International Journal of Ecology and Environmental Sciences 39 (3): 187-193, 2013
ISSN: 0377-015X (Print); 2320-5199 (Online)

Multivariate Statistical Techniques to Identify the Source of Pollution and Assessment of Surface Water Quality

M.A. HOSSAIN, I.M. SUJAUL* AND M.A. NASLY
Faculty of Civil Engineering and Earth Resources, University Malaysia Pahang, 26300Gambang, Pahang, Malaysia
Corresponding author; *Email: sujaulbd@gmail.com*

ABSTRACT

Principal component analysis (PCA) and multiple linear regression (MLR) analysis were applied on the data for 14 physico-chemical parameters of surface waters from Tunggak River adjacent to the Gebeng Industrial Estate, Pahang, Malaysia during February 2012-January 2013 with the objective of identifying sources of pollution and their contribution to the variation in water quality. Physico-chemical parameters were determined for a period of 12 months by following standard methods of analysis. Results revealed that most of the parameters including BOD, COD, conductivity, NH_4-N and phosphorus were in concentrations greater than the national standard of Malaysia. PCA was applied to identify the source and MLR analysis was done to determine their contribution. PCA yielded five VFs; which extracted 74.72% of total variance that established its validation. Results showed that, surface water quality was strongly influenced by ionic groups of salts, soil erosion and agricultural runoff, organic and nutrient pollutions from domestic wastewater, industrial sewage and wastewater treatment plants. Vicinity of industrial park resulted in low DO concentration all over the basin. MLR showed the contribution of every variable to be highly significant ($p < 0.01$).

Key Words: Principal Component Analysis, Varimax Rotation, Multiple Linear Regression, Dissolved Oxygen, Anthropogenic Activities, Industrial Effluents

INTRODUCTION

Surface water pollution is a major global concern nowadays. Population growth, urbanization, and increasing commercial and industrial activities are exerting tremendous pressure on the surface water resources and generate a large amount of pollutants that degrade the water quality. Rivers that are the main reservoir of surface water, serve as the recipients of excessive amount of wastes generated and discharged from anthropogenic activities (Milovanovic 2007, Satheeshkumar and Anisa 2011). Water pollution is ever increasing also threatens human health, aquatic ecosystems, and economic and social welfare (Milovanovic 2007).

For its essentiality surface water management and control of pollution has a great obligation for the environmentalist as well as the policy makers. It is obvious that, for better management and control of surface water pollution, authentic and reliable information of its sources is required (Sing et al. 2005). Apportioning the sources of pollution can provide ample opportunities to the policy makers and environmental managers to set priorities for sustainable water resources management (Huang et al. 2010). Again, identification of the sources of pollutants and their contribution in water quality is one of the major challenges for the researchers (Adamu and Abdu 2012). Use of principal component analysis (PCA) to identify pollution sources is now a common unbiased statistical method (Satheeshkumar and Anisa 2011) and eventually multiple linear regressions is useful in estimating the contribution of different parameters (Praveena et al. 2011). PCA helps in interpretation of complex data in the simplified way (Pejman et al. 2009) for better understanding by investigating the structural information of confusing data (Ragno et al. 2007). On the other hand, multiple linear regressions is the tool to examine the relationship between single dependent variable and a set of independent variables to best represent relationship in the each factor(Adamu and Abdu 2012). Shrestha and Kazama, (2007); Huang et al.

(2010); and Juahir et al. (2011) employed PCA to identify water pollution sources. In their study, Onojake et al. (2011) used PCA to identify the latent factor and found that Rivers in Delta State of Nigeria were heavily polluted due to anthropogenic source of pollution. Similarly, recent study conducted by Koklu et al. (2010) revealed that multiple regressions analysis also identifies important parameters to determine the major sources of pollution.

Malaysia is a member of the newly industrialized countries (NIC) and the study area Gebeng is one of the important industrial areas in the country. Again, surface water is the main source of water here. Nevertheless, no in-depth study regarding the surface water quality and source apportionment of pollution was yet done in this industrial area. In this study, multivariate statistics PCA and MLR were used to evaluate the pollution sources and to assess the contribution of significant parameters to water quality variation.

Study Area

Gebeng on the east coast of peninsular Malaysia lies between 3°55' 0" to 4°01' 0" N and 103° 22' 0" to 103° 27' 0" E (Figure 1). It comprises of one industrial estate - the Gebeng Industrial Estate (GIE), a residential area and a seaport. Since 1970s, multifarious industries have been established and the trend is ever increasing. These industries discharge their effluents into river Tunggak (Nasly et. al. 2012) which originates at the upper end of Gebeng and after passing through the GIE, it falls into South China Sea with another river Balok. Normal tide occurred every day twice and water goes at about 3 km. upstream from the sea. Ten Monitoring stations were selected ro cover the entire river basin and surroundings of GIE, based on the land use-pattern, point sources of pollution, vegetation and river network. Global Positioning System (GPS) was used to locate the stations.

METHODS

Water samples were collected from pre-selected 10 stations every month during January to December 2012. Grab samples were collected 10 cm below the water level with 3 replicates using 1000 mL HDPE bottles; for BOD samples the dark BOD bottles (300 mL) were used. Collected samples were immediately preserved in ice-boxes under low temperature conditions and quickly transported to the laboratory maintaining APHA (2005) standard methods. Bringing to the laboratory samples was stored at 4 °C until analysis.

Samples were analyzed for 14 physiochemical parameters. Of these, temperature, pH, dissolved oxygen (DO), total dissolved solids (TDS), conductivity, salinity and turbidity were estimated in the field during sampling using multi-parameters monitoring instrument (YSI Inc. Yellow Spring, Ohio, USA). Five-day biochemical oxygen demand (BOD_5), chemical oxygen demand (COD), total suspended solids (TSS), ammonia (NH_3-N) and nitrate-N (NO_3), inorganic phosphorus (PO_4-P) and sulphate (SO_4) were analyzed in the laboratory.

All analyses were made according to APHA (2005) and HACH (2005) standard procedure. For BOD_5 determination, samples were incubated for 5 days at 20±30 °C. COD was determined by reactor digestion method using HACH spectrometer 5000. Nitrogen (NH_3) was measured by Nessler's method; nitrate by cadmium reduction method, SO_4 by sulfavar 4 method and PO_4^{3-} was determined by ascorbic acid method using colorimetry (APHA 2005).

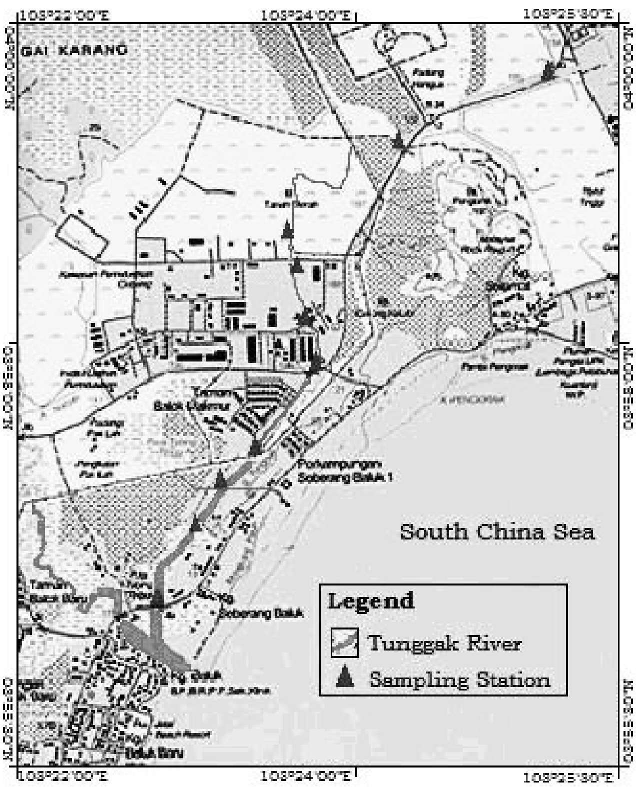

Figure 1. Study area showing the sampling locations

Principal Component Analysis (PCA)

Principal component analysis (PCA) was executed on the measured concentration of the physicochemical parameters of the study area (Onojake and Abrakasa 2012; Adamu and Abdu 2012). The method converts a set of possibly correlated variables into a set of linearly uncorrelated variables (principal components, called variance factors, VF). It can be used to reduce the number of variables and to explain the variance of large data set with a smaller set of variables (Wu and Wang 2007). VFs are the linear combinations of actual data and the eigenvectors. In this study PCA were applied on the estimated data standardized through z-scale transformation, so that misclassification due to wide differences in data dimensionality can be avoided (Wang et al. 2007, Liu et al. 2003, Simeonov et al. 2003).The z-scale equation is:

$$Zij = a_{i1}x_{1j}+a_{i2}x_{2j}+ ... + ... +a_{im}x_{mj} \quad (1)$$

where, Z is the component score, a is the component loading, x is the measured value of a variable, i is the component number, j is the sample number, and m is the total number of variables. For all statistical analysis SPSS 16.0 software was used.

Multiple Linear Regressions

Multiple linear regressions analysis is a statistical method to predict the relationship between a dependent variable and a set of explanatory (independent) variables (several predictors) (Koklu et al. 2010). It can be used for both explanatory and predictive purposes within experimental design and also in non-experimental design (Adamu and Abdu 2012). In multiple linear regressions the relationship between the dependent variable and the explanatory variables can be expressed by the following equation:

$$Y = \beta_0+\beta_1X_1+\beta_2X_2+ ... + ... + \beta_mX_m+\varepsilon \quad (2)$$

where, Y represents the dependent variable, X_1 …. X_m represent the several independent variables, β_0 is the constant term and β_1 …. β_m represent the regression coefficient and ε represents the random error.

RESULTS AND DISCUSSION

A descriptive statistics of the water quality parameters os given in Table 1 which shows the range, minimum and maximum values, mean, standard deviation and variance of the physico-chemical parameters and the Malaysian standards of those parameters, in the study area. The data reveal that the concentrations of conductivity, salinity, TDS, turbidity, NH_3-N, SO_4 and PO_4^{3-} were greater than the standard level of Malaysia (DOE 2008) with higher mean values of 2538.80 µS cm^{-1}, 1.27%, 4242.40 mg L^{-1}, 17.45 mg L^{-1}, 1.66 mg L^{-1}, 135.5 mg L^{-1} and 1.24 mg L^{-1}, respectively. This statistics shows that those variables may have a common source of origin (Adamu and Abdu 2012). However, the mean temperature (30.72 °C) was within the normal range (27-31^{0}C) (Saad et al. 2008) although at some stations it was observed to be higher (38.45 ^{0}C). The maximum and minimum pH were 9.12 and 4.15 respectively, with a mean of 6.70; which was almost in neutral level. The mean concentration of DO (3.27 mg L^{-1}) was very low compared to the standard level. It was likely because of the industrial and domestic wastewaters that caused high organic pollution and decay of large OM (Satheeshkumar and Anisa 2011, Wang et al. 2007). Moreover, at some stations, lower amount of DO was due to high temperature (Wu et al. 2009). Regarding the BOD_5 and COD values, the mean concentrations were 17.93 and 41.76 mg L^{-1} respectively; which were above the standard level. COD concentration was higher than BOD and both COD and BOD concentrations were higher than DO; this order of concentration reveals that anthropogenic pressure was associated with some natural pressure on the surface water quality (Adamu and Abdu 2012).

Principal Component Analysis

PCA was used to obtain composite variables (variance factors, VF). Before applying PCA, Kaisere-Meyere-Olkin (KMO) and Bartlett's test were performed to check the sampling adequacy. KMO test indicates the proportion of common variance and a value close to 1 denotes that PCA may be useful (Shrestha and Kazama 2007). For KMO value, there is a general thumb rule that it should be greater than 0.5 to precede a satisfactory PCA (Hinton et.al. 2004). In the present study, KMO was 0.663, indicating that the variables were correlated enough for appropriate PCA. Likewise, the Bartlett test of sphericity, significance level was 0.00 ($p < 0.01$) suggesting relationships between the variables. In this study, five variance factors (VF) were extracted by PCA based on Eigen value >1 and a varimax rotation was conducted to reduce the overlapping of genuine variables over every VF (Zhang et al. 2011). Factor loading matrix

Table 1 Descriptive statistics of the concentration of water quality parameters

Parameters	Range	Min.	Max.	Mean	Std. Dev.	Variance	Malaysian standard (max)
BOD (mg L^{-1})	38.25	0.10	38.35	17.93	12.02	144.51	1.00
COD (mg L^{-1})	139.00	1.00	140.00	41.76	31.58	997.27	10.00
Conductivity (µS cm^{-1})	2.71E4	16.00	27080.00	2.54E3	4756.67	2.26E7	1000.00
DO (mg L^{-1})	6.55	0.55	7.10	3.27	1.67	2.78	7.00 (min)
NH_3-N (mg L^{-1})	4.05	0.00	4.05	1.66	0.90	0.82	0.10
NO_3-N (mg L^{-1})	4.50	0.00	4.50	0.34	0.71	0.51	7.00
pH	4.97	4.15	9.12	6.70	1.21	1.47	6.5-8.5
Phosphate (mg L^{-1})	37.19	0.01	37.20	1.24	4.74	22.50	0.2
Salinity (%)	13.72	0.01	13.73	1.27	2.54	6.43	0.50
SO_4 (mg L^{-1})	1220.00	0.00	1220.00	1.36E2	216.32	4.68E4	250.0
SS (mg L^{-1})	74.00	1.00	75.00	15.46	12.915	166.80	25.0
TDS (mg L^{-1})	7.46E4	7.70	74600.00	5.24E3	12096.05	1.46E8	500.0
Temperature (^{0}C)	13.09	25.36	38.45	30.72	2.49	6.20	27.0-31.0
Turbidity (NTU)	198.41	1.59	200.00	17.45	18.79	353.19	5.0

Table 2 Rotated Component Matrix

Parameters	Components VF1	VF2	VF3	VF4	VF5
Conductivity	**0.952**	-0.096	0.032	0.130	0.061
Salinity	**0.949**	-0.105	0.029	0.081	0.077
TDS	0.866	0.000	-0.036	-0.051	0.080
SO_4	**0.814**	0.147	-0.024	0.150	-0.095
BOD	-0.121	**0.913**	-0.104	-0.069	-0.049
COD	0.224	**0.851**	-0.195	-0.010	0.070
pH	0.007	**0.665**	0.330	0.157	-0.108
Temperature	-0.143	**0.647**	0.296	0.043	-0.081
NO_3	-0.045	0.018	**0.813**	-0.047	-0.175
NH_3	0.110	0.130	**0.752**	0.028	0.414
Turbidity	-0.042	0.125	-0.036	**0.909**	0.034
TSS	0.440	-0.091	0.029	**0.791**	-0.010
Phosphate	-0.122	-0.118	0.180	0.007	**0.751**
DO	-0.186	0.008	0.170	0-.015	**-0.700**
Eigen value	3.756	2.589	1.573	1.358	1.185
Variance (%)	26.832	18.489	11.233	9.701	8.462
*CV (%)	26.832	45.321	56.554	66.255	74.717

Extraction method: principal component analysis
Rotation method: varimax with Kaiser Normalization
* CV means cumulative variance

is given in Table 2 and component plot of 14 variables iso shown in Figure 2. The summary of the PCA result (Table 2) after rotation demonstrated parameters loadings, Eigen values, % variance of each component and cumulative variance. It is seen that five significant factors (VF) extracted by PCA with Eigen value >1 altogether explained 74.72% of total variance.

The first factor (VF1) explained 26.832% of the total variance with strong loading of conductivity, salinity, TDS and SO_4. This VF indicated that this factor group was highly and positively contributed from the natural sources related variables like ionic groups of salts in the basin from inflows, soil erosion and runoff (Varol et al. 2012). This ionic factor group amounts a lots of

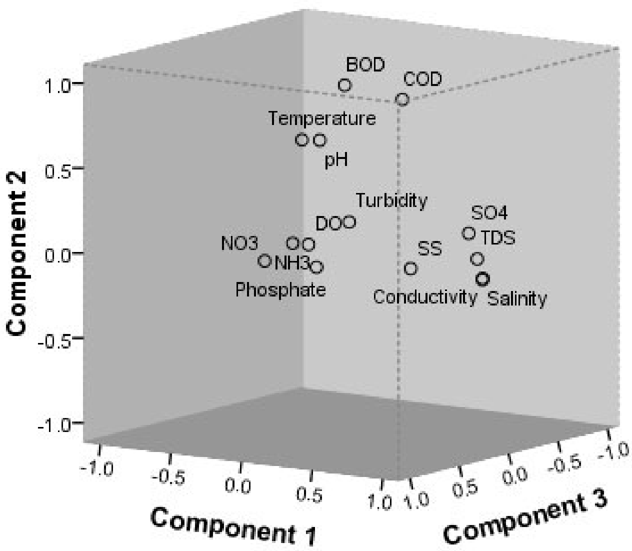

Figure 2. Component plot of 14 variables in rotated space

ions and their compounds that leads to high loading of those variables (Zhang et al. 2011). VF2 explained 18.49% of total variance and was strong positive loading with BOD_5 and COD, and, moderately loading with temperature and pH. High loading of BOD_5 and COD represented organic pollutant from industrial sewage and domestic wastewater (Zheng et al. 2008) and temperature and pH were due to the physiochemical source of variables and also strongly influenced by climatic and other environmental factors (Shrestha and Kazama 2007; Xiao-long et.al. 2007). VF3 had strong positive loading of NO_3-N and NH_3-N and explained 11.23% of total variance. It indicated the cause of these variables as non-point source originating from agricultural areas (Fuka-sawa 2005, Shrestha and Kazama 2007) and organic and nutrient pollution from domestic wastewater, industrial sewage and wastewater treatment plants (Sujaul et al. 2013, Zheng et al. 2008, Zhao et al. 2011). VF4 explained 9.7% of total variance and had strong positive loading on turbidity and TSS. This factor explained soil erosion from upland (Shrestha and Kazama 2007) and diluted to the river flow that increased the level of TSS and the significant positive correlation between TSS and turbidity might be the other cause. VF5 had strong positive loading of PO_4^{3-} and moderate negative loading of DO and explained 8.46% of total variance. High loading of inorganic phosphorus represented the influence from both point sources like industrial effluents, domestic wastewater (Shrestha and Kazama 2007) and non-point sources like agricultural runoff (Varol et al. 2012). Inverse relationship between PO_4^{3-} and DO indicated the difference in pollution sources, such as point sources and non-point sources. DO loading is extremely dependent upon point sources like organic wastes and some natural process like temperature variability. On the other hand, inorganic phosphorus largely comes from non-point sources as well as from point sources but not similar to those of DO.

Multiple Linear Regressions Analysis

Multiple linear regression analysis was done with SPSS 16.0 statistical software to identify the contribution of variables to water quality of Tunggak river basin. To detect the best predictors and remove the less significant variables (predictors) of water quality variation stepwise multiple linear regressions model was used (Hinton et.al. 2004). Classical assumptions of linear regressions were checked before the interpretation of MLR model results: normal p-p plot of regression-standardized residuals were analyzed (Figure 3a). It explained that all the observed values fall roughly along the straight line and indicated that the residuals are from normally distributed population. Furthermore, scatter plot of regression standardized predicted values against observed values (Figure 3b) also showed a linear relationship between the dependent variable and the predictors; and the residual variances are equal or constant.

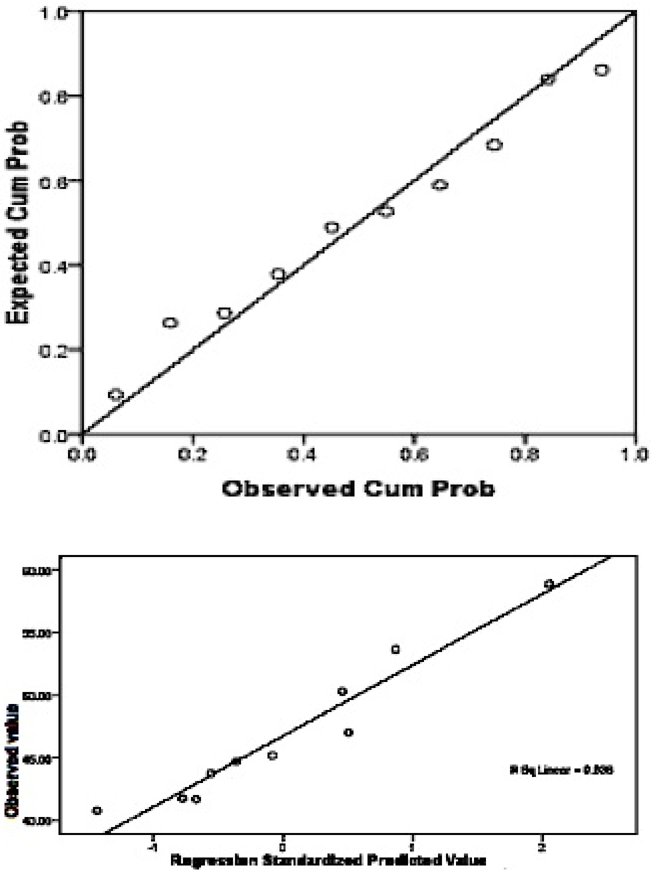

Figure 3. a) Normal p-p plot of regression standardized residuals and b) Scatter plot of regression standardized predicted values against observed values

Removing less significant variables using stepwise multiple linear regression model, four best predictors namely BOD_5, COD, pH and NH_3-N were detected (Table 3). Thus, maximum water quality variation (93.4%) in Tunggak River was explained by these four predictor variables. The coefficients of the predictors estimated in the model (Table 3) show that COD makes the strongest unique contribution to water quality variation with Beta coefficient value of -0.567. The second highest Beta value was for BOD_5 (-0.500) followed by NH_3-N (-0.454). The least contributor was pH with a Beta value 0.223. The negative sign of Beta value indicates that water quality was negatively associated / correlated with those predictors (Nathans et al. 2012). The analysis of variance (ANOVA) showed that the *F*-statistics value is 18.344 (df=4, $p = 0.003$), which is highly significant. This test confirms the linear relationship between the predictors of the applied models.

Table 3 Estimated coefficients of multiple linear model

.	Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Significance
(Constant)	60.043	5.216	11.511	0.000	
BOD	-0.354	0.186	-0.500	-1.902	0.116
COD	-0.190	0.083	-0.567	-2.307	0.069
pH	1.227	1.136	0.223	1.080	0.329
NH_3N	-4.337	1.412	-0.454	-3.072	0.028

CONCLUSION

The surface water quality of the study area was assessed and found to be deteriorated all over the area. Principal component analysis identified seven sources of pollution namely: ionic, soil erosion run-off, industrial sewage, domestic wastewater, wastewater treatment plant, dilution and agricultural run-off, and natural and climatic reasons. Multiple linear regression supported PCA result and identified four parameters that contributed higher to the water quality. The main source of those parameters was anthropogenic and from the analysis, it is clear that the major source of pollution was the industrial activities like poor waste management. As a newly industrialized country, Malaysia needs to take proper action plan for amicable association among the industrial activities and environmental protection; especially, control of water quality deterioration from industrial activities is a vital task. The findings of the study can be used, as baseline information for proper waste management (industrial and domestic) and emphasizing the monitoring of anthropogenic activities in Malaysia as well as all newly industrialized countries to confirm the least negative effects on surface water of those areas.

ACKNOWLEDGMENT

We are grateful to the University Malaysia, Pahang, and the Faculty of Civil Engineering and Earth Resources for their support and funding through the research project RDU 110354 and GRS 120363.

REFERENCES

Adamu, M. and Abdu, A. 2012. Application of principal component analysis & multiple regression models in surface water quality assessment. Journal of Environment and Earth Science. **2**2), 16-23.

APHA. 2005. Standard methods for the examination of water and wastewater. Washington, USA: American Public Health Association. 21st edition. 1368 pages.

Fukasawa, E. 2005. Determination of origin of nitrate nitrogen in Fuefuki river using stable isotope method. Bachelor thesis (unpublished). Department of Ecosocial System Engineering, University of Yamanashi, Japan.

HACH. 2005. Water analysis guide. HACH Company, USA.

Hinton, P.; Charlotte, B. and Isabella, M. M. 2004. SPSS Explained. Routledge, Taylor and Francis group. 400 Pages

Huang, F.; Wang X.; Liping, L.; Zhiqing, Z. and Jiaping, W. 2010. Spatial variation and source apportionment of water pollution in Qiantang River China) using statistical techniques. Water Research 44: 1562-1572.

Juahir, H.; Zain, M.S.; Yusoff, M.K.; Hanidza, T.I T.; Mohd Armi, A. S.; Toriman, M.E. and Mokhtar, M. 2011. Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. Environmental Monitoring and Assessment 173(1-4): 625-641.

Koklu, R.; Sengorur, B. and Topal, B. 2010. Water quality assessment using multivariate statistical methods, a case study: Melen River system. Water Resource management 24 (5): 959-978.

Liu, C.W.; Lin, K.H. and Kuo, Y.M. 2003. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. Science of the Total Environment 313(1-3): 77-89.

Milovanovic, M. 2007. Water quality assessment and determination of pollution sources along the Axios/Vardar River, Southeastern Europe. Desalinization 213 (1-3): 159-173.

Nasly, M.A.; Hossain, M.A. and Islam, Mir Sujaul. 2013. Water quality index of Sungai Tunggak: an analytical study. In the proceedings of 3rd International Conference on Chemical, Biological and Environment Sciences (ICCEBS'2013) January 8-9, 2013 Kuala Lumpur (Malaysia): 40-44.

Nathans, L.; Oswald, F. and Nimon, K. 2012. Interpreting Multiple Linear Regression: A Guidebook of Variable Importance. Practical Assessment, Research and Evaluation 17 (9). Available at: http://pareonline.net/getvn.asp?v=17&n=9

Onojake, M. C.; Ukerun, S. O. and Iwuoha, G. 2011. A statistical approach for evaluation of the effect of industrial and municipal wastes on Warri Rivers, Niger Delta, Nigeria. Water quality Exposure and Health 3(2): 91-99.

Pejman, A.H.; Nabi Bidhendi, G. R.; Karbassi, A.R.; Mehrdadi, N. and Esmaeili, Bidhendi M. 2009. Evaluation of spatial and seasonal variations in surface water quality using multivariate statistical techniques. International Journal of Environmental Science and Technology 6(3): 467-476.

Praveena, S.M.; Kwan, O.I. and Aris, A.Z. 2011. Effects of data pre-treatment procedures on principal component analysis: a case study for mangrove surface sediment datasets. Environmental Monitoring and Assessment. doi: 10.1007/s10661-011-2463-2.

Ragno, G.; De Luca, M. and Ioele, G. 2007. An application of cluster analysis and multivariate classification methods to spring water monitoring data. Micro-chemical Journal. 87: 119–127.

Saad, M.; Farah, N.; Nik, Abdul R.; Nik, Norulaini; Abdul Kadir; Mohd Omar and Mohd Omar, Fatehah. 2008. Project Report: Identification of Pollution Sources within the Sungai Pinang River Basin. Universiti Sains Malaysia. Available at: http://eprints.usm.my/4907/

Satheeshkumar, P. B. and Anisa, K. 2011. Identification of mangrove water quality by multivariate statistical analysis methods in Pondicherry coast, India. Environmental Monitoring and Assessment. doi: 10.1007/s10661-011-2222-4.

Shrestha, S. and Kazama, F. 2007. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environmental Modelling & Software 22: 464-475.

Simeonov, V.J.; Stratis, C.J.; Samara, G.J.; Zachariadis, D.; Voutsa, A.; Anthemidis, M.; Sofriniou, T. and Koumtzis, T. 2003. Assessment of the surface water quality in Northern Greece. Water Resources 37(17): 4119–4124.

Singh, K.P.; Malik, A.; Sinha, S.; Vinod, K. and Murthy, R. C. 2005. Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal components analysis. Water Air and Soil Pollution 166 (1-4): 321-341.

Sujaul, I.M.; Hossain, M. A.; Nasly M A, and Sobahan M A. 2013. Effect of industrial pollution on the spatial variation of surface water quality. American Journal of Environmental Science 9 (2): 120-129.

Varol Memet, Bülent Gökot, Aysel Bekleyen and Bülent Şen 2012. Spatial and temporal variations in surface water quality of the dam reservoirs in the Tigris River basin, Turkey. Catena. doi:10.1016/j.catena.2011.11.013.

Wang, X.; Yong, Y.L.; Jing-yi, H.; Gui-zhen, H. and Tie-yu, W.2007. Identification of anthropogenic influences on water quality of rivers in Taihu watershed. Journal of Environmental Management 19: 475-481.

Wu, M.L.; Wang, Y.S.; Sun, C.C.; Wang, H.L. and Dong, J.D. 2009. Using chemometrics to identify water quality in Day Bay, China. Oceanologia 52: 217–232.

Xiao-long, W.; Yong-long, L. U.; Jing-yi, HAN, Gui-zhen, HE and Tie-yu, W.2007. Identification of anthropogenic influences on water quality of rivers in Taihu watershed. Journal of Environmental Sciences 19: 475–481.

Zhang, Y.; Yin, Y.; Feng, L.; Zhu, G.; Shi, Z.; Liu, X. and Zhang, Y. 2011. Characterizing chromophoric dissolved organic matter in Lake Tianmuha and its catchment basin using excitation-emission matrix fluorescence and parallel factor analysis. Water Research 45: 5110–5122.

Zhao, Jian, Guo, Fu, Kun, Lei and Yanwu, Li 2011. Multivariate analysis of surface water quality in the three Gorges area of China and implication for water management. Journal of Environmental Sciences. 23(9): 1460–1471.

Zheng, B.H., Cao, C.J., Qin, Y.W. and Huang, M.S. 2008. Analysis of nitrogen distribution characters and their sources of the major input rivers of three Gorges Reservoir. Environmental Science. 29: 1–6.

Received 15 May 2013;
Accepted after revision 8 September 2013