Environ Monit Assess (2009) 153:179–185
DOI 10.1007/s10661-008-0347-x

Analysis of heavy metal concentrations in sediments of selected estuaries of Malaysia—a statistical assessment

Abbas F. M. Alkarkhi · Norli Ismail · Aness Ahmed · Azhar mat Easa

Received: 5 October 2007 / Accepted: 14 April 2008 / Published online: 27 May 2008

Abstract Statistical analysis of heavy metal concentrations in sediment was studied to understand the interrelationship between different parameters and also to identify probable source component in order to explain the pollution status of selected estuaries. Concentrations of heavy metals (Cu, Zn, Cd, Fe, Pb, Cr, Hg and Mn) were analyzed in sediments from Juru and Jejawi Estuaries in Malaysia with ten sampling points of each estuary. The results of multivariate statistical techniques showed that the two regions have different characteristics in terms of heavy metals selected and indicates that each region receives pollution from different sources. The results also showed that Fe, Mn, Cd, Hg, and Cu are responsible for large spatial variations explaining 51.15% of the total variance, whilst Zn and Pb explain only 18.93 of the total variance. This study illustrates the usefulness of multivariate statistical techniques for evaluation and interpretation of large complex data sets to get better information about the heavy metal concentrations and design of monitoring network.

A. F. M. Alkarkhi (✉) · N. Ismail · A. Ahmed
School of Industrial Technology,
Environmental Technology Division,
Universiti Sains Malaysia,
11800 Penang, Malaysia
e-mail: abbas@usm.my

N. Ismail
e-mail: norlii@usm.my

A. Ahmed
e-mail: anees@usm.my

A. m. Easa
School of Industrial Technology,
Food Technology Division, Universiti Sains Malaysia,
11800 Penang, Malaysia
e-mail: azhar@usm.my

Keywords Heavy metal · Sediments · Estuary · Factor analysis · Cluster analysis

Introduction

Estuaries are complex and dynamic environmental components which receive large amounts of contaminants from urban areas and industrial sites. Heavy metals are of particular concern due to their environmental persistence and biogeochemical recycling and ecological risks. Pollution of the natural environment by heavy metals is a worldwide problem as these metals are indestructible and have toxic effects on living organisms when they exceed a certain concentration limit (MacFarlane and Burchett 2000). Sediments are mixture of several components including different mineral species as well as organic debris. Sediments represent one of the ultimate sinks for heavy metals discharged into environment (Luoma and Bryan 1981; Bettinetti et al. 2003; Hollert et al. 2003).

The application of different multivariate statistical techniques, such as cluster analysis (CA) and factor analysis (FA) help in the interpretation of complex data matrices to better understand the heavy metals and ecological status of the systems studied, allows the identification of possible factors/ sources that influence heavy metals and offers a valuable tool for reliable management as well as rapid solution to pollution problems (Vega et al. 1998; Adams et al. 2001; Reghunath et al. 2002; Simeonov et al. 2004; Habes and Nugem 2006).

In present study different multivariate statistical techniques were applied to extract information about the similarities or dissimilarities between sampling sites, identification of heavy metals responsible for spatial variations in river estuaries, and the influence of possible sources (natural and anthropogenic) on the heavy metals of the two river estuaries in Malaysia.

Materials and methods

Description of study sites

Study site is located on the North West coast of Peninsular Malaysia, in the state of Penang and within a coastal mudflat in the Juru and Bukit Tambun district (Fig. 1). The sites are located adjacent to industrial

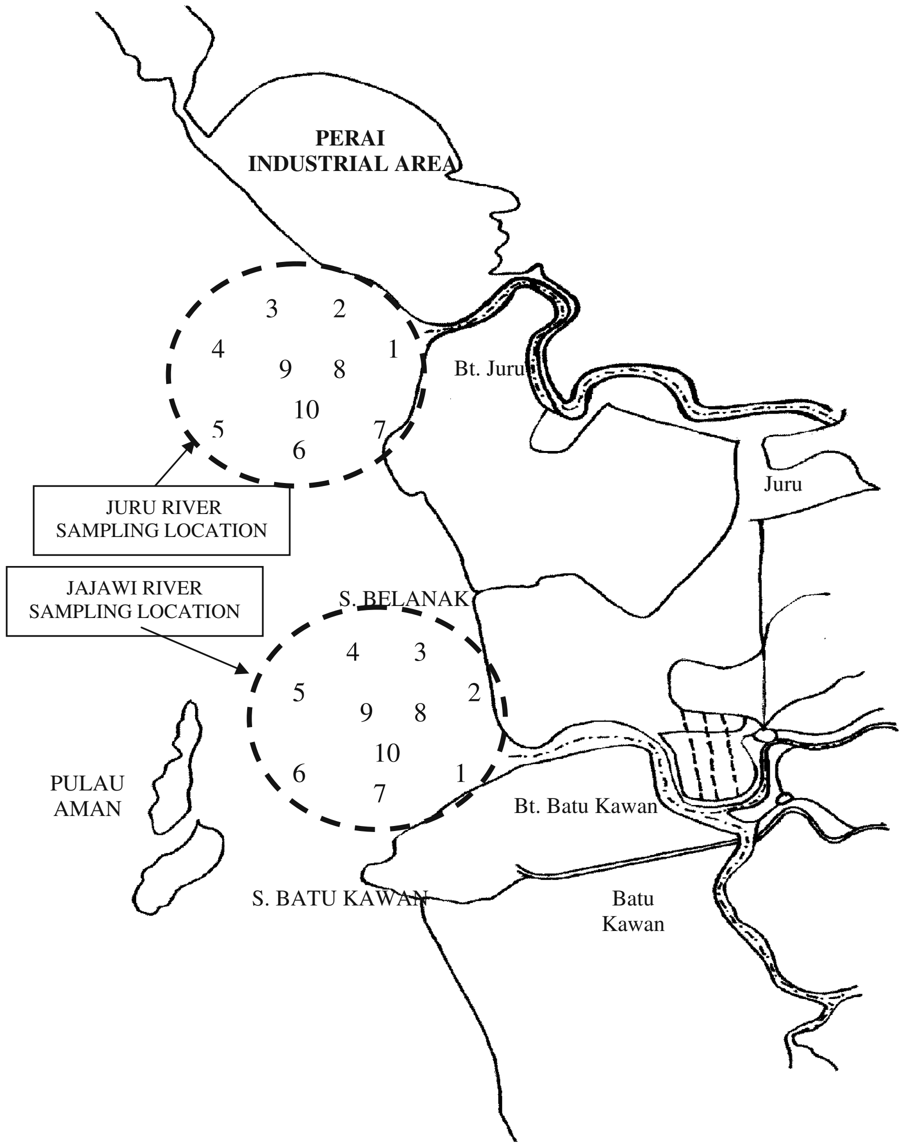

Fig. 1 Map of sampling sites for the study area

areas which were reclaimed from mangrove. The types of industry presently in operation include: electronics; textiles; basic and fabricated metal products; food processing and canning; processing of agricultural products; feed mills; chemical plants; rubber based industry; timber based wood products; paper products and printing works; and transport equipment. Other main activities that are operating in the vicinity of the cultured area are a ships' harbor with petroleum unloading and a red earth quarry which extends right up to the coastline. There are three main rivers flowing into the area, Juru River, Semilang River and Jejawi River where some fishing villages are situated.

Sampling and analytical procedure

Sediment sampling was carried out during five fieldtrips, at the Juru and Jejawi Estuaries in the year 2006. Sediment samples were collected at low tide with an Eijkelkamp gouge auger from each of the 20 sites (10 sites from each river estuary). After being collected, sediments were put into a sealed polyethylene bag in order to limit gaseous exchanges and were brought back to the laboratory and kept refrigerated at 4°C before analyses. Sediment samples were dried and grind with an agate pestle and mortar. Then the sediment samples were sieved through a 2 mm sieve. One gram of sediment samples (dry) was taken and digested for 4 h at 140°C with 5 ml concentrated nitric acid according to standard digestion technique (APHA 2000). The whole extract was place in Teflon vessels, previously washed with concentrated nitric acid. Flame atomic absorption spectrometer (FAAS; Perkin Elmer HGA-600) was employed for the analysis of Cu, Zn, Cd, Fe, Pb, Cr and Mn and cold vapor atomic absorption spectrometer (CV-AAS) method was employed for Hg analysis after sample digestion in acid solution. Blank acid mixtures were digested in the same way.

Multivariate statistical methods

Cluster analysis

Cluster analysis (CA) is a multivariate technique, whose primary purpose is to classify the objects of the system into categories or clusters based on their similarities, and the objective is to find an optimal grouping for which the observations or objects within each cluster are similar, but the clusters are dissimilar to each other. Hierarchical clustering is the most common approach in which clusters are formed sequentially. The most similar objects are first grouped, and these initial groups are merged according to their similarities. Eventually as the similarity decreases all subgroups are fused into a single cluster. CA was applied to heavy metals in sediment data using a single linkage method. In the single linkage method, the distances or similarities between two clusters A and B is defined as the minimum distance between a point in A and a point in B:

$$D(A,B) = \min\left\{d\left(y_i, y_j\right), \text{for } y_i \text{ in } A \text{ and } y_j \text{ in } B\right\} \quad (1)$$

Where $d(y_i, y_j)$ is the Euclidean distance in Eq. 1.

At each step the distance is found for every pair of clusters and the two clusters with smallest distance (largest similarity) are merged. After two clusters are merged the procedure is repeated for the next step: the distances between all pairs of clusters are calculated again, and the pair with minimum distance is merged into a single cluster. The result of a hierarchical clustering procedure can be displayed graphically using a tree diagram, also known as a dendrogram, which shows all the steps in the hierarchical procedure (Richard and Dean 2002; Alvin 2002).

Factor analysis

Factor analysis (FA) is designed to transform the original variables into new uncorrelated variables called factors, which are linear combinations of the original variables. The FA is a data reduction technique and suggests how many variates are important to explain the observed variances in the data. Principal components method (PCA) is used for extraction of different factors. The axis defined by PCA is rotated to reduce the contribution of less significant variables (Richard and Dean 2002; Alvin 2002). This treatment provides a small number of factors that usually account for approximately the same amount of information as the original set of observations. The FA can be expressed as:

$$F_i = a_1x_{1j} + a_2x_{2j} + \ldots + a_mx_{mj} \quad (2)$$

where F_i is the factor, a is the loading, x is the measured value of variable, i is the factor number,

j is the sample number and m the total number of variables.

And factor scores can be expressed as:

$$Z_{ij} = a_1 f_{1j} + a_2 f_{2j} + \ldots + a_m f_{mj} + e_{ij} \tag{3}$$

where z is the measured variable, a is the factor loading, f is the factor score, the residual term accounting for errors or other source of variation.

Result and discussion

Spatial similarity and site grouping

Cluster analysis was used to identify the similarity groups between the sampling sites. CA rendered a dendrogram as shown in Fig. 2, grouping all 20 sampling sites into two statistically significant clusters. Cluster 1 (Sites 1–10 for Juru Estuary) and cluster 2 (sites 11–20 for Jejawi Estuary) reveals that the two regions have different characteristics in terms of heavy metals selected and indicates that each region receives pollution from different sources (natural and anthropogenic). Figure 2 also shows that some sites in each cluster have similar characteristics and can be classified into groups, cluster 1 for Juru Estuary includes four different groups, group 1 (sites 6–10), group 2 (sites 1–2), group 3 (sites 3 and 5), and group 4 includes only site 4, while cluster 2 for Jejawi Estuary classified into four different groups, group 1 (sites11–13), group 2 (sites 14–15 and 17–18), group 3 (sites 19–20), and group 4 includes only site 16. This grouping gives evidence that sites in each group share each other the sources of pollution from point or non-point sources. It implies that for rapid assessment of heavy metals only one site in each group may serve as good in spatial assessment of the heavy metals as the whole network. It is evident that the CA technique is useful in affording reliable classification of heavy metals in the whole region and will make possible to design a future spatial sampling strategy in an optimal manner. Thus, the number of sampling sites in the monitoring network was reduced; hence, cost without losing any significance of the outcome.

Source identification and relationship between sampling sites

Factor analysis was carried out on the data set (eight variables) to compare the compositional patterns between analyzed heavy metal samples and to identify the sources of variation. FA yielded two factors with Eigen-value >1, explaining 70.08 of the total variance in heavy metals data set. An Eigen-value gives a measure of the significance of the factors; the factors with highest Eigen-value are most significant and responsible in explaining large variation in the data. The Eigen-values for different factors, percentage variance accounted, and cumulative percentage variance are given in Table 1. The factor analysis was actually performed on the correlation

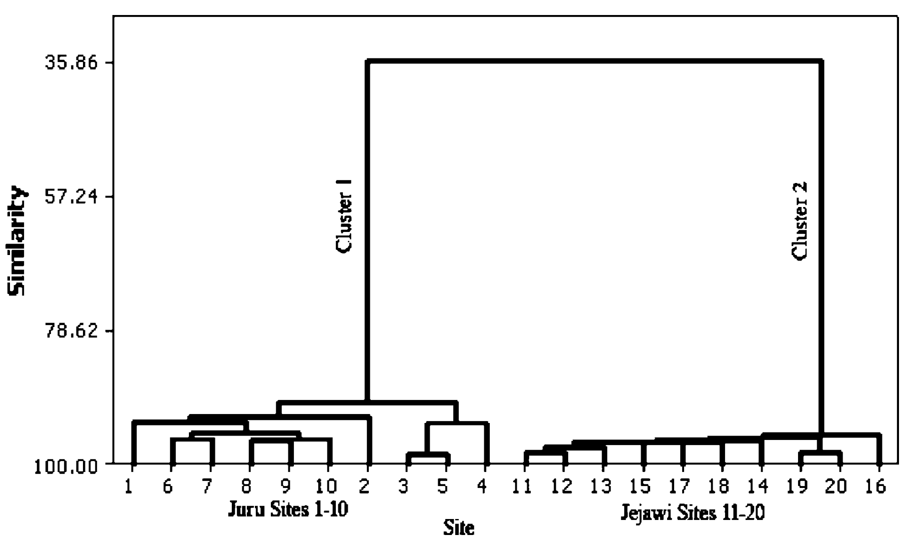

Fig. 2 Dendrogram showing clustering of sampling sites based on heavy metals in sediment characteristics of Juru and Jejawi Estuaries

Table 1 Extracted values of various factor analysis parameters for the Juru and Jejawi Estuaries

Component	Total variance explained before rotation			Total variation explained after rotation		
	Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	4.51	56.40	56.40	4.09	51.15	51.15
2	1.10	13.69	70.08	1.51	18.93	70.08

Extraction method: principal component analysis

matrix between different parameters followed by Varimax rotation.

The parameter loading for the two factors from the factor analysis of the data are given in Eq. 4 and 5, which illustrate that most of the variables associated with each factor are well defined and contribute very little to other factors, which helps in the interpretation of the result.

$$F_1 = 0.96\ \text{Fe} + 0.93\ \text{Mn} - 0.92\ \text{Cd} + 0.86\ \text{Hg} - 0.72\ \text{Cu} + 0.31\text{Cr} + 0.04\ \text{Zn} - 0.31\ \text{Pb} \quad (4)$$

$$F_2 = -0.22\ \text{Fe} - 0.32\ \text{Mn} + 0.24\ \text{Cd} - 0.10\ \text{Hg} + 0.05\ \text{Cu} - 0.30\ \text{Cr} + 0.87\ \text{Zn} + 0.68\ \text{Pb} \quad (5)$$

The first factor accounted for 51.15% of the total variance was positively correlated with Fe, Mn, and Hg, while negatively correlated with Cd and Cu. This factor may be termed as pseudo anthropogenic factor because both industrial as well as domestic factors contribute to their presence.

Factor 2 on the other hand, accounted for 18.93% of the total variance and is positively correlated with Zn and Pb. Since lead and zinc are considered as a toxic metal, this factor can be termed as anthropogenic factor because of industrial origin only. Further, the factor may be accompanied with metal removal by the natural recovery system of the river where, particles in suspension act as heavy-metal scavengers in the course of sedimentation, involving a series of chemical and physical processes including adsorption, the formation of complexes, ion exchange and inclusion within particles (Kuppusamy and Giridhar 2006).

The relationship between factor scores and the samples from different sites were studied to understand the behavior of heavy metals at different locations. Factor scores for the first factor from different sites are given in Fig. 3. It can be seen that all sites in cluster 1 (Juru Estuary) exhibited positive

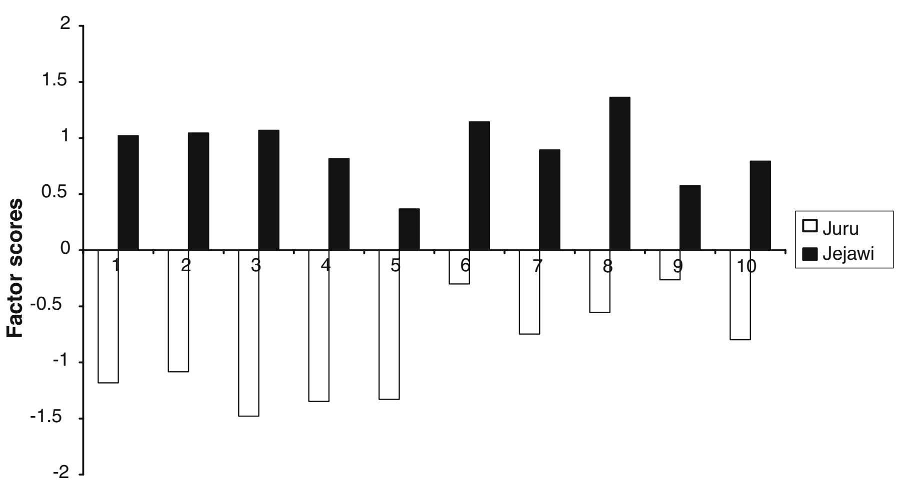

Fig. 3 Factor scores for factor 1 for different sampling sites along Juru and Jejawi Estuaries

contribution to pseudo anthropogenic factor, while all sites in cluster 2 (Jejawi Estuary) exhibited negative contribution to this factor. This may be due to changes in discharges along the study area. Positive contribution attributed to the concentrations of Fe, Mn and Hg, whereas negative contribution due to the concentrations of Cd and Cu. Thus, the two estuaries are different in terms of pollution load of heavy metals. This also indicates the pattern of distribution of polluting industries, agricultural and orchid plantation activities surrounding these two rivers.

The scores for the second factor are shown in Fig. 4, it can be seen that positive contribution mainly due to the concentrations of Zn and Pb. Heavy metal concentrations exhibited different behavior from site to site regardless to the region. It may be accounted for the different types of dynamics of the river water at different equilibrium conditions which results in irregular mixing and equilibrium for the heavy metal distribution in aqueous and solid phase matrices of the river. The high total concentrations of heavy metals in sediments may not necessarily indicate anthropogenic contamination, because of different background levels in parent materials and sediment properties.

The study area seemed to be highly polluted with Fe, Zn and Cd and this in agreement with previous studies at Juru River, Malaysia on metal accumulation on aquaculture (Yahya and Zubir 1994). According to report by DANCED (DANCED 1998) the electroplating, pulp and paper, textiles, food and beverages and auto-workshops industries were closely linked to industrial pollution in the Prai industrial area. This is in line with the report conducted by department of environment (DOE 1999) that the four predominant industries of Penang are electronics/electrical, textiles, fabricated metal products, plastic and plastic products. Other industries include rubber based, paper and paper products/printing works, chemical/fertilizers and basic metal industries.

Conclusion

Multivariate statistical techniques including Cluster analysis and factor analysis have provided satisfactory picture of the overall fate of heavy metals in sediments and give a better understanding about the sources of pollution. CA showed reliable classification of heavy metals in the whole region and made possible to design future spatial sampling strategy in an optimal way. Factor analysis provides valuable information about the source of pollution, since more than 70% of the total variance explained by two factors, pseudo anthropogenic and anthropogenic factor. The first factor explains most of the observed variation in the data (51.15%). The parameters correlated with this factor are Fe, Mn, Cd, Hg, and Cu. The second factor responsible for the presence of Zn and Pb is taking a secondary role in the evaluation of pollution status.

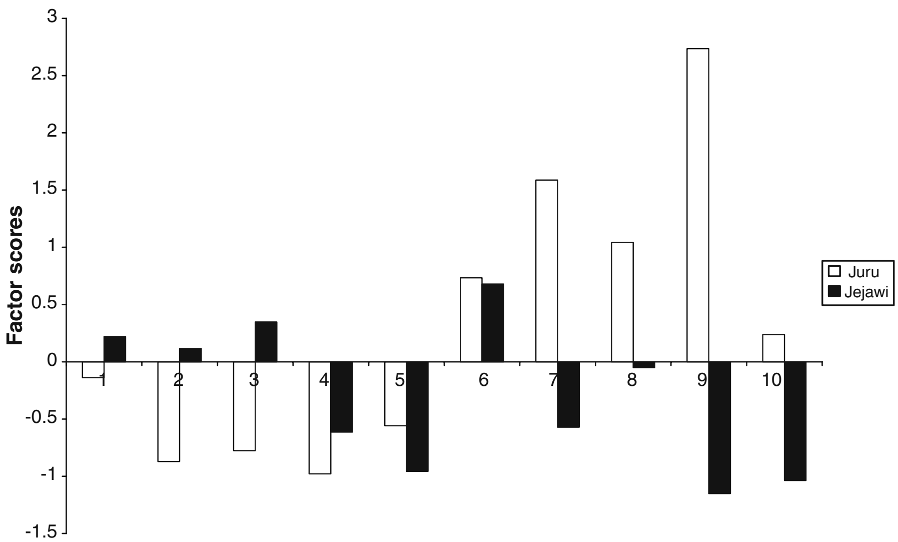

Fig. 4 Factor scores for factor 2 for different sampling sites along Juru and Jejawi Estuaries

Acknowledgement The authors gratefully acknowledge the financial assistance from Universiti Sains Malaysia and the research facilities by Dean School of Industrial Technology, USM. Grant number PTEKIND/ 638021.

References

Adams, S., Titus, R., Pietesen, K., Tredoux, G., & Harris, C. (2001). Hydrochemical characteristic of aquifers near Sutherland in the Western Karoo, South Africa. *Journal of Hydrology*, *241*, 91–103.

Alvin, C. R. (2002). *Methods of multivariate analysis*. USA: Wiley.

APHA (2000) Standard methods for the examination of water and wastewater, 18th ed. American Publich Health Association (APHA), American Water Works Association (AWWA) and Water Pollution Control Federation (WPCF), Washington, DC.

Bettinetti, R., Giarei, C., & Provini, A. (2003). A chemical analysis and sediment toxicity bioassays to assess the contamination of the River Lambro (Northen Italy). *Archives of Environmental contamination Toxicology*, *45*, 72–80.

DANCED (1998). Penang coastal profile, integrated coastal zone management project report. Penang State Government, Malaysia.

Department of Environment (DOE) (1999). Ministry of science and technology, Malaysia, Annual Report.

Habes, G., & nugem, Y. (2006). Assessing Mn, Fe, Cu, Zn, and Cd pollution in bottom sediments of Wadi Al-Arab Dam, Jordan. *Chemospher*, *65*, 2114–2121.

Hollert, H., Keiter, S., Konig, N., Rudolf, , Ulrich, M., & Braumbeck, T. (2003). A new sediment contact assay to assess particular-bound pollutants using zebrafish (*Danio rerio*) embryos. *Journal of Soils and Sediments*, *3*, 197–207.

Kuppusamy, M. R., & Giridhar, V. V. (2006). Factor analysis of water quality characteristics including trace metal speciation in the coastal environmental system of Chennai Ennore. *Environment International*, *32*, 174–179.

Luoma, S. N., & bryan, G. W. (1981). A statistical assessment of the form of trace metals in oxidized estuarine sediments employing chemical extractants. *Science of the Total Environment*, *17*, 165–196.

Macfarlane, G. R., & Burchett, M. D. (2000). Cellular distribution of Cu, Pb and Zn in the grey mangrove *Avicennia marina* (Frorsk.) Vierh. *Aquatic Botany*, *68*, 45–59.

Reghunath, R., Murthy, T. R. S., & Raghavan, B. R. (2002). The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. *Water Research*, *36*, 2437–2442.

Richard, A. J., & Dean, W. W. (2002). *Applied multivariate statistical analysis*. London: Prentice–Hall.

Simeonov, V., Simeonova, P., & Tsitouridou, R. (2004). Chemometric quality assessment of surface waters: Two case studies. *Chemical & Engineering Ecology*, *11*(6), 449–469.

Vega, M., Pardo, R., Barrado, E., & Deban, L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research*, *32*, 3581–359.

Yahya, M. N., Zubir, D. (1994). The effect of metals accumulation on aquaculture from Juru River, Research Report.