, runoff and land use around the river area were studied [7]. Moreover, several previous studies on the use of land and its impact on the river quality have been reported at the Nerus River basin [11–13].

Malaysia has developed National Rivers Conservation Program where water quality of river systems is observed at several selected stations for various of physico-chemical, bacteriological and hydrological parameters regularly with an output of remarkable databases of high complexity. This monitoring programs required high amounts of financial budget. Thus, there is a need to optimize the monitoring networks, a number of water-quality parameters, reducing these to representative ones without losing useful information [14,15]. One of the effective tools for a significant data reduction and interpretation of multi-constituent measurements is by applying multivariate statistical techniques and experimental data analysis [16]. Application of various multivariate statistical methods such as cluster analysis (CA), discriminant analysis (DA), principal component analysis (PCA), and factor analysis (FA) helps interpret complex data matrices to better comprehend water quality and ecosystems, as well as enable to identifies factors that affect water systems and provides a valuable tool for the effective management of water resources [4–5,17]. Further, Geographical Information System (GIS) with spatial and temporal interpolation, which combine different available water quality data into an easily comprehensive format provides a way to summarize overall water quality conditions in a manner that can be clearly communicated to policymakers [12]. In the present study, multivariate statistical techniques CA, DA, PCA and FA were applied to assess the spatial and temporal variation in the river water quality datasets of Nerus River basin.

2. Materials and methods

2.1. Study area

Nerus River is located in tributaries of Terengganu River basin passes through Setiu and Kuala Terengganu districts, on the east coast of Peninsular Malaysia. Nerus River basin comprises of tributaries namely Tepuh River, Tongkat River, Kasar River, Pelung River, Kulai River, Tayur River Telemong River, Las River, Tong River, Linggi River, Kepayang River, Temiang River and Semelang River) as shown in Fig. 1 [13,18,19]. Nerus River is the largest catchment in Terengganu River basin with an area of 851 km^2. In general, Nerus catchment is characterized by a low-lying slope over the area (mostly in the Eastern part) with a high degree of slope located at the West of the catchment. The long formation of the Nerus River with estimated distance of 92.85 km produces a large amount of surface runoff. Geographical location of Nerus River is located at latitudes of 103°00′E to 103°06′E and longitude of 05°13′N to 05°23′N in the northeastern coastal region of Peninsular Malaysia. Its source is at Gunung Sarut and flows southeastern towards the mouth of Nerus River which discharges its water into Terengganu River estuary before finally discharging into the South China Sea [19]. It passes through the populated urban area of northeastern Kuala Terengganu and receives and carries different kinds of agricultural and urban solid and liquid wastes produced by agriculturally based industries and domestic sewage [13,20].

Moreover, the climate is usually hot and fairly humid year-round, averaging from 28°C to 33°C/d [21]. The area is dominated by tropical climate with two seasons (the rainy and dry season), the rainy season is characterized by high rainfall, which sometimes cause flooding [11,13]. Nerus River is considered as water resource of cultivation and domestic water supply (DOE, 2011). Water quality of Nerus River was determined by including 28 parameters at eight sampling locations covering two different seasons (wet and dry) during 2016–2017. The study was conducted from August 2016 until August 2017 at eight sampling stations that have been determined based on the distance from the upstream to downstream (2 km for each station) to provide the real condition of water quality for the whole basin. Locations of the sampling stations are illustrated in Fig. 2, while the coordinates of the stations are shown in Table 1.

2.2. Data collection and treatment

Surface water samples were collected about 10 cm below the water surface. The sampling stations were selected covering the upstream to downstream. The selection criteria of the sampling locations were based on the characteristics of the water condition, land use and human activities along the river. Water samples were collected from eight sampling stations with a distance of 2 km for each station (Table 2).

On the other hand, certain parameters require onsite measurement due to the rapid changes in the reading caused by various factors such as activity of microorganisms etc. Physico-chemical parameters, such as temperature (TEMP), dissolved oxygen (DO), conductivity (EC), turbidity (TUR), salinity (SAL), total dissolved solids (TDS)

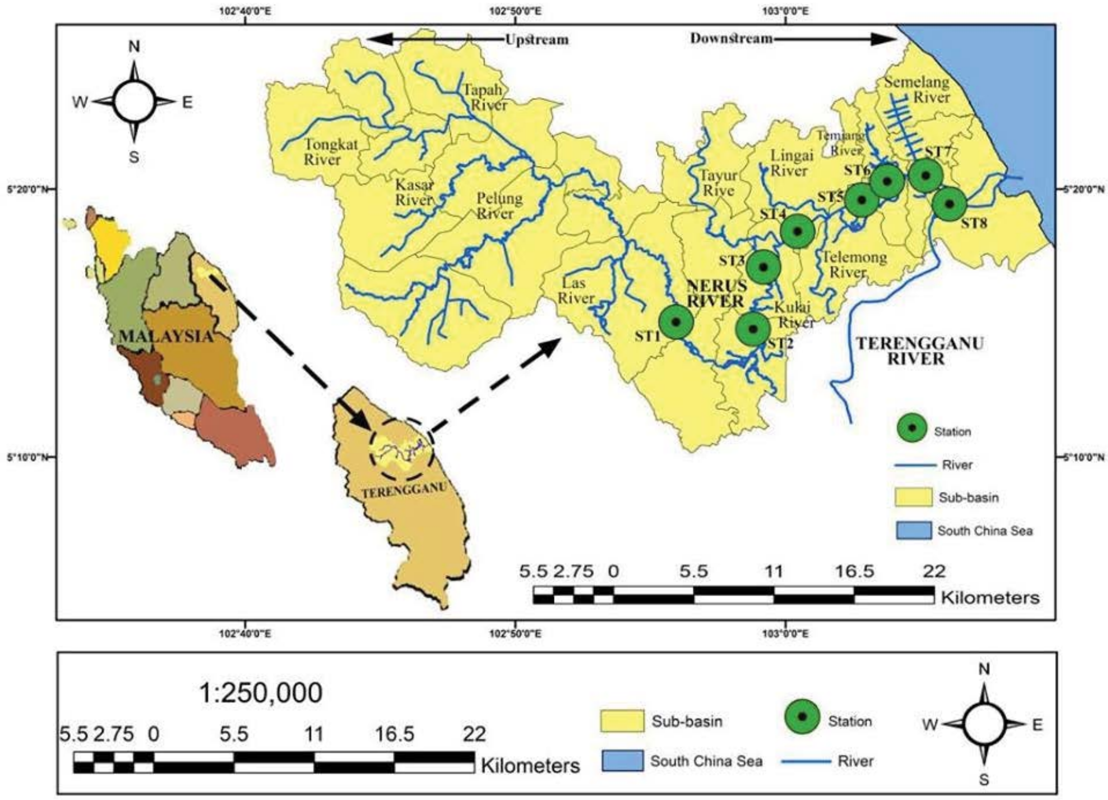

Fig. 1. Map of study area, locations of specific flow measurement and water sampling stations at Nerus River.

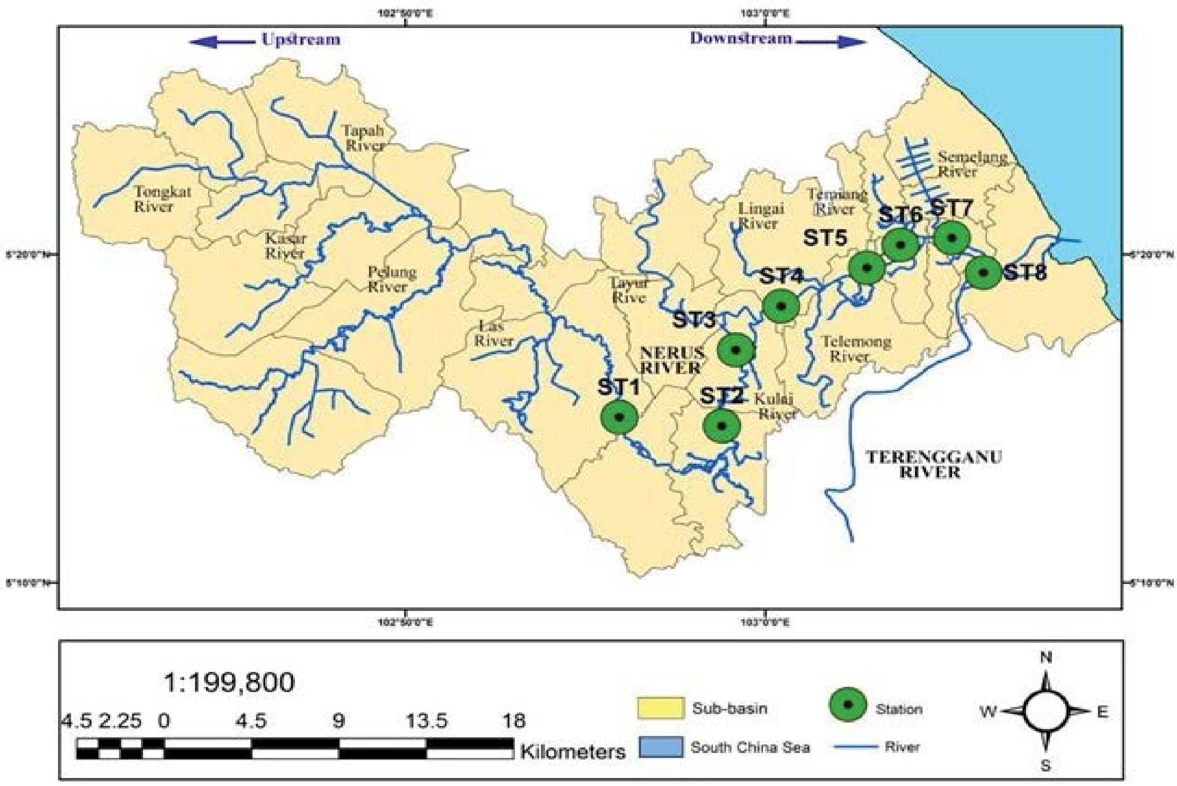

Fig. 2. Locations of Nerus River and sampling station. Schematic of upstream catchment delineation on the basis of the positions of the sampling station.

Table 1
Locations of waters sampling stations and characteristics of the surrounding Nerus River area

Station No.	Location				
	Latitude and longitude	Elevation (m)	Depth (m)	Width (m)	Land use in surrounding area
ST 1	5°17′57.6′′N 102°59′51.6′′E	8	3.4	70	Upstream, forest, recreation (Tepuh River, Tongkat River, Kasar River, Pelung River)
ST 2	5°18′34.4′′N 103°00′40.2′′E	11	5	75	Forest, agricultural activities, oil palm, and rubber (Kulai River, Tayur River)
ST 3	5°18′54.5′′N 103°01′29.6′′E	20	8.5	86	Mining activity and deforestation, palm plantation (Telemong River)
ST 4	5°19′22.5′′N 103°02′41.0′′E	15	6.6	80	Human activities, settlements, agricultural activities, mixed horticulture (Las River)
ST 5	5°20′09.8′′N 103°03′27.5′′E	14	4.8	106	Urban activity, elimination of forests, livestock farms (Tong River, Linggi River)
ST 6	5°21′00.2′′N 103°04′01.9′′E	20	8	80	Residential area, agricultural, livestock farms (Kepayang River)
ST 7	5°20′33.00′′N 103° 5′2.80′′E	18	5.5	100	Deforestation and construction of bridge, residential area (Temiang River)
ST 8	5°19′37.4′′N 103°05′56.0′′E	16	3.8	221	Downstream, residential area, urbanization, human activity (Semelang River)

Table 2
Sampling date sampling time

Sampling period	Date of sampling	Distance between stations	Season
1st sampling	8th Aug 2016	2 km	Rainy
2st sampling	19th Oct 2016	2 km	Rainy
3st sampling	16th Dec 2016	2 km	Rainy
4st sampling	21th Feb 2017	2 km	Dry
5st sampling	9th Apr 2017	2 km	Dry
6st sampling	24th Jun 2017	2 km	Dry

and pH, were directly measured in the field by using Multiparameter (YSI 556 meter). While, the *ex-situ* analysis for heavy metals and microbiological testing were conducted in the laboratory. Water quality parameters included biochemical oxygen demand (BOD), chemical oxygen demand (COD), ammoniacal nitrogen (NH_3–N), total suspended solids (TSS), nitrite (NO_2), nitrate (NO_3), phosphate (PO_4), sulfate (SO_4), oil and grease (O&G) and *E. coli*. In addition, the dissolved heavy metals analysis included iron (Fe), zinc (Zn), cadmium (Cd), manganese (Mn), chromium (Cr), nickel (Ni), lead (Pb), copper (Cu), sodium (Na) and cobalt (Co). The water quality parameters, units and methods of analysis are summarized in Table 3.

2.3. Statistical method

The spatial and temporal variations of the stream water quality parameters were first evaluated through season parameter correlation matrix. The water quality parameters were grouped into different seasons based on temporal CA and different stations based on spatial CA differences. River water quality data sets were subjected to three multivariate techniques: cluster analysis (CA), discriminant analysis (DA) and principal component analysis PCA/FA [22–24]. In contrast, the Geographic Information System (GIS) was used as a computational platform to analyse the land use map. The use of topographical maps and aerial photographs provide a good database for land use change information within the Nerus River basin. All statistical analysis was carried out using XLSTAT version 2014.

2.3.1. Cluster analysis

The CA, is a group of multivariate techniques whose primary purpose is to assemble objects based on the characteristics they possess. The CA classifies objects, so each object is similar to the others in the cluster with respect to a predetermined selection criterion [4]. Further, it is a data reduction tool that creates subgroups that are more manageable than individual datum. In cluster analysis there is no prior knowledge about which elements belong to which clusters. The grouping or clusters are defined through an

Table 3
Methods and instruments used for physical-chemical parameters of water quality

Parameter	Equipment/instruments method	Unit
Depth to surface water (m)	Depth gauge measuring tape	m
Temperature	YSI Model 5564 MPS Handheld Multiparameter Meter	°C
pH	YSI Model 5565 MPS Handheld Multiparameter Meter	–
Conductivity	YSI Model 5560 MPS Handheld Multiparameter Meter	μS/cm
Dissolved oxygen (DO)	YSI Model 559 MPS Handheld Multiparameter Meter	mg/L
Salinity	YSI Model 556 MPS Handheld Multiparameter Meter	ppt
Turbidity	HACH Model 2100 P Portable Turbidity Meter	NTU
Total dissolved solid (TDS)	YSI Model 556 MPS Handheld Multiparameter Meter	mg/L
Chemical oxygen demand (COD)	HACH Model DR/2500 Spectrophotometer (Method 430, LR)	mg/L
Biochemical oxygen demand (BOD)	YSI Model 5000 Dissolved Oxygen Meter (APHA Method 5210, 2012)	mg/L
Ammoniacal nitrogen (NH_3–N)	HACH Model DR/2500 Spectrophotometer (Nessler Method)	mg/L
Total suspended solids (TSS)	APHA Method 2540 D (1998)	mg/L
Nitrate (NO_3)	HACH Model DR/2500 Spectrophotometer	mg/L
Nitrite (NO_2)	HACH Model DR/2500 Spectrophotometer (Cadmium Reduction LR Method)	mg/L
Phosphate (PO_4)	HACH Model DR/2500 Spectrophotometer (Ascorbic Acid Method)	mg/L
Sulfate (SO_4)	HACH Model DR/2500 Spectrophotometer (SulfaVer 4 Method)	mg/L
Oil and grease	Gravimetric Separatory Funnel Extraction (APHA Method 5520 B, 2012)	μg/L
Heavy metals	ICP-MS analytical	mg/L
E. coli	APHA, 2012 (Coliform Agar (CCA))	Count/100 mL

analysis of the data. The Euclidean distance usually gives the similarity between two samples, and a distance can be represented by the difference between transformed values of the samples [3].

The resulting from CA clusters of objects should then exhibit high internal (within–cluster) homogeneity and high external (between clusters) heterogeneity. Hierarchical agglomerative clustering is the most common approach, which provides intuitive similarity relationships between any one sample and the entire data set, typically illustrated by a dendrogram [25]. Moreover, the dendrogram provides a visual summary of the clustering processes, presenting a picture of the groups and their proximity, with a dramatic reduction in dimensionality of the original data. The Euclidean distance usually gives the similarity between two samples and a distance can be represented by the difference between analytical values from the samples [5].

In this study, hierarchical agglomerative CA was performed on the normalized data set by means of the Ward's method, using squared Euclidean distances as a measure of similarity. The spatial and temporal variabilities of water quality in the whole river basin were determined using the linkage distance, which represents the quotient between the linkage distances for a particular case divided by the maximal linkage distance. The Euclidean distance (linkage distance) in this study reported in equation:

$$\frac{D_{\text{link}}}{D_{\text{max}}} \times 100 \tag{1}$$

where D_{link} is a linkage distance and D_{max} is a maximal distance. The quotient is multiplied by 100 to standardize the linkage distance represented by the y-axis [5,16,26].

2.3.2. Discriminant analysis

The basic objectives of discriminant analysis (DA) are to differentiate between occurring groups and to assign new observations into the classified groups [27]. In DA module, multiple quantitative attributes are used to discriminate between two or more naturally occurring groups. In contrast to CA, whereas, DA provides statistical classification of samples and it is performed with prior knowledge of membership of objects to a particular group or cluster. Furthermore, DA helps in grouping samples sharing common properties. In data mining, DA is also employed to discover most important quantitative variables separating the groups and for testing the hypothesis on the differences between the groupings expected [27,28]. Thus, DA becomes not only an effective technique to evaluate relationships between different clusters, but also to perform the best discrimination of individuals into groups defined a priori [29]. Three modes of DA were applied, which were standard mode, forward stepwise mode and backward stepwise mode. A standard mode was performed to construct DFs for evaluating spatial variations in the water quality. Variables were gradually eliminated starting with the most significant variable until no significant changes were found. What about forward stepwise mode and backward stepwise mode.

In this study, two groups for temporal (two seasons) and groups for spatial (sampling regions) evaluations have been selected and the number of analytical parameters used to assign a measure from a monitoring site into a group (season or monitoring area). The site (spatial) and the season (temporal) were the grouping (dependent) variables, whereas all the measured parameters constituted the independent variables, which can be determined by the equation:

$$f\left(G_i\right) = K_i \sum_{j=1}^{n} w_{ij}\ P_{ij} \tag{2}$$

where i is the number of groups (G), K_i is the constant inherent to each group, n is the number of parameters used to classify a set of data into a given group and w_j is the weight coefficient assigned by DF analysis (DFA) to a given parameter (P_j) [30].

DA was performed on each raw data matrix using standard, forward stepwise and backward stepwise modes in constructing DFs to evaluate both the spatial and temporal variations of water quality in Nerus River basin.

2.3.3. Principal component analysis

Principal component analysis (PCA) and factors analysis (FA) are statistical approaches that can be applied to analyze interrelationships among a large number of variables based on their common underlying dimension by providing empirical estimates of the variables [22]. Further, it provides information on the most significant parameters due to spatial and temporal variations which describes the whole data set by excluding the less significant variables parameters are excluded from the whole data set with very minimum loss of original information [5,16,26].

The PCA is designed to transform the original variables into new and uncorrelated variables (axes, known as principal components), which are linear combination of the original variables. It also provides an objective way of finding indices of this type so that the variation in the data set can be accounted for as concisely as possible [31]. PCA gives the information on the most significant parameters describing the majority of the data set; affording data reduction with minimum loss of original information [32]. Factor analysis (FA) is usually applied as a method to interpret a large complex data matrix and offers a powerful means of detecting similarities among variables or samples [33]. PCs generated by PCA are sometimes not readily interpreted, therefore, it is advisable to rotate PCs by varimax rotation which is carried out using FA.

Varimax rotations are applied on PCs with eigenvalues more than 1. These are considered significant to obtain new groups of variables call varimax factors (VFs) [34–36]. The principal component (PC) can be expressed as:

$$z_{ij} = a_{f1} f_{1i} + a_{f2} f_{21} + \dots\ +\ a_{\text{fm}} f_{\text{mi}} + e_{fi} \tag{3}$$

where z is the measured value of a variable, a is the factor loading, f is the factor score, e is the residual term accounting for errors or other sources of variation, i is the sample number, j is the variable number and m is the total number of factors.

In this study, the PCA technique was applied on the water quality data set to evaluate the compositional patterns among the studied water quality variables and to recognize the factors that manipulate each of the discovered regions produced by HACA. Further, the correlation analysis CA was used to evaluate the strength of the relations between water quality variables with each other? PCA normally involves three main steps: the standardization of measurements to ensure that they have equal weights in the analysis by auto scaling the data to produce new variables, the calculation of the co-variance matrix by identifying the Eigenvalues and their corresponding eigenvectors and the elimination of components that account only for a small proportion of the variation in datasets [36,37]. In addition to, PCA/FA was applied to the normalized datasets (27 variables) separately for three different spatial regions labelled as HPS, MPS and LPS and delineated by the HACA technique [38,39]. Nevertheless, in this study, all of these analyses were made concerning the need for and selection of, appropriate transformations to achieve a better approximation of the normal distribution for CA and FA [40]. Also, GIS was used for handling geographic data in digital forms.

In this study, Nerus River classification is designed to capture, store, manipulate, analyse and display diverse sets of spatial and temporal or georeferenced data as a spatial analysis tool.

3. Results and discussion

3.1. Spatial similarity and site grouping

Cluster analysis (CA) was used to detect the similarity groups between the sampling stations. It yielded a dendrogram as shown in Fig. 3. All of 27 samples were grouped and three clusters were generated from the clustering method in a convincing way, as the eight stations in these clusters, share the homogeneity characteristics. Also, CA was used to test the data of water quality and to determine the similarity of sampling stations, as well as to classify specific areas of pollution, with different land use activities [22,37]. HACA was executed on the standardized data set using Ward's method with Euclidean distances for the determination of similarity. Where, all eight sampling stations along the Nerus River were grouped into three significant different clusters at ($D_{\text{link}}/D_{\text{max}}$). Distinguished via the Ward method that uses the squared Euclidean distance as a similarity measure. However, cluster 1 (only station 1) cluster 2 (comprised stations 2, 3, 4 and 5), and, cluster 3 (including three stations: 6, 7 and 8) were shown in Fig. 3. Water sampled at different sampling stations did not show a apparent trend. The findings in this test may be attributed to the complex of impacts taking place on the river, including the various pollution sources and disturbances caused by human activities.

Relationship between water quality and GIS-HACH based analysis recorded between 2017 and 2018 in the Nerus River basin was conducted. The relationship between spatial and temporal changes in land use and its impact on water quality shows that urbanization was a key factor affecting the river water quality, followed by horticultural anthropogenic activities (rural area). These are often in the vicinity of rivers, due to higher urbanization and agricultural activities. GIS results showed greater significance for the sampling site groups (land use activities) than for the sampling events. In this study, it was found that the based on the change in the Nerus River basin land

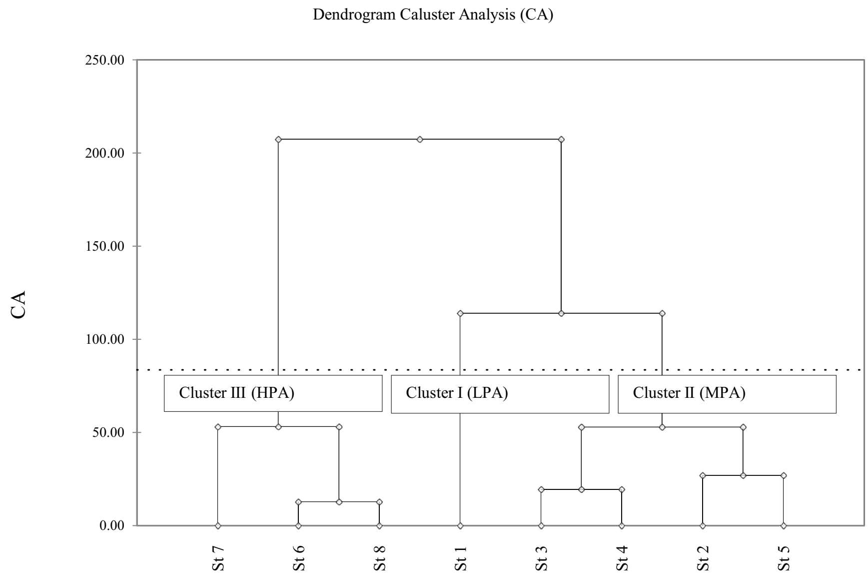

Fig. 3. Cluster analysis dendrogram showing sampling stations in the Nerus River water quality data.

use recorded between 2017 to 2018 especially at the middle and at downstream the area close to the river banks (urbanisation, Settlement and Residential Expansion, mixed horticulture, agricultural activity) and considerable many land use activities as reported by [13]. While, the upstream region is (forestry, characterised by intensive palm oil agricultural activity).

As shown in Fig. 4a and b water quality at cluster 1, the sampling station in particular station 1 (Tepuh River, Tongkat River, Kasar River and Pelung River) was located in the upstream associated with less pollution area (LPA), was surrounded by extended forest covering dense and diverse trees. In addition, human activities at this station were limited. Therefore, the condition of water quality was slightly clean and optimized at upstream. These results obtained were expected as the upstream section of a river usually exhibits better water quality when the area is less populated by humans and their activities [12–13,41]. Sampling stations in cluster 2 moderate pollution area (MPA), are located in the middle stations 2, 3, 4 and 5 which was influenced by parallel pollution sources as a result of the land use changed from forest to agriculture and unplanned settlements rural areas (Kulai River, Tayur River Telemong River and Las River). Deforestation, logging and mining activities, palm plantation, thus, the entrance of pollution increased to the level of compared to the upstream area during the study period, also, with the of that contained in runoff from the forest area [42]. In addition, station 4 receives pollution from mining and domestic effluents of areas, agricultural activities and domestic effluents of neighbouring areas, deforestation, wastewater discharge, and surface runoff. Increase of the flow rate, deforestation, palm plantation, runoff from agricultural fields have negatively impacted on the water quality at these stations. Station 5 was impacted by industrial activities, surface runoff, and wastewater from the residential areas around Nerus River. Moreover, station 5 appears to be largely influenced by the livestock husbandry, palm plantation, wastewater discharge, horticulture, runoff from agricultural fields and small dumps leeches placed on the banks. Thus, water quality of Nerus River was found to be highly polluted by livestock farming activities and waste from the settlements [12,13].

Cluster 3 stations (6, 7 and 8) are located in the downstream river section (Tong River, Linggi River, Kepayang River, Temiang River, and Semelang River), and considered as the high pollution area (HPA). Stations 6 and 7 are located at the end of the middle section, correspond to moderate polluted stations, which were affected by land use and anthropogenic activities. Station 6 with industrial wastes, wastewater discharge and erosion from urban runoff and agricultural land contributes to the high levels of pollution when combined with river surface water, and mining activities that take place around this station, as well as extensive urban-industrial-commercial land use. Station 7 was affected by industrial activities, runoff, and wastewater of the residential neighbourhoods surrounding the river. Station 7 was also influenced by deforestation and bridge construction and surface runoff. Increase of the flow rate, deforestation, palm plantation, areas of entertainment and playgrounds, runoff from agricultural fields and discharge of vehicles washing and workshops were negatively impacted on the water quality at station 7. Moreover, extensive urban-industrial-commercial land use, which is mixed urban and rural areas and commercial and government buildings at station 8 in downstream, also influenced by accumulated pollutants from the previous stations, discharges from vehicles washing and workshops, industries, construction of buildings, and settlements residential and commercial area [13,19,43]. Therefore, the minimum water quality was recorded at station 8 that received contamination can be attributed to several sources including (i) wastewater from sewage treatment plants septic systems, drainage fields on-site wastewater treatment and disposal systems, (ii) urban runoff from roads and (iii) agricultural runoff. However, the three different clusters along the Nerus River can be described as less pollution area (LPA),

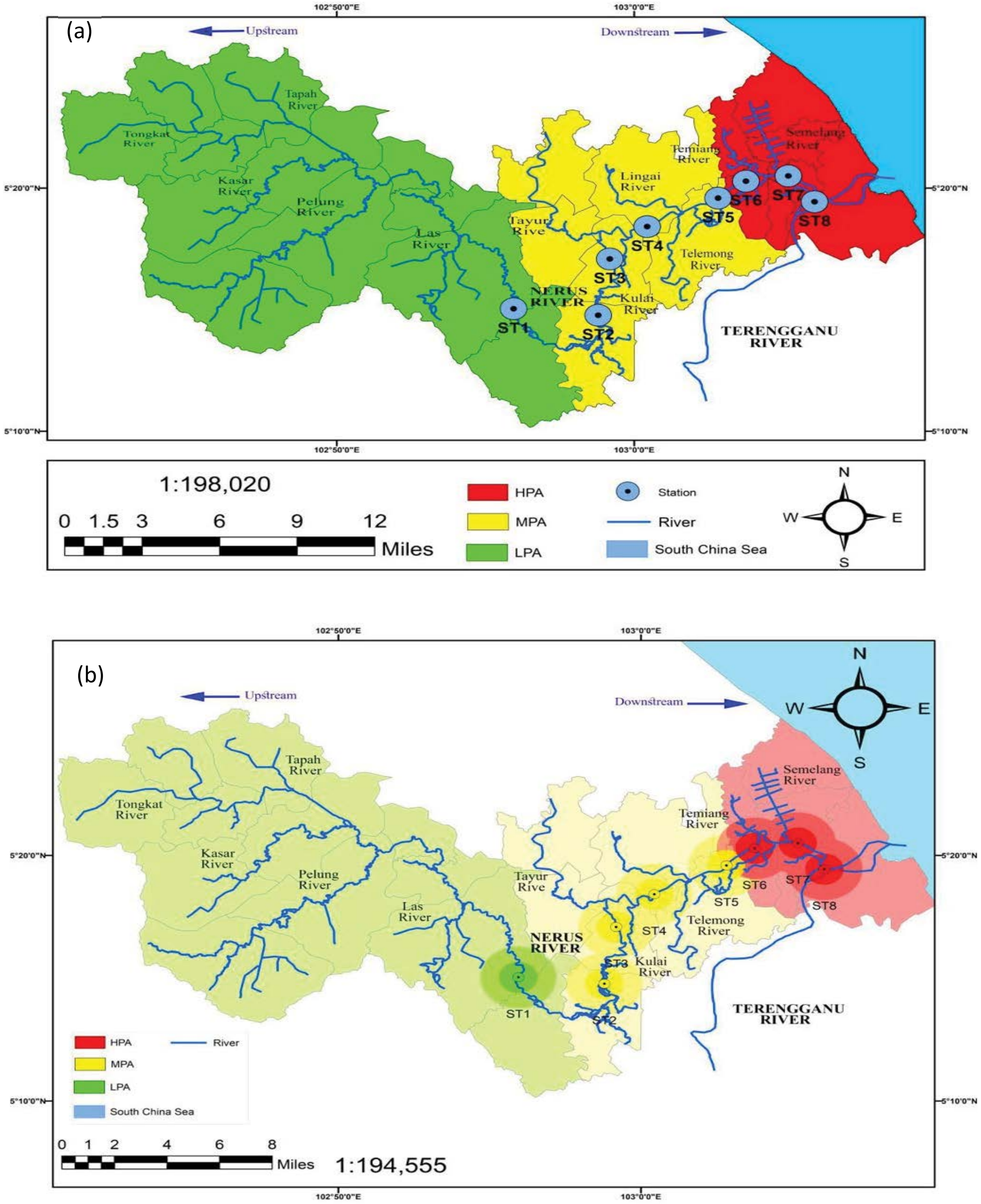

Fig. 4. Classification of regions due to surface river water quality by HACA for Nerus River basin.

moderate pollution area (MPA), or high pollution area (HPA) more polluted based on the measured values of pollution levels.

3.2. Spatial variations of Nerus River water quality

Discriminant analysis (DA) was used to determine the spatial variation among various sampling stations. Spatial variations for river water quality for 27 parameters were identified through stations–parameter correlation matrix. It was found that those parameters were significantly ($p < 0.01$) correlated with the spatial variations in water quality, where further evaluated through DA. The aims of the DA were to examine the significance of discriminant functions among clusters and identify the most significant variables of water quality in the study area. Subsequently, DA can be utilized as input parameters in prediction modelling. Besides, the findings from DA can optimize the monitoring procedures by decreasing the sampling parameters.

The parameters of water quality acted as independent variables, while the three significant groups (HPA, MPA, and LPA) were considered as dependent variables. In this study, DA was conducted on the original data based on three different modes to construct the best discriminant functions (DFs) for checking the clusters decided by HACA and evaluate the spatial variations. The accuracy of spatial classification was done using standard, forward stepwise and backward stepwise mode. In contrast, DFA were 68.09% of standard 7 discriminant variables, 76.60% of forward 3 discriminant variables, and 82.98% of backward 6 discriminant variables, respectively (Table 4). Using backward stepwise mode discriminant analysis, TSS, NO_2, TDS, Cu, Na and Ni were the most significant parameters for discriminating among the three groups (clusters) and accounted for most of the predictable spatial variation in water quality. This indicates that these parameters have high variation in terms of their spatial distribution (Table 5).

Fig. 5 shows box and whisker plots of discriminating parameters identified by spatial DA (backward stepwise mode) that was used to examine different patterns associated with spatial variations surface water quality in Nerus River. The average TSS is the highest in MPA because MPA close to road construction and residential area, considered as urbanized zones which carry domestic wastewater, industrial effluents and wastewater treatment plants. Anthropogenic activities like construction and residential activities highly affecting the TSS, especially during the wet season, as well as runoff from the upstream (forest) area towards the downstream. On the other hand, the decrease in TSS was observed in LPA and the significant decrease was also found in the HPA. Backward stepwise DA showed that NO_2 was the discriminating parameter showed by spatial and box and whisker plots of discriminating parameters determined by the spatial backward stepwise DA.

The NO_2 reveals elevation in the spatial in LPA, whereas there are significantly decrease of NO_2 in MPA and HPA. This is the result from the agricultural activities where nitrogenous fertilizers run off produced during the wet season in the study area. It is believed that nitrite–nitrogen might also present due to the anaerobic wastewaters, low DO levels and urban runoff containing poorly degraded organic and NH_3–N waste during the wet season, but the leaching of agricultural fertilizers has the most effect on the presence of NO_2 in river water.

The level of TDS was comparatively higher in sampling stations HPA of Nerus River. The level of TDS was also comparatively higher in HPA compared to LPA and MPA. At HPA, the increase in the level of TDS were attributed to the agricultural fields, agricultural activities, domestic sewage pollution, especially in the wet season. The TDS levels were found to be higher in HPA, reveals a subtle elevation during the wet compared with the dry season, as well as LPA and MPA. This could be a result of soil erosion from disturbed land or inflow of effluent from LPA and MPA by the runoff from these areas especially in the rainy season. This could be also explained by considering the chemical components of various anthropogenic activities that constitute point sources pollution especially from industrial, domestic and commercial and agricultural runoff areas located at the HPA in Nerus River.

Table 4
Classification matrix by DA for spatial variations in Nerus River

From/to	LPA	MPA	HPA	Total	% correct
Standard DA mode (7 variables)					
LPA	4	1	1	6	66.67%
MPA	3	16	5	24	66.67%
HPA	0	5	12	17	70.59%
Total	7	22	18	47	68.09%
Forward stepwise mode (3 variables)					
LPA	3	3	0	6	50.00%
MPA	2	22	0	24	91.67%
HPA	0	6	11	17	64.71%
Total	5	31	11	47	76.60%
Backward stepwise mode (**6** variables)					
LPA	5	1	0	6	83.33%
MPA	2	21	1	24	87.50%
HPA	0	4	13	17	76.47%
Total	7	26	14	47	**82.98%**

Table 5
Factor loadings of the LPA, MPA, and HPA by DA on varimax rotation in the Nerus River

Variable	Lambda	*F*	DF1	DF2	*p*-value
TSS (mg/L)	0.757	7.048	2	44	0.002
NO_2 (mg/L)	0.737	7.852	2	44	0.001
TDS (mg/L)	0.843	4.103	2	44	0.023
Cu (μg/L)	0.743	7.612	2	44	0.001
Na (μg/L)	0.541	18.665	2	44	<0.0001
Ni (μg/L)	0.870	3.296	2	44	0.046

The average concentration of Cu ions was observed to be higher in dry season and lower in wet season which was a manifestation of the dilution effect of rainfall and discharge. On the contrary, Na and Ni ions were observed to be higher in wet season and lower in dry season. Cu, Na and Ni levels were higher in HPA compared to the LPA and MPA. Further, downstream, HPA commonly had higher Na, Ni and Cu concentrations due to the continual inflow and natural disturbance of surface water at LPA and MPA due to the occurrence of heavy rainfall. Thus, rainfall intensity influences the Na, Ni and Cu concentrations of surface water. Finally, the agricultural activities and oil palm plantations, untreated sewage and wastewater from the rural areas could be resulted in the increasing of Na, Ni and Cu levels in HPA area.

3.3. Temporal variations of river water quality

Temporal variations in 27 samples from river water quality parameters were evaluated through a season–parameter correlation matrix, which shows that most of

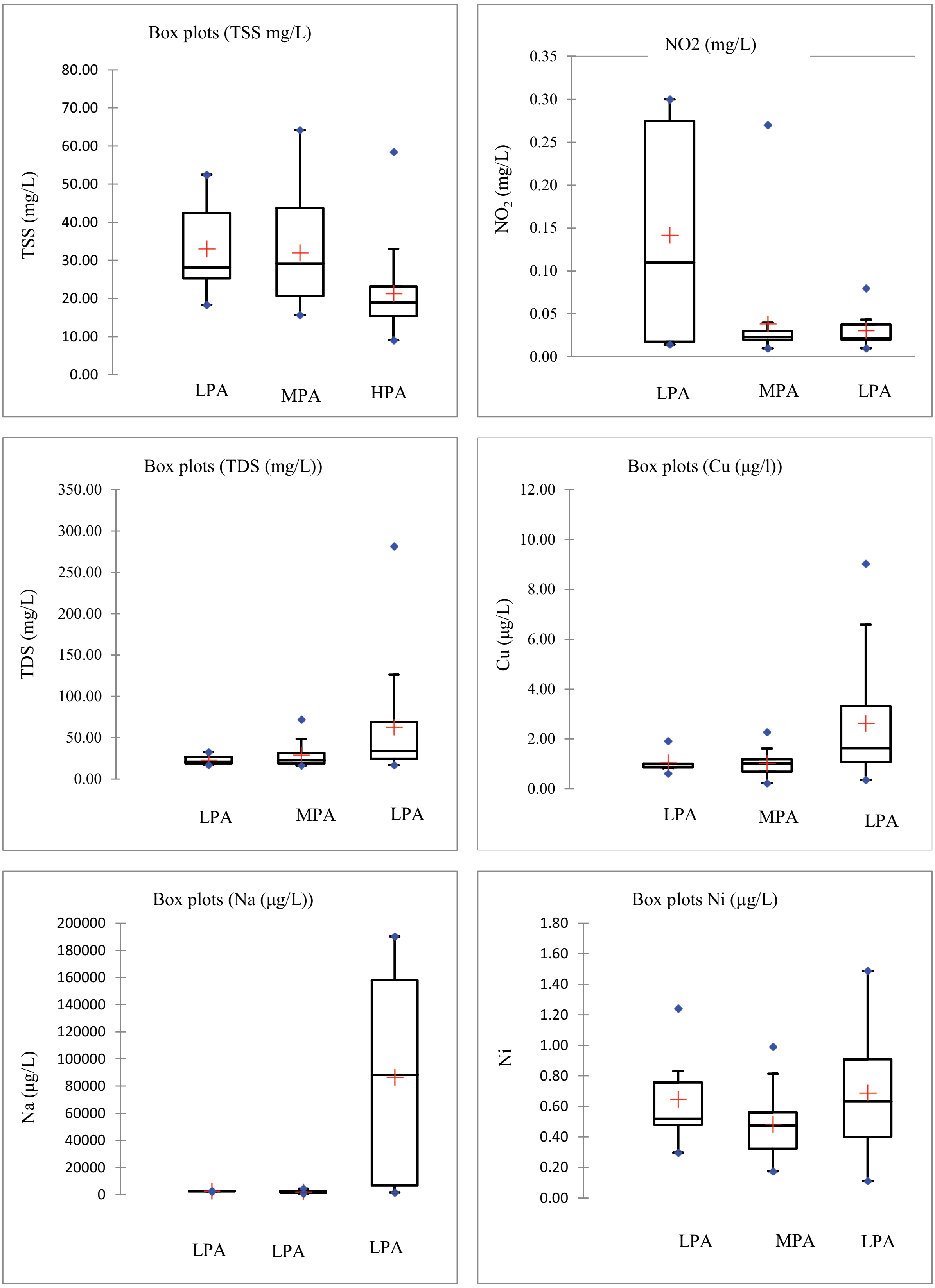

Fig. 5. Box and whisker plots of some parameters separated by spatial DA associated with the water quality data of Nerus River.

the measured parameters were found to be significantly correlated with the season ($p < 0.01$). DA was performed for temporal analysis on data after dividing the whole dataset into two seasonal groups (wet and dry). Further, the standard, forward stepwise and backward stepwise DFs using 27 variables. In contrast, standard (14 variables), forward (4 variables) and backward (10 variables) were significant variables, respectively. These standards, forward stepwise and backward stepwise mode DFs, respectively gave a classification matrix assigning 97.92%, 95.83%, and 100.00% cases correctly (Table 6). Thus, the temporal, DA results in standard DA Mode are suggested that Total forward stepwise TEMP, pH, DO, and PO_4 are the most significant parameters to discriminate between the wet and dry these four parameters account for most of the expected temporal variations (Table 7).

In contrast, a solid relationship between temperature and DO is also identified, due to the seasonal effect. Since evaporation process occur optimally when water achieved warmer temperature, causing water to have lack of capacity in retaining oxygen, proved that the relationship between temperature and DO is a natural process. A decreasing trend of pH can be observed in dry over wet season where enhanced weathering process during the dry season in the catchment area can be the cause of the pH variation. During wet season, the higher average of pH is possibly due to extreme dilution.

The season-correlated parameters can be used to signify the main sources of temporal variations in water quality. Whereas, the average TEMP, pH and DO are the highest during the wet season compared to the dry season as a result of the runoff from the upstream (forest) area towards the downstream. Forward stepwise DA shows that PO_4 is the discriminating parameter according to season. Fig. 6 shows box and whisker plots of discriminating parameters analysed by the seasonal forward stepwise DA. Further, the results of the temporal variations can be used to identify the polluted areas and set the priority areas for the river water quality management in the study area.

Table 6
Classification matrix by DA for temporal variations in Nerus River

From/to	Dry	Wet	Total	% correct
Standard DA mode (14 variables)				
Dry	23	1	24	95.83%
Wet	0	24	24	100.00%
Total	23	25	48	97.92%
Forward stepwise mode (**4** variables)				
Dry	24	0	24	100.00%
Wet	2	22	24	91.67%
Total	26	22	48	**95.83%**
Backward stepwise mode (10 variables)				
Dry	24	0	24	100.00%
Wet	0	24	24	100.00%
Total	24	24	48	100.00%

Table 7
Classification matrix by DA for temporal variation in Nerus River

Variable	Lambda	F	DF1	DF2	p-value
Temp. (°C)	0.787	12.418	1	46	0.001
pH	0.555	36.953	1	46	<0.0001
DO (mg/L)	0.825	9.773	1	46	0.003
PO_4 (mg/L)	0.872	6.758	1	46	0.013

3.4. Principal component analysis

In this study, the data set of 27 water quality variables encompasses nine factors, loading values and percentage of variance for all stations listed in Table 8. PCA produced nine significant PCs were obtained for the study area with total variance of 82.237% and Eigen values > 1. In addition, factor loading was classified based on loading values to be strong (greater than 0.75), moderate (0.75–0.50) and weak (0.50–0.30) [44]. The first accounting for (16.515%), the total variance explaining (82.237%) of the total variance in the corresponding water quality data sets.

Factor 1 represents 16.515% of the total variance, showing strong positive loadings on Cd, Cr, Cd, Cu and Zn. Furthermore, it shows moderate positive loading on COD and Co. These parameters are mainly represented the loading associated with increased pollution runoff in the form of untreated sewage, industrial waste and agricultural runoff. Pollution is also originated from agriculture and the use of fertilizers and agricultural pesticides that seep into the river by runoff surface irrigators, mining activity, river's silt as well as organic and inorganic contaminations from domestic (village) effluent. Point and non-point sources of contamination are associated with anthropogenic input and land use activities. Also, it could be contributed to the pollution. In contrast, this factor indicates negative loadings on pH, DO, BOD, TSS, TUR, PO_4, NO_2 and *E. coli*. Moreover, temperature is most possibly associated with seasonal influences of solar radiation from the early morning to the afternoon that affects the temperature. The factor loadings between each element are affected by the presence of the other element. *E. coli* affected by BOD, DO and other elements [45]. The positive loadings on Cr, Cd and Cu ions could be attributed to weathering of soils and rocks, industrial and municipal activities [46].

Factors 2, 3 and 4 are represented 9.409%, 10.067% and 10.720% of the total variance, respectively. These factors have strong positive loadings on SO_4 and SAL in factor 2. Strong positive loadings on BOD, NH_3–N and TUR are observed in factor 4, while a strong negative loading on DO is obtained in factor 3. This factor could thus be an indicator of anthropogenic activities related to different types of land use [47,48]. Whereas, positive loadings on SAL, NH_3–N, BOD and TUR could be attributed to the anthropogenic sources. Particularly, the organic pollutants from point sources such as the discharge of domestic wastewater, water treatment plants, livestock farms and untreated

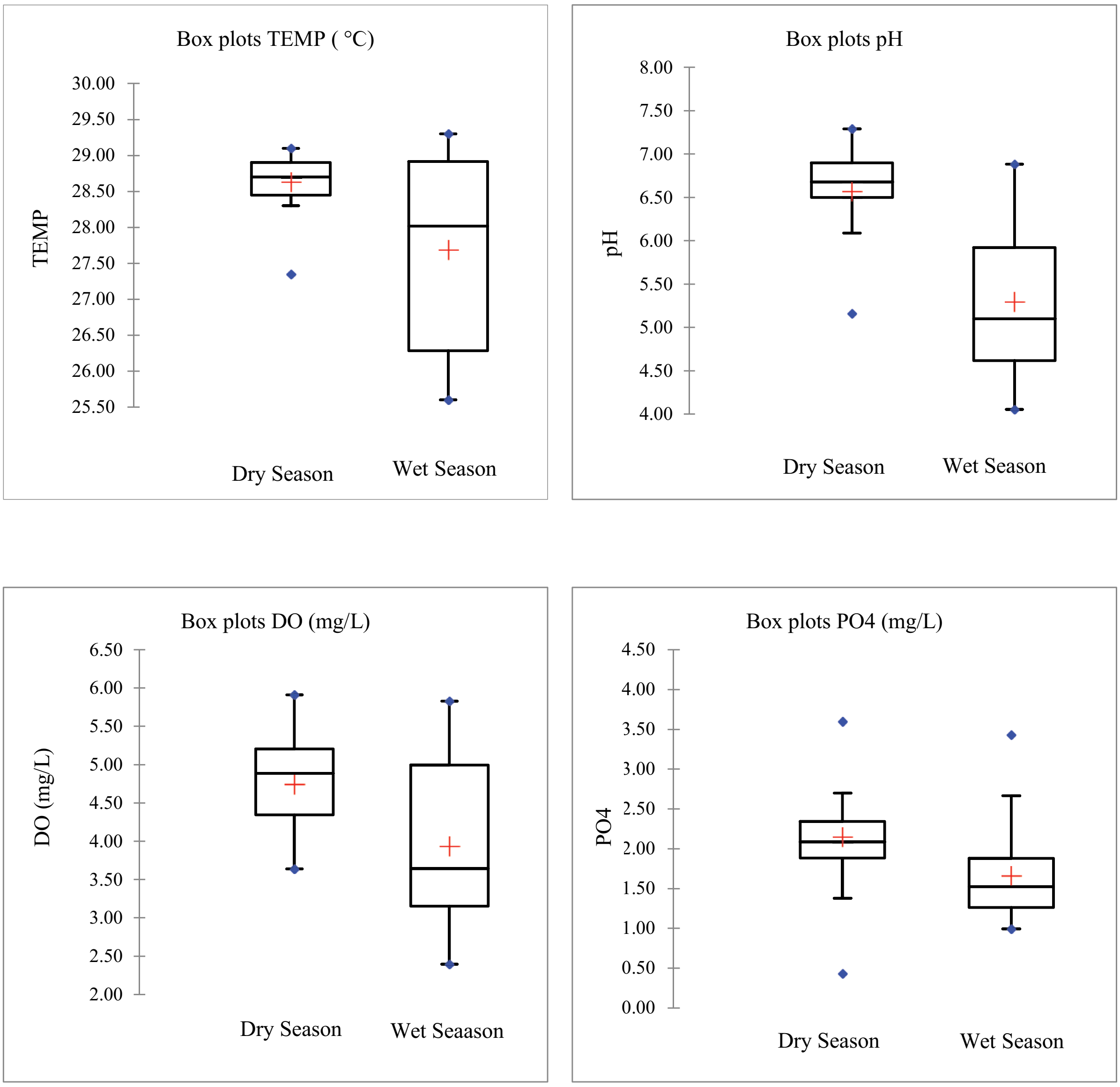

Fig. 6. Box and whisker plots of some parameters separated by temporal DA associated with the water quality data of Nerus River.

sewage of rural settlement areas as well as industrial effluents were reported [38]. A strong loading of BOD and NH_3–N is typically related to organic sources of pollution with organic matter levels, which are extremely high in runoff in areas dominated by horticultural activities. The TUR and SAL contents of water are directly related to each other and are very much connected to agricultural activities and domestic wastewater, while agricultural sites differ from urban sites in terms of exhibiting higher water TUR [49].

Factors 5 and 6 are stands for 6.564% and 6.615% of the total variance and have a moderate positive loading on Pb in factor 5. Furthermore, the moderate positive loading of Pb ions is associated with industrial wastes, industrial and municipal activities. Pb level in natural water increases mainly through anthropogenic activities related to municipal, industrial wastes [50] and the automobile exhausts [51,52]. In addition, the presence of Na ions in Factor 6 could be attributed to anthropogenic activities such as mining. The Na compounds is applied in agricultural activities, in untreated sewage and untreated wastewater. In contrast, factors 5 and 6 indicate strong negative loadings on *E. coli* and TSS [53].

Factor 7 stands for 7.131% of the total variance with strong positive loading on Mn and moderate positive loading of Fe. The strong positive loading of Mn is associated with industrial wastes. Moreover, the presence of Fe in waterways is associated with the bedrocks weathering, soils erosion, as well as the direct discharge from residential area and industrial operations, leakage from polluted sites and landfills and the use of sludge and fertilizers in agriculture. Besides, the presence of Mn and Fe indicates pollution from anthropogenic sources because of the industrial effluents, agricultural runoff and animal's farm waste causes Co pollution in receiving water [54].

Table 9
Loadings of the pollutant variable on the varimax-rotated PCs for water quality data collected from LPA, MPA, and HPA in Nerus River

Variables	LPA		MPA			HPA		
	VF1	VF2	VF1	VF2	VF3	VF1	VF2	VF3
TSS (mg/L)	0.074	**–0.942**	–0.087	–0.718	–0.387	0.139	–0.564	–0.493
NO_2 (mg/L)	–0.876	0.205	–0.003	**0.829**	–0.175	0.001	0.723	0.308
TDS (mg/L)	**–0.786**	0.126	0.016	–0.077	**0.885**	0.175	–0.058	**0.889**
Cu (µg/L)	**0.861**	0.319	**0.902**	–0.076	0.007	**0.970**	–0.054	0.014
Na (µg/L)	**–0.184**	0.875	0.481	0.372	0.628	0.000	**0.846**	–0.262
Ni (µg/L)	**0.914**	0.258	**0.933**	0.119	0.130	**0.960**	0.022	0.135
Variability (%)	**49.640**	**31.670**	**32.059**	**22.794**	**22.911**	**31.873**	**26.058**	**20.262**
Cumulative (%)	**49.640**	**81.310**	**32.059**	**54.854**	**77.764**	**31.873**	**57.931**	**78.193**

Factor 8 elucidates 4.440% of the total variance with strong positive loading on NO_2, ascribed to the anthropogenic sources such as urban effluents, which then becomes a pollutant of aquatic ecosystems and thus harmful to aquatic organisms [52]. Whereas, likely attributable to organic pollutants derived from industrial and domestic waste [55].

Factor 9 elucidates 10.775% of the total variance with strong positive loading on COND and TDS. The presence of COND and TDS are due to agricultural runoff, such as livestock waste and fertilizers [56]. COND and TDS are important constituents of detergents that discharge into the river by municipal sewage, industrial effluents and water treatment plants [57]. On the other hand, COND and TDS are attributed to natural sources and anthropogenic activities. The presence of TDS is associated with rocks weathering and discharges from industrial areas and water treatment plants as well as leaching of soil contamination and agricultural and residential runoff that are the primary sources of TDS in rivers water [58]. Additionally, as well as COND is released to water by discharge from industrial facilities or from landfills and soil. The above factors obtained from factor analysis indicate that the parameters responsible for water quality variations in the Nerus River are mainly related to urban and industrial activities, domestic waste, as well as agricultural activities such as the use of PO_4 and nitrogen fertilisers [44].

3.5. Identification of source of variation in HPA, MPA and LPA using PCA technique

In this study, PCA technique was applied to the water quality data set to determine the major sources of variation in each cluster (HPA, MPA and LPA) produced by HACA. Six selected water quality parameters that gave high variations (the most significant) by backward stepwise DA were then used for further discussion. PCA was employed on the data set to compare the compositional patterns between the examined water parameters and to identify the factors that influence each of the identified regions (HPS, MPS and LPS). Parameters of water quality were tested for each cluster include TSS, NO_2, TDS, Cu, Na and Ni. From the findings, two VFs have been chosen from LPA and three VFs have been chosen from MPA and HPA with the eigenvalues larger than 1. Whereas, the percent cumulative of total variance in the data sets in LPA, MPA and HPA were 81.310%, 77.764% and 78.193%, respectively (Table 9). In this study, only strong factor loadings (>0.75) were selected for the PC interpretation. The pollutants such as TDS, Cu and Ni were presented in each of clusters (LPA, MPA and HPA) as the addition to the variables water pollutants.

3.5.1. Less pollution area

Less pollution area (LPA) consists of one station 1 with the first factor VFl in this cluster explains 49.640% of the total variance, shows strong positive factor loadings for Cu (0.861) and Ni (0.914), but negative loadings on NO_2 (–0.876) and TDS (–0.786). This factor contains chemical parameters and those attributed to products from anthropogenic activities and can be identified to be originated from both wastewater treatment plants (PS) and (NPS) pollution sources [19] (Table 9). The presence of Cu and Ni is related to the influence of domestic waste and agricultural runoff [20].The presence of Cu and Ni may be possibly linked to parent rock materials (earth crust) [59]. The second factor VF2 with 31.670% of the total variance shows strong positive loadings for Na (0.875) and negative loadings on TSS (–0.942), thus represents the influence of organic pollutants from point sources such as discharge from wastewater treatment plants, domestic wastewater and industrial effluents.

3.5.2. Moderate pollution area

Moderate pollution area (MPA) involves stations 2, 3, 4 and 5 with VF1 explains 32.059% of the total variance and has strong positive loadings on Ni (0.933) and Cu (0.902). Whereas, this factor could be explained by considering the chemical components of various anthropogenic activities which constitute pollution especially from domestic, commercial and agricultural runoff (Table 10). The second factor (VF2) explains 22.794% of the total variance, shows a strong positive factor loading for NO_2 (0.829). This is possibly due to agricultural runoff because NO_2 is commonly used in agricultural practice. VF3 explains 22.911% of the total

Table 8
Loadings of the 27 variables on varimax rotation in the Nerus River factor

Factor loadings after varimax rotation									
	F1	F2	F3	F4	F5	F6	F7	F8	F9
Temp. (°C)	0.167	–0.104	0.350	**–0.751**	–0.326	0.113	–0.176	0.029	0.175
pH (unit)	–0.103	0.279	–0.004	–0.671	0.039	0.076	–0.080	0.186	–0.410
DO (% Sat)	0.037	–0.108	**–0.841**	0.024	–0.263	–0.018	–0.135	0.052	–0.336
DO (mg/L)	–0.056	–0.141	**–0.874**	0.055	0.001	0.010	–0.081	0.008	–0.270
BOD (mg/L)	–0.048	–0.333	0.384	**0.719**	0.088	0.071	0.073	0.090	0.001
COD (mg/L)	**0.521**	–0.541	0.121	–0.034	0.096	–0.061	–0.167	0.062	0.394
TSS (mg/L)	–0.113	–0.109	–0.012	0.046	0.210	**–0.828**	–0.087	–0.111	–0.174
NH_3–N (mg/L)	0.130	–0.110	–0.154	**0.738**	0.078	0.232	–0.094	0.008	0.172
NO_2 (mg/L)	–0.072	0.051	–0.038	0.082	–0.007	0.050	–0.052	**0.927**	–0.022
NO_3 (mg/L)	0.421	–0.161	0.443	0.433	–0.277	–0.152	–0.084	0.187	0.145
PO_4 (mg/L)	–0.036	–0.381	–0.430	–0.107	0.241	0.529	0.187	0.124	–0.186
SO_4 (mg/L)	0.034	**0.761**	0.317	–0.030	0.021	0.048	–0.051	0.099	0.469
TUR (NTU)	–0.136	0.287	0.109	**0.730**	–0.246	–0.238	0.162	0.303	–0.032
COND (uS)	0.181	0.103	0.267	0.056	–0.027	0.057	0.028	–0.017	**0.901**
TDS (mg/L)	0.190	0.146	0.276	0.045	–0.044	0.086	0.016	–0.018	**0.892**
SAL (ppt)	0.157	**0.849**	0.046	–0.124	–0.094	0.128	–0.106	0.037	0.069
O&G (mg/L)	0.390	–0.345	0.437	0.170	–0.185	0.080	–0.199	–0.085	0.289
Cd (µg/L)	**0.861**	–0.033	0.151	–0.028	–0.071	–0.062	0.208	–0.024	0.059
Co (µg/L)	**0.544**	–0.165	0.092	0.070	–0.036	0.050	0.606	–0.123	0.404
Cr (µg/L)	**0.870**	0.060	–0.088	–0.013	–0.007	0.228	–0.037	–0.036	0.038
Cu (µg/L)	**0.760**	0.301	0.068	0.044	0.315	0.069	–0.021	–0.117	0.093
Fe (µg/L)	0.209	–0.049	–0.221	0.052	0.352	–0.310	**0.633**	0.113	0.177
Mn (µg/L)	0.005	–0.027	0.166	0.128	0.029	0.091	**0.922**	–0.054	–0.086
Na (µg/L)	0.040	0.330	0.118	–0.008	0.337	0.630	–0.210	–0.178	0.082
Ni (µg/L)	**0.864**	0.151	–0.019	–0.037	0.167	0.095	0.010	–0.077	0.096
Pb (µg/L)	0.299	–0.002	0.244	–0.076	**0.700**	0.112	0.187	–0.181	0.072
Zn (µg/L)	**0.759**	–0.260	–0.002	–0.072	–0.079	–0.262	0.128	0.126	0.388
E. coli (CFU/100 mL)	–0.012	0.234	–0.090	–0.299	**–0.696**	0.181	–0.038	–0.165	0.240
Variance (%)	**16.515**	**9.409**	**10.067**	**10.720**	**6.564**	**6.615**	**7.131**	**4.440**	**10.775**
Cumulative (%)	**16.515**	**25.924**	**35.991**	**46.711**	**53.275**	**59.890**	**67.021**	**71.461**	**82.237**

variance and has strong positive loading on TDS. Thus, TDS increments may be due to overland inputs, increased streambank erosion and increased entrainment of bedload sediments during stormflow especially in forested area [13].

3.5.3. High pollution area

High pollution area (HPA) consists of three stations 6, 7, and 8 in downstream with the first factor VFl in this cluster explains (31.873%) of the total variance, shows strong positive factor loadings for Ni (0.960) and Cu (0.970). This loading was contributed by the construction site and agricultural runoff. Agricultural and construction sites are more common near watercourses and as a result of these activities, the sediment was deposited.VF2 explains 26.058% of the total variance and has strong positive loading on Na. This is due to the present of buildings and houses in the area, anthropogenic activities and wastewater treatment plants (Table 10). Finally, the third factor VF3 explains 20.262% of the total variance, shows a strong positive factor loading for TDS (0.889). This factor, within the HPA region caused by the high load of soil and waste disposal activities, industrial activities and certain industrial processes and residential areas, urban developments, wastewater treatment plants, domestic wastewater, and recreation activities along the river. All of these factors contributed to the increase in the loading rate TDS along the river. It was identified that urban land use is a major pollution source contributing to the HPA.

4. Conclusions

The present study applied principal component analysis technique to evaluate the potential impact of natural and anthropogenic point and non-point sources of pollution on the Nerus River. The HACA rendered the sampling stations into 3 clusters LPA, MPA and HPA based on similar characteristics of water quality. Certain stations should

Table 10
Results of pollution sources identification at LPA, MPA and HPA region.

Regions	V1	V2	V3
LPA	Natural pollution, organic pollution	Organic pollution, chemical pollution	
MPA	Anthropogenic sources, organic pollution, chemical pollution	Anthropogenic sources, organic pollution	Natural pollution
HPA	Anthropogenic sources, chemical pollution	Anthropogenic sources	Natural pollution, anthropogenic sources

be emphasized if treatment efforts are to be undertaken as per the cluster results recommended. On the other side, DA was used to examine the spatial and temporal variation in river water quality. The spatial DA discovered significant data reduction as it gives six parameters (TSS, NO_2, TDS, Cu, Na and Ni) with 982.98% correct assignation. While, temporal DA revealed significant data reduction as it gives four parameters (TEMP, pH, DO and PO_4) with 95.83% correct assignation. In contrast, PCA/FA identified nine latent factors that explained more than 82.237 of the total variances of 27 parameters. PCA identify the major sources of water pollution, and it was found that the major contribution was from urban land use and other non-point sources of pollution might explain the low water quality recorded throughout the study area. This finding confirms that one of the greatest causes of water quality problems is derived from the increasing intensity of human activities taking place in the region. Although variance was acceptable for a domestic and agricultural use, it was found to be high for certain parameters, particularly at anthropogenic activities sites. Areas of varying land use could be differentiated on the basis of these water quality parameters with rainfall events contribute substantially to major pollutant loading via surface runoff input. Nerus River has received high level of pollutants from rural, agricultural, urban and minor industrial land use areas in the form of discharged waste. This study shows the effectiveness of multivariate statistical methods for analysis and interpretation of complex data sets, and also for determination sources of pollution and better comprehension of spatial and temporal variations in water quality for effective and sustainable river water quality management.

Acknowledgments

The author would like to thank UniSZA and KPM for providing financial support for this research on the SRGS: (UniSZA/2017/SRGS/17) – R0019-R017, FRGS 2017 (FRGS/1/2017/WAB05/UNISZA/01/1) – RR22 and Department of Irrigation and Drainage Malaysia for the secondary data and East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin (UNISZA) for giving permission to use research facilities and supporting in this research.

References

[1] J. Das, B.C. Acharya, Hydrology and assessment of lotic water quality in Cuttack City, India, Water, Air Soil Pollut., 150 (2003) 163–175.
[2] J. Yisa, T. Jimoh, Analytical studies on water quality index of River Landzu, Am. J. Appl. Sci., 7 (2010) 453–458.
[3] A. Mahmood, W. Muqbool, M.W. Mumtaz, F. Ahmad, Application of multivariate statistical techniques for the characterization of ground water quality of Lahore, Gujranwala and Sialkot (Pakistan), Pak. J. Anal. Environ. Chem., 12 (2011) 102–112.
[4] A. Azid, C.N. Che Hasnam, H. Juahir, M.A. Amran, M.E. Toriman, M.K.A. Kamarudin, A.S.M. Saudi, M.B. Gasim, A.D. Mustafa, Coastal erosion measurement along Tanjung Lumpur to Cherok Paloh, Pahang during the northeast monsoon season, J. Teknol., 74 (2015) 27–34.
[5] S. Shrestha, F. Kazama, Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji River basin, Japan, Environ. Modell. Software, 22 (2007) 464–475.
[6] M.K.A. Kamarudin, M.E. Toriman, M.H. Rosli, H. Juahir, N.A.A. Aziz, A. Azid, S.F.M. Zainuddin, W.N.A. Sulaiman, Analysis of meander evolution studies on effect from land use and climate change at upstream reach of Pahang River, Malaysia, Mitigation Adapt. Strategies Global Change, 20 (2015) 1319–1334.
[7] S. Suratman, A.N.A.R. Hussein, N. Mohd Tahir, M.T. Latif, R. Mostapa, K. Weston, Seasonal and spatial variability of selected surface water quality parameters in Setiu wetland, Terengganu, Malaysia, Sains Malaysiana, 45 (2016) 551–558.
[8] S. Muangthong, Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River Basin, Thailand, J. Ind. Technol., 11 (2015) 25–37.
[9] M.K.A. Kamarudin, M.E. Toriman, N.A. Wahab, H. Juahir, A. Endut, R. Umar, M.B. Gasim, Development of stream classification system on tropical areas with statistical approval in Pahang River Basin, Malaysia, Desal. Water Treat., 96 (2017) 237–254.
[10] N.A. Wahab, M.K.A. Kamarudin, M.B. Gasim, R. Umar, F.M. Ata, N.H. Sulaiman, Assessment of total suspended sediment and bed sediment grains in upstream areas of Lata Berangin, Terengganu, Int. J. Adv. Sci., Eng. Inf. Technol., 6 (2016) 757–763.
[11] M.H.M. Nazir, W.N.A. Sulaiman, H. Juahir, Hydrologic response characteristics of a tropical catchment to land use changes: a case study of The Nerus catchment, Environ. Earth Sci., 73 (2015) 7533–7545.
[12] H.M.A. Alssgeer, M.B. Gasim, M.M. Hanafiah, E.R.A. Abdulhadi, A. Azid, GIS-based analysis of water quality deterioration in the Nerus River, Kuala Terengganu, Malaysia, Desal. Water Treat., 112 (2018) 334–343.
[13] M.E. Toriman, H.M.A. Alssgeer, M.B. Gasim, K.A. Kamarudin, M.M. Daw, L.O.M. Alabyad, Impacts of land-use changes on water quality by an application of GIS analysis: a case study of Nerus River, Terengganu, Malaysia, Int. J. Eng. Technol,. (UAE), 7 (2018) 155–164.
[14] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y. Michotte, Kaufman, Chemometrices: A Textbook, Elsevier, Amsterdam, 1988.
[15] M. Camara, N.R. Jamil, A.F.B. Abdullah, Impact of land uses on water quality in Malaysia: a review, Ecol. Processes, 8 (2019) 1–10.
[16] K.P. Singh, A. Malik, D. Mohan, S. Sinha, Multivariate statistical techniques for the evaluation of spatial and temporal variations

in water quality of Gomti River (India)—a case study, Water Res., 38 (2004) 3980–3992.
[17] S. Shrestha, S. Muangthong, Assessment of surface water quality of Songkhram River (Thailand) using environmetric techniques, Int. J. River Basin Manage., 12 (2014) 341–356.
[18] N.A.M. Shazili, H.M. Sharipudin, N. Tahir, K. Yunus, Dissolved Cd, Cu, Pb and Zn concentration in Sungai Nerus, Terengganu, Malaysian J. Anal. Sci., 8 (2004) 112–117.
[19] F. Azaman, H. Juahir, K. Yunus, A. Azid, M.K.A. Kamarudin, M.E. Toriman, A.D. Mustafa, M.A. Amran, C.N.C. Hasnama, A.S.M. Saudi, Heavy metal in fish: analysis and human health-a review, J. Teknol., 77 (2015) 61–69.
[20] S.C. Poh, S. Suratman, C.K. Chew, N.A.M. Shazili, N.M. Tahir, Metal geochemistry of Nerus River, Terengganu, Malaysian J. Anal. Sci., 12 (2008) 593–599.
[21] S. Suratman, M.M. Sailan, Y.Y. Hee, E.A. Bedurus, M.T. Latif, A preliminary study of water quality index in Terengganu River basin, Malaysia, Sains Malaysiana, 44 (2015) 67–73.
[22] A. Ibrahim, H. Juahir, M.E. Toriman, M.K. Kamarudin, H.A. Isiyaka, Surface water quality assessment of Terengganu River Basin using multivariate techniques, Adv. Environ. Biol., 8 (2014) 48–58.
[23] M.H. Sayadi, Impact of land use on the distribution of toxic metals in surface soils in Birjand city, Iran, Proc. Int. Acad. Ecol. Environ. Sci., 4 (2014) 18–29.
[24] P. Simeonova, V. Simeonova, G. Andreev, Environmetric analysis of the Struma River water quality central, Eur. J. Chem., 2 (2003) 121–126.
[25] H. Zhang, C.Y. Jim, Species adoption for sustainable forestry in Hong Kong's degraded countryside, Int. J. Sustainable Dev. World Ecol., 20 (2013) 484–503.
[26] K.P. Singh, A. Malik, S. Sinha, Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques—a case study, Anal. Chim. Acta, 538 (2005) 355–374.
[27] K. Manoj, P.K. Padhy, Multivariate statistical techniques and water quality assessment: discourse and review on some analytical models, Int. J. Environ. Sci., 5 (2014) 607–626.
[28] K. Manoj, S. Ghosh, P.K. Padhy, Characterization and classification of hydrochemistry using multivariate graphical and hydrostatistical techniques, Res. J. Chem. Sci., 3 (2013) 32–42.
[29] T. Jombart, S. Devillard, F. Balloux, Discriminant analysis of principal components: a new method for the analysis of genetically structured populations, BMC Genet., 11 (2010) 94–109.
[30] P.R. Kannel, S. Lee, S.R. Kanel, S.P. Khan, Chemometrics application in classification and assessment of monitoring locations of an urban river system, Anal. Chim. Acta, 582 (2007) 390–399.
[31] S. Shrestha, F. Kazama, T. Nakamura, Use of principal component analysis, factor analysis and discriminant analysis to evaluate spatial and temporal variations in water quality of the Mekong River, J. Hydroinf., 10 (2008) 43–56.
[32] B. Helena, R. Pardo, M. Vega, E. Barrado, J.M. Fernandez, L. Fernandez, Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis, Water Res., 34 (2000) 807–816.
[33] R. Reghunath, T.R.S. Murthy, B.R. Raghavan, The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India, Water Res., 36 (2002) 2437–2442.
[34] D. Dominick, H. Juahir, M.T. Latif, S.M. Zain, A.Z. Aris, Spatial assessment of air quality patterns in Malaysia using multivariate analysis, Atmos. Environ., 60 (2012) 172–181.
[35] H. Juahir, M.S. Zain, M.K. Yusoff, T.I. Hanidza, A.S. Mohd, M.E. Toriman, M. Mokhtar, Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques, Environ. Monit. Assess., 173 (2011) 625–641.
[36] D. Phung, C. Huang, S. Rutherford, F. Dwirahmadi, C. Chu, X. Wang, T.A.D. Dinh, Temporal and spatial assessment of river surface water quality using multivariate statistical techniques: a study in Can Tho City, a Mekong Delta area, Vietnam, Environ. Monit. Assess., 187 (2015) 229, doi: 10.1007/s10661-015-4474-x.
[37] C. Vialle, C. Ablayrolles, M. Lovera, C. Jacob, C.M. Montrejaud-Vignoles, Monitoring of water quality from roof runoff: Interpretation using multivariate analysis, Water Res., 45 (2011) 3765–3775.
[38] F. Zhou, G.H. Huang, H. Guo, W. Zhang, Z. Hao, Spatio-temporal patterns and source apportionment of coastal water pollution in eastern Hong Kong, Water Res., 41 (2007) 3429–3439.
[39] E.M. De Andrade, H.A.Q. Palácio, I.H. Souza, R.A. de Oliveira Leão, M.J. Guerreiro, Land use effects in groundwater composition of an alluvial aquifer (Trussu River, Brazil) by multivariate techniques, Environ. Res., 106 (2008) 170–177.
[40] H. Bu, X. Tan, S. Li, Q. Zhang, Temporal and spatial variations of water quality in the Jinshui River of the South Qinling Mts., China, Ecotoxicol. Environ. Saf., 73 (2010) 907–913.
[41] A.N. Kamarudzaman, V.K. Feng, R.A. Aziz, M.F.A. Jalil, Study of Point and Non-Point Sources Pollution – A Case Study of Timah Tasoh Lake in Perlis, Malaysia, International Conference on Environmental and Computer Science, Singapore, 2011.
[42] L.I. Ruzhong, S.H.U. Kun, L.U.O. Yueying, S.H.I. Yong, Assessment of Heavy Metal Pollution in Estuarine Surface Sediments of Tangxi River in Chaohu Lake Basin, Chin. Geogr. Sci., 20 (2010) 9–17.
[43] N.M. Abdullah, M.E. Toriman, H.M. Din, N.A.A. Aziz, M.K.A. Kamarudin, N.S.A. Rani, NF.M. Ata, M.H. Saad, N.W. Abdullah, M. Idris, N.R. Jamil, Influence of spatial and temporal factors in determining rainfall interception at Dipterocarp Forest Canopy, Lake Chini, Pahang, Malaysian J. Anal. Sci., 17 (2013) 11–23.
[44] M. Li, Q. Zhu, C.-W. Hu, L. Chen, Z.-L. Liu, Z.-M. Kong, Cobalt and manganese stress in the microalga *Pavlova viridis* (Prymnesiophyceae): effects on lipid peroxidation and antioxidant enzymes, J. Environ. Sci., 19 (2007) 1330–1335.
[45] A. Mustapha, A.Z. Aris, H. Juahir, M.F. Ramli, N.U. Kura, River water quality assessment using environmentric techniques: case study of Jakara River Basin, Environ. Sci. Pollut. Res., 20 (2013) 5630–5644.
[46] R.N. Kumar, R. Solanki, J.N. Kumar, Seasonal variation in heavy metal contamination in water and sediments of river Sabarmati and Kharicut canal at Ahmedabad, Gujarat, Environ. Monit. Assess., 185 (2013) 359–368.
[47] I. Bruneau, D.R. Corkal, E. Pietroniro, B. Tothand, G.V.D. Kamp, Human activities and water use in the South Saskatchewan River Basin, Prairie Forum, 34 (2009) 129–152.
[48] I.O. Akpan, E.A. Thompson, Assessment of heavy metal contamination of sediments along the cross-river channel in Cross River state, Nigeria, J. Environ. Sci. Toxicol. Food Technol., 2 (2013) 20–28.
[49] G. Walsh, V. Wepener, The influence of land use on water quality and diatom community structures in urban and agriculturally stressed rivers, Water, 35 (2009) 579–594.
[50] J.K. Nzeve, S.G. Njuguna, E.C. Kitur, Assessment of heavy metal contamination in surface water of Masinga Reservoir, Kenya, J. Nat. Sci. Res., 5 (2015) 101–108.
[51] M. Shuhaimi-Othman, M.B. Gasim, Heavy metals concentrations in water of Sungai Semenyih Watershed, Selangor Malaysia, Sains Malaysiana, 34 (2005) 49–54.
[52] N.E. Gray, Drinking Water Quality Problems and Solutions, 2nd ed., University of Dublin, 2008.
[53] M. Cempel, G. Nikel, Nickel: a review of its sources and environmental toxicology, Pol. J. Environ. Stud., 15 (2006) 375–382.
[54] M. Varol, B. Şen, Assessment of nutrient and heavy metal contamination in surface water and sediments of the upper Tigris River, Turkey, Catena, 92 (2012) 1–10.
[55] I.S. Lotfinasabas, Y.R. Gunale, N.S. Rajurkar, Water quality assessment of Alibaug Mangrove Forest using multivariate statistical technique, Maharashtra, India, Indian J. Geo-Mar. Sci., 42 (2013) 915–923.
[56] E.M. Al-Sabahi, A Study of Surface Water and Groundwater Pollution in Ibb City, Yemen, Ph.D. Thesis (Unpublished), Universiti Kebangsaan Malaysia, 2007.

[57] M. Vega, R. Pardo, E. Barrado, L. Deban, Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis, Water Res., 32 (1998) 3581–3592.
[58] R.F. Spellman, Water and Wastewater Treatment Plant Operations, CH 14, CRC Press, Boca Raton, 2003.
[59] R.A. Zampella, N.A. Procopio, R.G. Lathrop, C.L. Dow, Relationship of land-use/land-cover patterns and surface-water quality in the Mullica River Basin, JAWRA J. Am. Water Resour. Assoc., 43 (2007) 594–604.