

Using multivariate statistical techniques to assess water quality of Nhu Y river in Thua Thien Hue province

● **Nguyen Minh Ky**

College of Sciences, Hue University

● **Nguyen Hoang Lam**

Danang University of Technology

(Bài nhận ngày 07 tháng 08 năm 2014, nhận đăng ngày 08 tháng 10 năm 2014)

ABSTRACT

The aims of this research are to assess water quality by organic and nutrient matters and identifying the environmental pressures, examine the impact of the loads to Nhu Y River, Thua Thien-Hue Province. Five stations were sampled at Nhu Y River, the research had monitoring of water quality parameters such as Temperature (Temp), Dissolved Oxygen (DO), Biological Oxygen Demand (BOD₅), Chemical Oxygen Demand (COD), Nitrate (NO₃⁻) and Phosphate (PO₄³⁻). The research used multivariate statistical techniques such as correlation analysis, principal component analysis (PCA) and cluster analysis (CA) to assess water quality. The correlation analysis shown a strong positive correlation exists between water quality parameters such as Temp↔DO and

BOD₅↔COD (p<0.01). The PCA technique was applied to water quality data sets, which was obtained from Nhu Y River and the results show that the indices which has changed water quality. The results of the PCA using a varimax rotation technique were illustrated with two principal components (PC) and accounts for 62.207% of the overall total variance. The first PC accounted for 40.873% of the total variance, which was loaded with Temp, DO, BOD₅ and COD. The second PC consists of NO₃⁻ and PO₄³⁻ which accounts for 21.334% of the total variance, it can be due to the discharge of agricultural activities. Similarly, the CA has identified two major clusters involving: BOD₅, COD, Temp, DO (the first cluster) and NO₃⁻, PO₄³⁻ (the second cluster).

Keywords: Monitoring, Nhu Y River, environment, multivariate statistical techniques, water quality.

1. INTRODUCTION

Nhu Y River is located in the northeast of Hue City and runs in Thua Thien Hue Province. Generally, it actually plays an important role in people's life and productive activities of a large

area including Hue City, Phu Vang district, and Huong Thuy town's wards in Thua Thien Hue Province. However, Nhu Y River was limited with Huong River by Dap Da dam and it rarely

has to added water to make flow's dilution. Nhu Y River is only received a small flow from Loi Nong River through Phat Lat River's branch. Moreover, Nhu Y River is also received a wastewater volume from the surround living public. Typically, the wastewater from the process of living, farming, agricultural and traditional crafts activities. In the past, Nhu Y River was polluted because of organic and nutrient matters. The document analysis process and field surveys has showed that the reason make those status because the river is limited by flow, as a result, the river is more and more polluted, especially during dry season. On the other hand, water resource at Nhu Y River is used for many purposes such as daily activities as well as agricultural irrigation, etc. Therefore, the necessary requirement is to carry out monitoring activities and assess surface water quality to Nhu Y River which aim at preserving its water quality from pollution and protecting public health. Generally, water quality parameters of

environmental concerns including DO, BOD₅, COD, NO₃⁻ and PO₄³⁻. Obviously, the nutrients and organic matters contamination in the river is one of the major quality issues in many fast growing cities. Organic matters contamination has negative effects due to their potential toxicity for the environment, aquatic animal and plants. The assessment of water quality in developing countries has become a critical issue in recent years, especially due to the concern that fresh water will be scarce resource in the future. The increasing in population can result in some environmental issues and water pollution is an alarmed problem in many developing countries. The rapid urbanization and industrial development during last decade can lead to some serious concerns for the environment in Hue City. In this study, therefore, multivariate statistical techniques are the methods of rating that shows the composite influence of individual parameters on the overall water quality.

2. MATERIAL AND METHODS

2.1. Research area

Thua Thien Hue Province is located in the central Vietnam. This province has an area of 5.053 square kilometers in which 49.107 hectares of that are used as agricultural land [8]. Thua Thien Hue has the features of a tropical monsoon climate with the dry season which is from March to August and has high temperatures between 35 and 40 °C. The rainy season is from August to January, especially a flood season start from October. The average rainy season temperature is 20 °C; sometimes it decreases to 9 °C. Spring lasts from January to late February. The climate is similar to central Vietnam in general: a tropical monsoon climate. The cool season is from November to March with cold northeasterly

winds. The relative humidity is high which is between 85 and 95 percent. It is very humid in July but the relative humidity is lower, sometimes it may down to 50 percent. The annual precipitation in the province is 3200 mm but there are important variations. Depending on the year, the annual average rainfall can be 2500 to 3500 mm in the plains and from 3000 to 4500 mm in the mountains. In some years, the rainfall may be much higher and reach more than 5000 mm in the mountains. Rainfall often occurs in short heavy bursts causing flooding and erosion which can result in a number of serious social, economic, and environmental consequences [9].

2.2. Monitored parameters and analytical methods

The water quality's monitoring process has been taken place at Nhu Y River during the certain period of time which is from March to August in 2012. A total of five sampling sites have selected and all the samples have been labeled properly for the indication of the source. The samples have been taken from 10 to 20 cm below the water surface using acid-washed and wide-mouth polyethylene plastic bottles. Standard procedures have been followed for the collection of water samples. Water samples have been stored at 4°C and have been transported to the laboratory immediately. Water samples have been collected, preserved and analyzed in

accordance with Standards Methods (APHA, 1998). Six water quality parameters are used for the index involve Temp, DO, BOD₅, COD, NO₃⁻, and PO₄³⁻. The Temp and DO has been determined and recorded immediately at the field sites. BOD₅ has been determined by DO, where DO is measured initially and after incubation, and the BOD₅ is computed from the differences between initial and final DO. The COD test measures the oxygen equivalent consumed by organic matter in a sample during strong chemical oxidation. The nitrate and phosphate (NO₃⁻, PO₄³⁻) have been determined by UV spectrophotometer method (APHA, 1998).

Table 1. Water quality monitoring stations in Nhu Y River

	Sampling stations	Latitude	Longitude
S ₁	Dap Da, Vy Da commune, Hue city	16° 28.402'N	107° 35.721'E
S ₂	Van Duong village, Xuan Phu commune, Hue city	16° 28.463'N	107° 36.412'E
S ₃	Chiet Bi village, Phu Thuong commune, Phu Vang district	16° 29.329'N	107° 36.659'E
S ₄	Cong Luong village, Thuy Van commune, Huong Thuy town	16° 29.463'N	107° 37.328'E
S ₅	Van The village, Thuy Thanh commune, Huong Thuy town	16° 28.895'N	107° 38.963'E

2.3. Data treatment and multivariate statistical methods

A general water quality assessment by QCVN 08:2008/BTNMT is used to indicate the overall water quality conditions. At the same time, the research also uses the multivariate statistical techniques such as correlation analysis, principal component analysis (PCA) and cluster analysis (CA). Pearson's correlation coefficients also calculate to assess relationship among Nhu Y River's water quality parameters. The CA technique has performed on the values of the water quality parameters. The PCA technique extracts the eigenvalues and eigenvectors from the covariance matrix of original variables, thus,

reducing the dimensionality of the data set. The eigenvalues of the principal components (PCs) are the measure of their associated variance, the participation of the original variables in the PCs is given by the loadings. According to former studies of Singh *et al.*, 2008 [6] and Amadi, 2012 [2], the PCs with eigenvalues >1 have been retained and are used to assess the compositional, temporal and spatial variations in the river quality due to anthropogenic activities. The PCA with varimax rotation has also been applied to the water quality data set to form a correlation matrix for different variables and assists in the identification of sources of various pollutants.

	S3	6	4.0	1.21	0.49	2.3	5.6
	S4	6	6.8	1.48	0.60	4.7	8.6
	S5	6	5.1	0.78	0.32	3.8	6.0
BOD ₅ (mg/L)	S1	6	10.9	2.52	1.03	6.3	13.4
	S2	6	7.7	2.34	0.96	5.4	11.7
	S3	6	10.9	2.57	1.05	7.8	14.1
	S4	6	7.0	0.88	0.36	5.7	7.9
	S5	6	12.1	1.18	0.48	10.7	14.2
COD (mg/L)	S1	6	16.2	0.71	0.29	15.4	17.2
	S2	6	14.2	2.11	0.86	11.7	17.2
	S3	6	14.4	1.99	0.81	11.9	17.0
	S4	6	10.4	2.48	1.01	7.7	14.2
	S5	6	17.5	3.36	1.37	14.0	22.1
NO ₃ ⁻ (mg/L)	S1	6	0.25	0.048	0.019	0.16	0.29
	S2	6	0.25	0.054	0.022	0.19	0.34
	S3	6	0.40	0.113	0.046	0.27	0.54
	S4	6	0.30	0.094	0.038	0.15	0.40
	S5	6	0.26	0.078	0.032	0.13	0.35
PO ₄ ³⁻ (mg/L)	S1	6	0.019	0.013	0.005	0.002	0.033
	S2	6	0.086	0.072	0.029	0.008	0.181
	S3	6	0.117	0.053	0.022	0.073	0.220
	S4	6	0.121	0.043	0.017	0.043	0.164
	S5	6	0.060	0.034	0.014	0.016	0.096

*Note: Mean- average value, SD- standard deviation, Std. Error- standard error

Table 3. Vietnamese surface water quality standard [7]

Parameters	Unit	QCVN 08: 2008/BTNMT			
		A1	A2	B1	B2
DO	mg/L	≥ 6	≥ 5	≥ 4	≥ 2
BOD ₅	mg/L	4	6	15	25
COD	mg/L	10	15	30	50
NO ₃ ⁻	mg/L	2	5	10	15
PO ₄ ³⁻	mg/L	0.1	0.2	0.3	0.5

Most of the parameters like BOD₅ and COD has shown average concentration varied and just satisfied of Colum B1 (For irrigation or other purposes requiring the similar quality of water or other purposes like B2) to B2 (For water

transport and other purposes requiring low quality) in Vietnam's Surface Water Quality Standard (QCVN 08:2008) [7]. In general, high BOD₅ levels in the water river can be indicates the decline by DO content because the oxygen is consumed by the bacteria leading to fish and

other aquatic organisms cannot survive [12]. In contrast, NO_3^- content is lower than standard (Colum A1: For residential use and other purposes like A2, B1 and B2). The current QCVN 08:2008 (Vietnam Ministry of Natural Resources and Environment, 2008) permissible limit of parameters of river water quality are 2.0 mg/L for NO_3^- (Colum A1). Usually, the NO_3^- concentration in surface water is normally low and varied less than 18 mg/L; however, it can also reach high levels because of the agricultural runoff, refuse dump runoff, contamination with human or animal wastes. The concentration often fluctuates follow the season and can increase in case if the river is fed by nitrate rich aquifers [14]. The surface water standard of the PO_4^{3-} content are 0.1 mg/L and 0.2 mg/L for Colum A1 and A2, respectively (where Colum A2: For residential use with proper treatment; preservation of aquatic plants or other purposes like in B1 and B2). Meanwhile, PO_4^{3-} content varies between 0.002 mg/L to 0.220 mg/L, and this has shown that Nhu Y River has slight polluted signal, especially at sampling sites S2 to S5. It indicates that Nhu Y River's water cannot use directly for the daily life demands. The DO values of the water are low and just satisfied Colum B1 and B2 in QCVN 08:2008, excepting sampling site at S4 (satisfied Colum A1 and A2). The DO plays an important role to maintain the river's life process. According to Hach *et al.*, the DO is an important parameter, and must have a minimum value of about 2 mg/L to maintain

higher life forms [4]. The DO content can be used as indicator signal of values such as BOD_5 , COD level in the river water flow. The values of COD indicate water pollution and it cannot be used for drinking supply purposes. The reasons could be linked to sewage effluents discharged from Hue City's wards, small scale industry and agricultural practice (surrounding Nhu Y River at Phu Vang district and Huong Thuy town). It can be explained that surface water quality in the area has been affected by activities such as market, a part of agriculture and domestic wastewater. In addition, according to Gleick, 1993 [3], the runoff also is the common form of non-point sources of surface water pollution. Consequence of the runoff from land surface carries the residues into river system which known as non-point sources pollution [10]. The research has based on the variability in range of all the parameters distributions as compared with their respective means which is an indication of a water quality parameters level of the river. The increasing trend of average parameters levels are as follows: $\text{PO}_4^{3-} < \text{NO}_3^- < \text{DO} < \text{BOD}_5 < \text{COD}$.

Before applying PCA and CA techniques, correlation analysis has been carried out. Regarding of relationship between water quality parameters, the study has used Pearson's correlation coefficients by correlation analysis. As a result, the positive correlation between BOD_5 and COD has demonstrated the relationship between these parameters.

Table 4. Pearson's correlation coefficients between water quality parameters

	Temp ($^{\circ}\text{C}$)	DO (mg/L)	BOD_5 (mg/L)	COD (mg/L)	NO_3^- (mg/L)	PO_4^{3-} (mg/L)
Temp ($^{\circ}\text{C}$)	1					
DO (mg/L)	0.646(**)	1				
BOD_5 (mg/L)	-0.239	-0.394(*)	1			

COD (mg/L)	-0.378(*)	-0.356	0.644(**)	1		
NO ₃ ⁻ (mg/L)	0.078	-0.035	0.198	-0.081	1	
PO ₄ ³⁻ (mg/L)	0.174	0.370(*)	-0.277	-0.201	0.247	1

** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

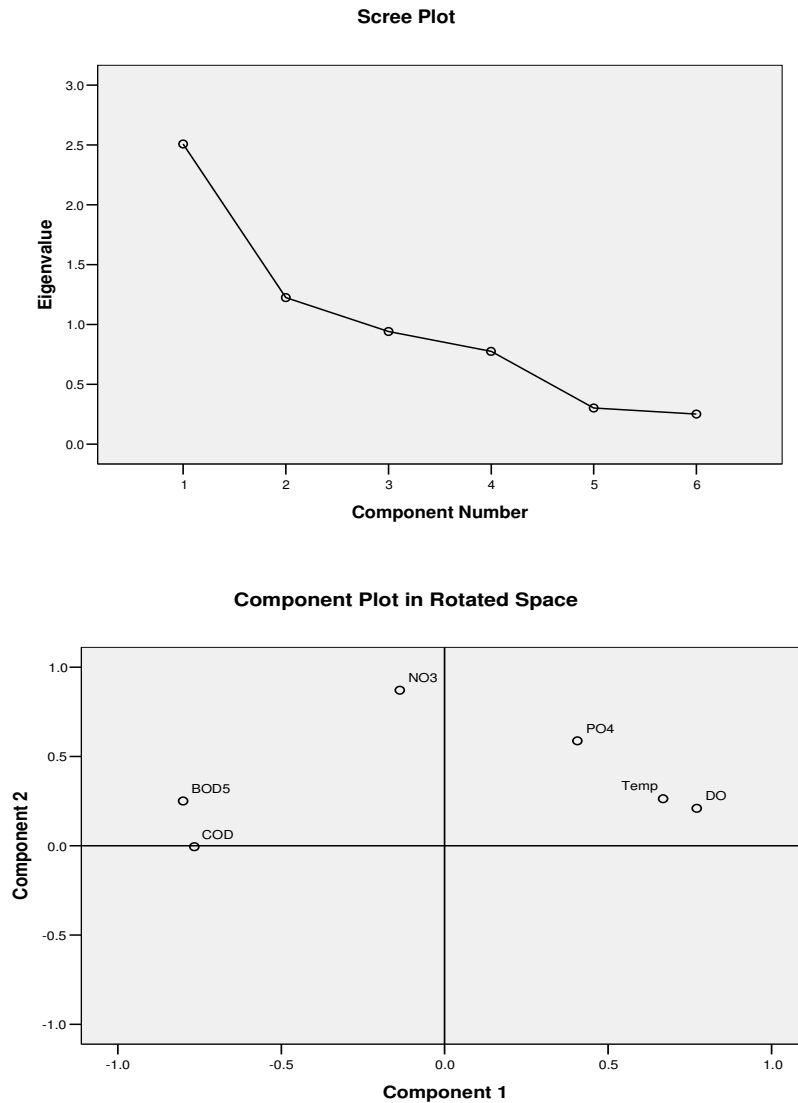


Fig. 2. Scree plot of the eigenvalues of PCs for water quality parameters and Bi-plot for the components in rotated space

The Pearson's correlation coefficients of the six parameters in Nhu Y River water have been summarized in the Table 4. Interrelationships have been established between physicochemical parameters indicators in which reliable correlations have been established using regression analysis. According to the research of Ajibade et al., (2008) [1] has also shown that the good correlations between these parameters as BOD₅ with COD ($r = 0.757$, $p < 0.05$). In this study, correlation analysis has showed a good correlation between Temp and DO ($r = 0.646$), BOD₅ and COD ($r = 0.644$) at $p < 0.01$ level. The relationship between BOD₅ and COD is known, to depend on the contaminants dissolved in the river water. The reason contributed water pollution can be due to organic compound. The results also show the negative relationship between DO and BOD₅ as well as between Temp and COD ($p < 0.05$).

Table 5. The varimax rotated factor loadings for water quality parameters of the Nhu Y river

Parameters	PC-1	PC-2
Temp (⁰ C)	0.668	0.263
DO (mg/L)	0.772	0.210
BOD ₅ (mg/L)	-0.800	0.250
COD (mg/L)	-0.766	-0.006
NO ₃ ⁻ (mg/L)	-0.137	0.872
PO ₄ ³⁻ (mg/L)	0.407	0.587
Eigenvalues	2.452	1.280
% of Variance	40.873	21.334
Cumulative %	40.873	62.207

Table 5 shows the results of the varimax rotated factor loadings for water quality parameters at monitoring sites in Nhu Y River. According to former studies of Singh *et al.*, 2008 [6] and Amadi, 2012 [2], the PCs with eigenvalues > 1 have been retained and are used to assess the compositional, temporal and spatial variations in the river quality due to

anthropogenic activities. In this study, the two PCs have been identified to be responsible for the deterioration of the river water and accounts for 62.207 percent of the overall total variance and with eigenvalues = 1.280 > 1 . Besides, this reduced the dimensionality of the total data from six to two (about 66.7% reduction) and resulted in a 37.793% loss of information contained in the dimensions. Though the significant PCs can still provide information on the most meaningful parameters and describes a whole data set affording data reduction with minimum loss of original information. PCA has determined a reduced number of two PCs that explain high percentage of the experimental data set variance (Explaining 62.207% of spatial and temporal variations with two PCs). The first PC accounts for 40.873 percent of the total variance and is characterized by high loading for Temp, DO (positive loadings) and BOD₅, COD (negative loadings). The first PC includes parameters Temp, DO, BOD₅ and COD; especially BOD₅ and COD due to the pollutant sources related to discharges process from sewage in the markets, household. It has indicated the component which is related to organic pollutants from domestic and traditional crafts wastewater. The PC involving parameters DO, BOD₅ and COD is due mainly to parameters of organic pollution and reflects contributions from waste water drainage [15]. It is also quite similar in research of Yeung, I.M.H., (1999), the PC explain BOD₅ and COD with strong factor loadings representing the anthropogenic input typically organic pollution and it can be due to the runoff or waste disposal activities [11]. The second PC consists of NO₃⁻ and PO₄³⁻ (positive loadings) which accounts for 21.334 percent of the total variance; it can be due to discharge of agricultural activities which is used chemical fertilizers. NO₃⁻ has showed a high

positive loading in the second rotated factor with 0.872, whereas PO_4^{3-} loading on this factor is 0.587. PC-1 involves in organic matters which are household sources while PC-2 is due to nutrients with artificial sources by runoff and agricultural activities. These results are also similar to the hierarchical cluster analysis in Figure 3. According to Munirah Abdul Zali *et al.*, 2011, the PC consists of BOD, COD, NH_3 , PO_4^{3-} , can be named as anthropogenic activities. It means that, Nhu Y River's pollution can be explained by anthropogenic effects. In addition, the research have shown that NO_3^- content describes industrial activities and densely populated housing areas surface runoff from agricultural related activities [5]. Moreover, according to the WHO, the nitrate concentration can easily reach several hundred milligrams per liter because of agricultural activities [13].

CONCLUSION

The study has shown general water quality situation at Nhu Y River by typical parameters such as DO, BOD_5 , COD, NO_3^- and PO_4^{3-} . Most of parameters content in stations has reached the value of B1 and B2 in national technical regulation on surface water quality (QCVN 08:2008/BTNMT). The content of PO_4^{3-} has been satisfied Vietnamese surface water quality national standard (Colum A2) with the average value of 0.082 mg/L. The results of correlation analysis with Pearson's correlation coefficients of the six parameters have been shown, however,

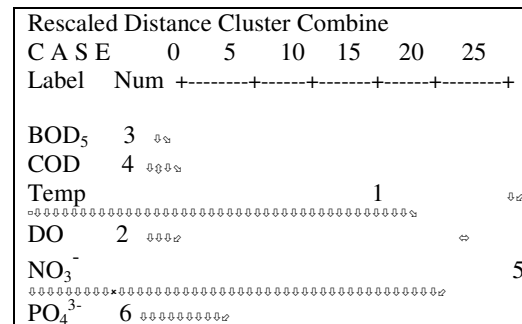


Fig. 3. A dendrogram using Average Linkage (between groups) from hierarchical cluster analysis

The CA technique has identified two major clusters involving: BOD_5 , COD, Temp, DO (the first cluster) and NO_3^- , PO_4^{3-} (the second cluster). The second cluster has shown pollutants sources which are related to agricultural activities.

just indicating strong relationship between Temp with DO and BOD_5 with COD. Specifically, Pearson's correlation coefficients of the BOD_5 and COD is 0.644 ($p < 0.01$). The PCA has been used in evaluating overall pollution of surface water pollution at Nhu Y River. The study shown that, two PCs (Eigenvalues > 1) have emerged which account for 62.207 percent of cumulative variance. The results of study suggest that the river has been mainly affected by nutrient and organic matters.

Sử dụng kỹ thuật thống kê đa biến đánh giá chất lượng nước sông Như Ý tỉnh Thừa Thiên Huế

● Nguyễn Minh Kỳ

Trường Đại học Khoa học, Đại học Huế

● Nguyễn Hoàng Lâm

Trường Đại học Bách Khoa Đà Nẵng

TÓM TẮT

Mục đích của nghiên cứu nhằm đánh giá chất lượng nước bởi các chất dinh dưỡng và hữu cơ cũng như xác định áp lực môi trường, xem xét tác động tải lượng chất ô nhiễm lên sông Như Ý, tỉnh Thừa Thiên Huế. Với năm trạm lấy mẫu, nghiên cứu đã tiến hành quan trắc các thông số chất lượng nước như nhiệt độ, hàm lượng oxi hòa tan (DO), nhu cầu oxi sinh hóa (BOD_5), nhu cầu oxi hóa học (COD), nitrate (NO_3^-) và phosphate (PO_4^{3-}). Nghiên cứu sử dụng kỹ thuật thống kê đa biến như phân tích tương quan, phân tích thành phần chính (PCA), và phân tích cụm cluster (CA) để đánh giá chất lượng nước. Phân tích tương quan chỉ ra sự tồn tại mối liên hệ có ý nghĩa thống kê giữa các thông số chất lượng nước như nhiệt độ

với DO và BOD_5 với COD ($p < 0,01$). Kỹ thuật PCA được áp dụng để xem xét phân nhóm dữ liệu và chỉ ra các nhóm nhân tố làm thay đổi chất lượng nước. Kết quả PCA trích xoay nhân tố gồm hai nhóm chính với tổng phương sai 62,207%. Trong đó, nhóm nhân tố đầu tiên chiếm 40,873% tổng phương sai gồm các thông số nhiệt độ, DO, BOD_5 và COD. Nhóm nhân tố thứ hai bao gồm NO_3^- và PO_4^{3-} với 21,334% tổng phương sai, đồng thời được đặt tên và giải thích bởi quá trình xả thải liên quan đến các hoạt động nông nghiệp. Tương tự, kết quả phân tích CA cũng xác lập và phân làm nhóm lần lượt BOD_5 , COD, Temp, DO (nhóm 1) và NO_3^- , PO_4^{3-} (nhóm 2).

Từ khóa: Quan trắc, sông Như Ý, môi trường, kỹ thuật thống kê đa biến, chất lượng nước.

REFERENCES

- [1]. Ajibade, W. A., Ayodele, I. A. and Agbede, S. A., Water quality parameters in the major rivers of Kainji Lake National Park, Nigeria, *African Journal of Environmental Science and Technology*, 2, 7, 185-196 (2008).
- [2]. Amadi, A. N., Quality Assessment of Aba River using heavy metal pollution index, *American Journal of Environmental Engineering*, 2, 1, 45-49 (2012).

-
- [3]. Gleick, P.H., Water in crisis: A guide to the world's fresh water resources, Oxford University Press, New York (1993).
- [4]. Hach C. C., Klein R. L. Jr. and Gibbs C. R., Introduction to Biochemical Oxygen Demand, Technical Information Series, Booklet No. 7, Hach Company, U.S.A (1997).
- [5]. Munirah Abdul Zali, Ananthy Retnam and Hafizan Juahir, Spatial Characterization of Water Quality Using Principal Component Analysis Approach at Juru River Basin, Malaysia, *World Appl. Sci. J.*, 14, 55-59 (2011).
- [6]. Singh, A. K., Mondal, G. C., Kumar S., Singh, T. B., Tewary, B. K., & Sinha, A., Major ion chemistry, weathering processes and water quality assessment in upper catchment of Damodar River basin, *Indian Environ. Geo.*, 54, 4, 745-758 (2008).
- [7]. Vietnam Ministry of Natural Resources and Environment, QCVN 08:2008/BTNMT: National technical regulation on surface water quality, Hanoi (2008).
- [8]. Thua Thien Hue People's Committee, Geographical conditions and natural resources of Thua Thien Hue Province, Hue (2012).
- [9]. Thua Thien Hue Statistical Office, Statistics yearbook of Thua Thien Hue Province 2011, Thanh Nien Publishing House (2012).
- [10]. Tong, S.T.Y. and W. Chen, Modeling the relationship between land use and surface water quality, *J. Environ. Manage*, 66, 377-393 (2002).
- [11]. Yeung, I.M.H., Multivariate analysis of the Hong Kong Victoria Harbour water quality data, *Environmental Monitoring and Assessment, Netherlands*, 59, 331-342 (1999).
- [12]. Waziri. M. and Ogugbuaja V.O., Interrelationships between physicochemical water pollution indicators: A case study of River Yobe-Nigeria, *Am. J. Sci. Ind. Res.*, 1, 1, 76-80 (2010).
- [13]. World Health Organization, Health hazards from nitrate in drinking-water, Report on a WHO meeting, Environmental Health Series No. 1, Copenhagen (1984).
- [14]. World Health Organization, Nitrate and nitrite in drinking-water, WHO Press, Geneva, Switzerland (2011).
- [15]. Wu M.-L., Y.-S.Wang, C.-C.Sun, H.Wang, Z.-P.Lou, J.-D.Dong, Using chemometrics to identify water quality in Daya Bay, China, *Oceanologia*, 51, 2, 217-232 (2009).