Science on Natural Resources and Environment **33** (2020) 52-61

Science on Natural Resources and Environment
Journal homepage: tapchikhtnmt.hunre.edu.vn

APPLICATION OF MULTIVARIATE STATISTICAL TECHNIQUES IN SELECTING SURFACE WATER QUALITY MONITORING SITES AT BUNG BINH THIEN RESERVOIR, AN GIANG, VIETNAM

Nguyen Thanh Giao
Can Tho University, Viet Nam

Received 18 December 2019; Accepted 16 November 2020

Abstract

This study applied principal components analysis (PCA) and cluster analysis (CA) to assess water quality and recommend sites and water parameters for water quality monitoring at Bung Binh Thien reservoir, An Giang province, Vietnam based on 11 water samples collected in the dry season. The water variables including temperature (T), pH, electricity conductivity (EC), total dissolved solids (TDS), total suspended solids (TSS), depth (H), turbidity (NTU), dissolved oxygen (DO), biological oxygen demand (BOD), chemical oxygen demand (COD), sulfate (SO_4^{2-}), chloride (Cl^- ammonia (NH_4^+-N), nitrate (NO_3^--N), total nitrogen (TN), orthophosphate (PO_4^{3-}-P), total phosphorus (TP), and coliforms were used for interpretation of surface water quality and multivariate statistical analysis. The results indicated that water quality in the reservoir contaminated by suspended solid, organic matter was and fecal microbes. PCA identified five sources of water pollution accounting for 93.4% of water quality variation. CA showed that only two sampling points are needed for water monitoring in the reservoir. At each sampling point, water quality variables of temperature, pH, BOD (or COD), NO_3^--N, TP, and coliforms should be monitored and recommended by using PCA. This study suggests that multivariate statistical analysis could be useful for recommending locations and selecting water parameters for water quality monitoring.

Keywords: Bung Binh Thien; Cluster analysis; Water monitoring; Principal component analysis; Water quality

Corresponding author. Email: ntgiao@ctu.edu.vn

1. Introduction

Bung Binh Thien (BBT) reservoir is the biggest natural reservoir in Mekong delta, Vietnam. It is located in An Phu district, An Giang province with total surface areas of 200 ha in the dry and 800 ha in the wet season. The average depth of the reservoir is 4 m while the length and width are 2,900 m and 430 m, respectively (An Giang DONRE, 2012). At the BBT reservoir, where is currently a wetland with a diverse ecosystem has the potential to become the Ramsar site of Vietnam (Tri et al., 2012). Besides, BBT reservoir is also facing risks and challenges in the environment due to impacts of daily life, tourism and agricultural production. Therefore, monitoring the quality of water at BBT is essential for sustainable development of the local communities. However, the methods for selecting monitoring stations as well as water quality parameters were not well studied. Multivariate statistical techniques, for example cluster analysis (CA) and principal component analysis (PCA), could be used efficiently for supporting in making decisions related to water quality monitoring network. These methods could be used to handle a large number of gauging locations and several water quality variables (Zeinalzadeh and Rezaei, 2017) and designing sampling networks (Singh et al., 2005; Zeinalzadeh and Rezaei, 2017). This study aimed to use CA and PCA to design water monitoring sites and parameters at BBT.

2. Methodology

2.1. Water sampling and analysis

Water quality parameters, including temperature (ºC), pH, electricity conductivity (EC, µS/cm), total dissolved solids (TDS, mg/L), total suspended solids (TSS, mg/L), depth (H, m), turbidity (NTU), dissolved oxygen (DO, mg/L), biological oxygen demand (BOD, mg/L), chemical oxygen demand (COD, mg/L), sulfate (SO_4^{2-}, mg/L), chloride (Cl^-, mg/L) ammonia (NH_4^+-N, mg/L), nitrate (NO_3^--N, mg/L), total nitrogen (TN, mg/L), orthophosphate (PO_4^{3-}-P, mg/L), total phosphorus (TP, mg/L), and Coliform (MPN/100mL) were analyzed in this study. These water quality parameters were selected based on the literature review (Cho et al., 2009; Chounlamany et al., 2017; Zeinalzadeh and Rezaei, 2017). Ten water samples (S1 - S10) were collected inside the reservoir and one sample (S11) was collected in the river directly connected to the reservoir (Fig. 1).

Figure 6: Sampling locations in the Bung Binh Thien reservoir (Google Earth, 2019)

Table 3. Coordinates and description of study sites (long/lat)

No.	Site	X-coordinate	Y-coordinate	Description
1	S1	10°55'17.38"N	104°49'54.89"E	15 - 20 m from the reservoir bank influenced by domestic activities of the local people
2	S2	10°55'10.16"N	104°49'55.91"E	At the middle of the reservoir
3	S3	10°55'7.61"N	104°49'58.81"E	Close to the reservoir bank, 5 meters from the floating restaurant
4	S4	10°55'12.76"N	104°49'37.90"E	20 m from the reservoir bank
5	S5	10°55'7.33"N	104°49'37.04"E	At the middle of the reservoir
6	S6	10°55'1.50"N	104°49'40.59"E	20 m from the reservoir bank
7	S7	10°55'4.43"N	104°49'21.95"E	20 m from the reservoir bank influenced by duck and corn farming
8	S8	10°54'58.51"N	104°48'50.53"E	At the middle of the reservoir
9	S9	10°54'52.6"N	104°49'24.06"E	20 m from the reservoir bank influenced by tourist activities
10	S10	10°54'50.93"N	104°48'35.35"E	The intersection between Binh Di river and the reservoir
11	S11	10°54'47.75"N	104°48'21.03"E	On the Binh Di river

Water samples inside the reservoir were collected at the onset (S10) and at the middle (S4 - S9) and at the end of the reservoir (S1 - S3). Water samples were also collected at the positions close to the reservoir banks (S3, S6, S9, S1, S4, and S7) and at the middle of the reservoir (S2, S5, and S8). Samples were collected in the dry season (12 January 2019). Temperature, EC, TDS, DO, and depth were measured by handheld meters at the field. The other parameters were analyzed based on Standard methods for the Examination of Water and Wastewater (APHA, 1998).

2.2. Data Analysis

Water quality data among the sampling locations were analyzed mean and standard deviation (SD) and then compared with the National technical regulation on surface water quality (QCVN 08-MT:2015/BTNMT). Mean values of samples were also used as inputs for the analysis of Cluster Analysis (CA) and Principal Component Analysis (PCA). CA was used to classify spatial distribution of water quality according to the physical, chemical and biological variables. Sampling sites with similar characteristics in terms of pollution levels were grouped into the same cluster whereas different characteristics of pollution levels were grouped into another cluster based on Euclidean distance (a similarity measure) (Feher et al., 2016; Chounlamany et al., 2017). CA was performed according to the Ward's method (Salah et al. 2012). PCA was extensively used in multivariate analysis to extract important information from the original dataset (Cho et al., 2009; Feher et al., 2016; Chounlamany et al., 2017, Zeinalzadeh and Rezaei, 2017). PCA reduces initial data variables that do not contribute significantly to data variation while creating a group of new variables called principal components or primary factors (PCs or PFs). These PCs are not interconnected and appear in descending order of importance. The important factor to consider is the eigenvalue coefficient which measures the significance of the components. The larger the coefficient, the greater the contribution it could make to explain the variation of the original dataset. Varimax was widely used as a rotation method, each of the original data variables will be associated with one component and each component represents only a small set of variables (Feher et al., 2016). The correlation between the main components and the initial data variables was indicated by loadings (Feher et al., 2016). In this study, CA and PCA were performed using Primer 5.2 for Windows (PRIMER-E Ltd, Plymouth, UK).

2.3. Water quality index

Water Quality Index (WQI) calculated to compare with results of cluster analysis. It was calculated based on Equation 1 which is based on Decision No. 879/QD-TCMT of Vietnam National Environmental Agency (2011).

$$\mathrm{WQI} = \frac{\mathrm{WQI_{pH}}}{100}\left[\frac{1}{5}\sum_{a=1}^{5}\mathrm{WQI_a} * \frac{1}{2}\sum_{a=1}^{2}\mathrm{WQI_b} * \mathrm{WQI_c}\right]^{1/3} \qquad (1)$$

where WQIa is the WQI value of five parameters (i.e., DO, BOD_5, COD, N-NH_4^+, and P-PO_4^{3-}); WQIb is the WQI value of TSS; WQIc is the WQI value of Coliforms; WQI_{pH} is the WQI value of pH parameters (ranging from 6 to 8.5).

The WQI is ranging from 0 to 100 dividing water quality into five levels. Level 1 (100> WQI> 91) is good water quality that could be used for purposes of water supply. Level 2 (90>WQI>76) is also used for water supply for domestic uses but suitable treatment measures are required. Level 3 is for irrigation and other similar purposes (75>WQI>51). Level 4 (50>WQI>26) is the water suitable for transport and equivalent purposes while Level 5 (25>WQI>0) is considered to be heavily polluted water that proper treatment measures are urgently needed.

3. Results and discussion

3.1. Water quality in Bung Binh Thien reservoir

Table 2 presented the 18 water quality parameters at 11 sampling points in the dry season (January 2019). The depth of sampling locations ranged from 1.61 m to 4.6 m. Water temperature was in the range of 28.07 ± 0.06 - 30.33 ± 1.36 ºC. pH in water ranged from 7.55 ± 0.03 to 7.85 ± 0.01. pH values in the reservoir were higher than those in An Giang's rivers (6.9 to 7.1) during 2009 - 2016 (Ly and Giao, 2018). pH and temperature do not highly fluctuate and this is the common condition in tropical regions (Singh et al., 2005; Chounlamany et al., 2017). Total dissolved solids varied from 61.8 ± 0.26 to 68.57 ± 0.46 mg/L and the lowest value was found at the river point (S11). This could be considered as indication of ongoing conversion of organic pollutants into inorganic compounds which are taking place in the reservoir. High turbidity was found at S10 (11.43 ± 0.06 NTU) and S11 (9.03 ± 0.09 NTU) since these two points were in close relation with the river. Total suspended solids at S10 - S11 (53.0 ± 1.0 - 53.3 ± 0.6 mg/L) were significantly higher ($p<0.05$) than those of S1 - S9 (44.0 ± 1.0 - 50.3 ± 0.5 mg/L). It was found that DO ranged from 5.33 ± 0.06 to 9.17 ± 0.38 mg/L. The significant higher DO ($p<0.05$) were observed at the points inside the reservoir while DO at sites close to the river (S10) and in the river (S11) were significantly lower ($p<0.05$). Higher values of DO in the reservoir could be resulted from the diverse and abundant presence of phytoplankton and water hyacinth which release and diffuse oxygen into the water environment. DO in Bung Binh Thien reservoir were higher than those in several water bodies (4.0 to 5.2 mg/L) belonging to An Giang province in 2009 - 2016 (Ly and Giao, 2018). The high DO concentration could indicate better self-purification capacity of the reservoir. Biological oxygen demand was in the range of 9.33 ± 0.58 - 11.67 ± 0.58 mg/L, whereas chemical oxygen demand was in the range of 14.33 ± 0.58 - 17.67 ± 0.58 mg/L. Both BOD and COD were used as indicators of organic waste concentration in water (Galal-Gorchev et al., 1993; Kazi et al., 2009). BOD and COD were found higher at the end of the reservoir where there are presence of active human activities such as restaurant and cafeteria. BOD averagely accounts for 65.2 ± 1.1% of the COD indicating almost 35% of organic matter presence in the reservoir are recalcitrant substances. The values of organic matter in the reservoir exceeded the national standard 2.6 and 1.6 times for BOD and COD, respectively and this could potentially pose high threat to ecological and human health. Fortunately, dissolved oxygen is high, generating good environmental condition for decomposition of organic matter. BOD in BBT reservoir (9.33 ± 0.58 - 11.67 ± 0.58 mg/L) was significantly higher than that in Hau river and field canals (4.1 - 5.5 mg/L) (Ly and Giao, 2018) and this indicated that BBT reservoir is more organically polluted than other water bodies in An Giang province.

Chloride at study sites ranged from 7.2 ± 0.0 to 8.4 ± 0.0 mg/L and the lowest value was found at the river site (S11). The statistical analysis showed that chloride concentration in the river was significantly lower ($p< 0.05$) than those in the reservoir. It was also found that chloride concentrations at the end of the reservoir (S1 - S3) were significantly lower than those of the other gauging locations. In natural rivers, chloride concentration is often lower than 2 mg/L. This could mean that chloride concentration in the reservoir is relatively high because of the influence of human activities from the cafeteria and restaurant, especially using supply water containing chloride and release of wastewater. In some cases, chloride concentration could indicate fecal contamination. Surface water influenced by landfill leachate showed higher chloride concentrations of 17.2 - 43.3 mg/L (Chounlamany et al., 2017) while chloride concentrations in water body influenced daily activity, urban, and industry were from 32.8 ± 44.6 - 54.2 ± 32.9 mg/L (Zeinalzadeh and Rezaei, 2017). Sulfate concentration at the river site was significantly higher ($p<0.05$) than those in the reservoir which were in the range of 1.04 ± 0.00 - 2.25 ± 0.04 mg/L. Previous study reported that sulfate concentration in natural rivers was from 2 to 80 mg/L (UNICEF, 2008). As can be seen sulfate in the BBT reservoir is lower than the normal range.

Ammonium concentration were not detected (detection limit of 0.03 mg/L) at S1, S3, S4, S5, S7, S8, and S9 while it was detected at S2 (0.2 mg/L), S6 (0.04 mg/L), S10 (0.10 mg/L) and S11 (0.22 mg/L). Previous study found that ammonium concentration in the river ranged from 0.02 ± 0.0 - 0.9 ± 1.9 mg/L (Zeinalzadeh and Rezaei, 2017) which were higher than that in this study because the river was affected by many human activities. Nitrate concentration was very low at study sites, ranging from not detected (detection limit of 0.01 mg/L) at S4, S6 and S8 to 0.07 mg/L. Water bodies in An Giang province during 2009 - 2016 had nitrate concentration ranged from 0.03 to 1.76 mg N/L (Ly and Giao, 2018). Total nitrogen in the reservoir was found less than the detection limit of 1 mg/L. Zeinalzadeh and Rezaei (2017) found that total nitrogen ranged from 1.3 ± 1.8 - 3.0 ± 1.3 mg/L in the river. The difference of total nitrogen found in this study and the literature could be due to the presence or absence of nitrogen-degrading microbes and phytoplankton. The reservoir receives domestic wastewater from cthe ommunity but nitrogen could not be detected because the consumption of phytoplankton, microorganisms and aquatic plants (mainly water hyacinth). The national water quality standard advised ammonia and nitrate levels of 0.3 mg/L and 2 mg/L, respectively for environmental protection. Orthophosphate was also not detected (detection limit of 0.03 mg/L) at all sampling sites except at S11 (0.05 mg/L). During 2009 - 2016, orthophosphate concentration was detected in the river system in An Giang province ranging from 0.02 to 0.47 mg/L (Giao and Ly, 2018) which was higher than that detected in the BBT Reservoir in the dry season. The allowable concentration of orthophosphate according to the national regulation (QCVN 08-MT: 2015/BTNMT, Column A1) is 0.1 mg/L which is far exceeding the actual measurement at the study area. However, total phosphorus (TP) was found at very high concentrations (3.09 ± 0.01 - 9.2 ± 0.13 mg/L). The lowest TP were found at S7, S8, and S11 whereas the highest TP were recorded at S1, S6, and S10. TP at S11 was low, implying that TP could originate from internal sources such as domestic wastewater, agricultural, aquaculture and restaurant activities. In this river, phosphate is a nutrient limit. In this study, the high concentration of TP compared to orthophosphate indicated that phosphorus was not in bioavailable form but in polyphosphates or in the bodies of living organisms, especially phytoplankton. The main source of phosphorus could also be from detergent that is released from daily human activities. In natural water, TP is less than 0.1 mg/L and the discharge of domestic and industrial water and the drainage of agricultural land fertilized contributing to increase the concentration (Barakat et al., 2016). Further investigation is needed to explain for the unusual high of TP in the reservoir. Coliform density at the study site ranged from 1900 ± 346.41 to 9300 ± 0.00 MPN/100mL. Coliform levels at S4, S8, S10, and S11 exceeded the national standard surface water quality (allowable limit of 2500 MPN/100 mL) from 1.72 to 3.72 times. Previous study found that coliform density in rivers in An Giang province often exceeded the national regulation by 2.14 - 7.04 times (Ly and Giao, 2018). Findings revealed that the river water was more seriously contaminated with fecal microorganisms than the reservoir water. The sources of coliform contamination are from human and animal wastes, especially the fecal materials (Bolstad and Swank, 1997; UNICEF, 2008). Results from this study indicated that total suspended solids, organic matter, and coliform could impair water quality in Bung Binh Thien Reservoir.

Table 4. Water quality in Bung Binh Thien reservoir

Parameter	**Unit**	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**
Depth	m	1.61	3.22	3.22	2.99	3.68	4.37
Temp	°C	29.2 ± 0.0	29.0 ± 0.0	30.3 ± 0.1	29.5 ± 0.0	29.23 ± 0.3	30.1 ± 0.3
pH	-	7.64 ± 0.01	7.79 ± 0.01	7.85 ± 0.01	7.81 ± 0.04	7.56 ± 0.03	7.55 ± 0.03
EC	µS/cm	132.7 ± 0.2	133.1 ± 0.3	137.3 ± 0.9	131.7 ± 0.1	130.8 ± 0.1	130.4 ± 0.6
TDS	mg/L	66.3 ± 0.0	66.4 ± 0.1	68.6 ± 0.5	65.8 ± 0.1	64.9 ± 0.3	65.5 ± 0.1
Turbidity	NTU	4.67 ± 0.06	3.53 ± 0.06	3.28 ± 0.04	4.33 ± 0.08	6.25 ± 0.06	3.27 ± 0.09
TSS	mg/L	46.3 ± 0.6	46.7 ± 0.6	44.0 ± 1.0	50.3 ± 0.5	48.0 ± 1.0	47.7 ± 0.6
DO	mg/L	8.8 ± 0.2	8.9 ± 0.2	7.7 ± 0.2	9.0 ± 0.3	8.1 ± 0.1	9.2 ± 0.4
COD	mg/L	17.3 ± 0.6	17.7 ± 0.6	17.0 ± 0.0	17.7 ± 0.6	15.3 ± 0.6	15.0 ± 0.0
BOD_5	mg/L	11.3 ± 0.6	11.7 ± 0.6	11.0 ± 0.0	11.7 ± 0.6	10.0 ± 0.0	10.0 ± 0.0
Chloride	mg/L	8.4 ± 0.0	8.4 ± 0.0	8.4 ± 0.0	7.8 ± 0.0	7.8 ± 0.0	7.8 ± 0.0
Sulfate	mg/L	1.37 ± 0.01	1.29 ± 0	1.13 ± 0	1.45 ± 0.01	1.48 ± 0.01	1.04 ± 0
NH_4^+-N	mg/L	0.0 ± 0.0	0.2 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.04 ± 0.0
NO_3^--N	mg/L	0.02 ± 0.0	0.01 ± 0.0	0.01 ± 0.0	0.0 ± 0.0	0.04 ± 0.0	0.0 ± 0.0
PO_4^{3-}-P	mg/L	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
TN	mg/L	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
TP	mg/L	8.4 ± 0.09	6.92 ± 0.1	4.2 ± 0.09	4.22 ± 0.06	6.63 ± 0.1	9.2 ± 0.13
Coliforms	MPN/ 100mL	1900 ± 346	2400 ± 0	2200±173	4300±0	2300±173	2300±0

Parameter	**Unit**	**S7**	**S8**	**S9**	**S10**	**S11**
Depth	m	2.99	3.68	3.45	1.70	4.60
Temp	°C	29.4 ± 0.3	30.3 ± 1.4	29.6 ± 0.3	29.3 ± 0.1	28.1 ± 0.1
pH	-	7.71 ± 0.03	7.74 ± 0.01	7.56 ± 0.01	7.7 ± 0.01	7.75 ± 0.03
EC	µS/cm	133.9 ± 0.7	129.6 ± 0.4	130.7 ± 0.4	128.2 ± 0.3	123.7 ± 0.6
TDS	mg/L	66.9 ± 0.4	64.8 ± 0.3	65.3 ± 0.3	64.1 ± 0.2	61.8 ± 0.3

Turbidity	NTU	5.33 ± 0.05	6 ± 0.05	4.9 ± 0.06	11.43 ± 0.06	9.03 ± 0.09
TSS	mg/L	47.0 ± 0.0	48.3 ± 0.6	49.7 ± 0.6	53.0 ± 1.0	53.3 ± 0.6
DO	mg/L	8.0 ± 0.1	7.6 ± 0.1	8.0 ± 0.1	6.1 ± 0.1	5.3 ± 0.1
COD	mg/L	16.0 ± 0.0	14.3 ± 0.6	15.3 ± 0.6	17.3 ± 0.6	15.3 ± 0.6
BOD_5	mg/L	10.0 ± 0.0	9.3 ± 0.6	10.0 ± 0.0	11.3 ± 0.6	10.0 ± 0.0
Chloride	mg/L	7.8 ± 0.0	7.8 ± 0.1	8.4 ± 0.0	7.8 ± 0.0	7.2 ± 0.0
Sulfate	mg/L	1.41 ± 0.00	1.50 ± 0.00	1.29 ± 0.00	1.96 ± 0.01	2.25 ± 0.04
NH_4^+-N	mg/L	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.0	0.22 ± 0.0
NO_3^--N	mg/L	0.05 ± 0.01	0.0 ± 0.0	0.07 ± 0.0	0.02 ± 0.0	0.07 ± 0.0
PO_4^{3-}-P	mg/L	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.05 ± 0.0
TN	mg/L	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
TP	mg/L	3.09 ± 0.01	3.20 ± 0.0	5.50 ± 0.0	8.05 ± 0.15	3.17 ± 0.06
Coliforms	MPN/ 100mL	2200 ± 173	4600 ± 0	2300 ± 0	9300 ± 0	9300 ± 0

3.2. Determination of locations for water monitoring

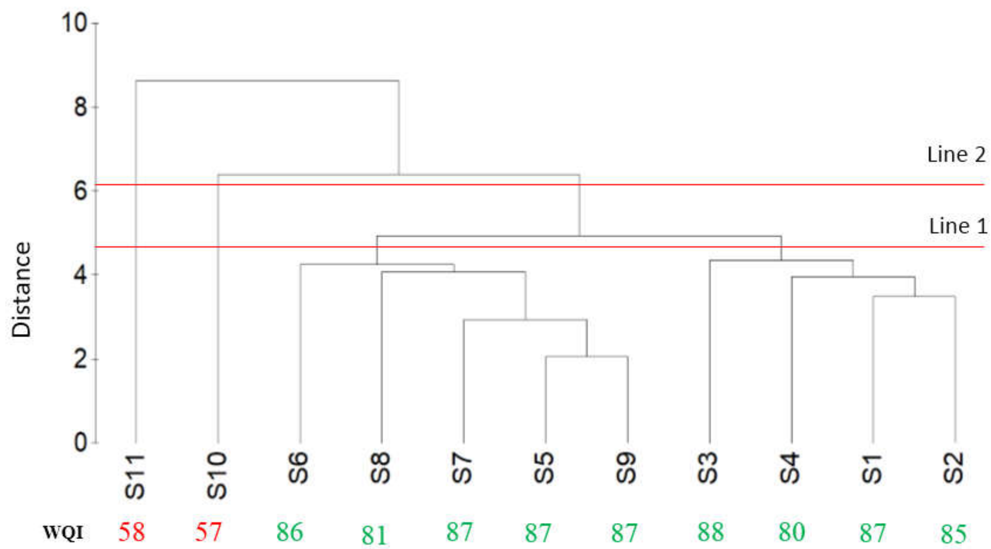

Figure 7: Spatial variation of water quality at Bung Binh Thien reservoir

Water quality data at 11 sampling sites in BBT reservoir was analyzed using cluster analysis. The water quality dataset was grouped into four separated clusters (Figure 2, Line 1).

Cluster 1 and Cluster 2 comprise of S11 and S10, respectively while Cluster 3 included location S5 to S9 and Cluster 4 contained from S1 to S4. Cluster 1 represents for river water quality that was more polluted than Cluster 3 and 4. Cluster 2 is at the site directly influenced by the river water at the sampling time (high tide) thus showing similar water quality characteristics with the Cluster 1. Cluster 3 included sampling points at the middle of the reservoir while Cluster 4 comprised of points at the end of the reservoir. Cluster analysis results were compared with water quality index (WQI) which was calculated based on 9 water quality parameters (Temperature, pH, turbidity, TSS, DO, BOD, COD, NH_4^+-N, PO_4^{3-}-P, and coliform) at every sampling location according to the guideline of National Environmental Agency (2011). It was found that water quality at S1 - S9 (WQI = 80 - 88) fell into the category of green (good for the purposes of human use but it should be treated before supplying for domestic use). Water quality at S10 (WQI = 57) and S11 (WQI = 58) was not suitable for domestic water supply and it can be used for irrigation only. Water quality in BBT reservoir could be simply categorized in two groups based on WQIs (Figure 2, Line 2). Non-metric multi-dimensional scaling (NMDS) was also used to present the multidimensional data visually. In this study, multidimensional scaling analysis was used to represent the difference and similarity of the water quality at the sampling sites through a non-metric multidimensional scale. As can be seen from Figure 3, points that are close together represent samples that are very similar in water quality characteristics while points that are apart correspond to very different water quality traits. Besides that, cluster and Non-Metric Multi-Dimensional Scaling (NMDS) analyses showing the clustering of the sampling sites within the groups were almost similar.

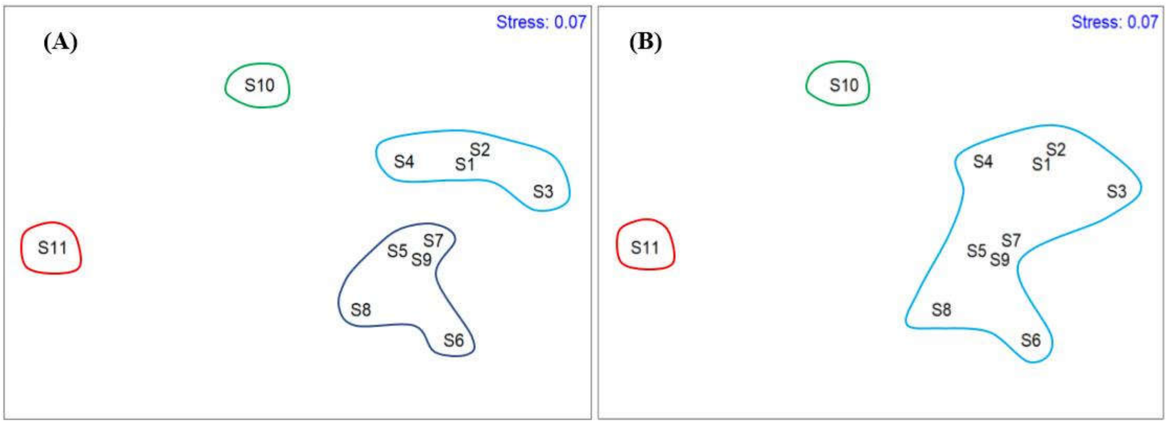

Figure 8: Results of NMDS analysis for water quality monitoring at BBT reservoir

Based on cluster analysis and water quality index, the clusters in Figure 2 could be reorganized using MDS (Figure 3A and B). Figure 3A corresponds to Line 1 which group sampling sites into three clusters while Figure 3B corresponds to Line 2 (group them into two clusters). Separation of sampling sites in the reservoir into two groups are still valid since WQI also indicated the same outcomes. The latter separation could lead to a design for optimal monitoring strategy with minimal sampling points at a reducing financial spending (Wang et al., 2013). Results suggested that two water quality monitoring stations in the reservoir (1 station at the onset and one at the end of the reservoir) could be established. This can reduce the cost (approximate 30%) for water quality monitoring at BBT reservoir since the middle sampling sites are removed.

3.3. Determination of water pollution sources and water parameters for monitoring

PCA was applied with 18 water quality parameters at 11 sampling locations. The eigenvalues each PC and loadings of water quality parameters were shown in Table 2. Five PCs significantly contributed to the explanation for 93.4% of the total variation of water quality in the BBT Reservoir from the original dataset. The PC1 explained 48.6%, PC2 19.0%, PC3 11.0%, PC4 8.3% and PC5 6.5% of total variance of the whole set of water quality data at the reservoir. This could indicate that variation of water quality in the reservoir is highly complicated that could be from contribution of various pollution sources. Contribution of water quality parameters to the PCs are identified using correlation coefficients between PCs and variables, also called loadings (Shaw, 2003). The strong correlation between PCs and each water quality parameter was established when absolute loading factors greater than 0.75, moderate correlation when absolute loading factors between 0.75 and 0.5, weak when absolute loading factors 0.50 - 0.30 (Liu et al., 2003). As can be seen from Table 3, PC1 was weakly contributed by EC, TDS (negative) and TSS (negative), chloride (negative), sulfate, phosphate and coliform (positive). PC1 is contributed by several water quality parameters with low loadings that could indicate PC1 representing diffuse sources of pollution including biogenic and anthropogenic origins. PC2 was moderately contributed by BOD, COD (positive), and weakly by depth (negative) and pH (positive). The PC2 represents organic pollution related to anthropogenic sources (Li and Liao, 2003). PC3 was moderately contributed by pH (positive) and total phosphorus (negative). PC4 was weakly contributed by temperature (negative) and NH_4^+-N, NO_3^--N (positive). PC5 was weakly contributed by depth, NH_4^+-N (positive) and moderately by NO_3^--N (negative). PC3 and PC5 could represent sources of pollution from human activities since nutrients (NH_4^+-N, NO_3^-N, and TP) show most influenced on these PCs. Depth (PC5), temperature (PC4) and pH (PC3) are attributed to physical-chemical properties of water in the reservoir. PCA results showed that the water parameters should be tracked including temperature, pH, BOD (or COD), NO_3^--N, and TP because these parameters show the most influenced on water quality in the study area. These parameters have also been recorded in several previous studies (Minh et al., 2019; Giao, 2020a; Giao, 2020b). In addition, coliform should also be included in the monitoring program.

Table 3. PCA for water quality parameters at Bung Binh Thien reservoir and river

Water quality Variables	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**
Depth	0.108	-0.309	0.245	0.217	**0.406**
Temp	-0.237	-0.190	0.030	**-0.486**	0.133
pH	0.002	0.313	**0.578**	-0.157	0.137
EC	-0.320	0.062	0.198	0.010	-0.184
TDS	-0.323	0.063	0.193	-0.010	-0.142
Turbidity	0.282	0.145	-0.191	-0.269	-0.239
TSS	0.303	0.086	-0.184	-0.133	0.039
DO	-0.295	-0.054	-0.109	0.162	0.257
COD	-0.107	**0.516**	0.033	0.108	-0.036
BOD_5	-0.102	**0.511**	-0.027	0.122	0.111
Chloride	-0.280	0.123	-0.061	0.174	-0.215
Sulfate	0.326	0.153	0.015	-0.082	-0.135
NH_4^+-N	0.214	0.202	0.086	0.388	0.332
NO_3^--N	0.186	-0.163	0.060	0.389	**-0.586**
PO_4^{3-}-P	0.290	-0.023	0.221	0.279	0.075
TP	-0.095	0.117	**-0.616**	0.167	0.245
Coliform	0.295	0.168	0.031	-0.323	0.151
Eigenvalue	8.26	3.23	1.87	1.42	1.10
Variability (%)	48.6	19.0	11.0	8.3	6.5
Cumulative (%)	48.6	67.6	78.5	86.9	93.4

4. Conclusion

The BBT reservoir is risk of eutrophication where the limiting factor is caused mainly by nitrogen rather than phosphorus. Water quality in the reservoir is still suitable for domestic water supply but proper treatments are required. The main concern for water quality in the reservoir in the dry season was TSS, organic matters, and coliforms. Cluster analysis revealed that two sampling station (S1 and S10) within the reservoir should be selected for water quality monitoring. PCA identified five factors occupying 93.4% of total variation in the quality of water. PCA results also showed that water quality variables including temperature, pH, BOD (or COD), NO_3^--N, TP, and coliform should be monitored. Future study should investigate sources of pollution for ensuring good water quality due to the important role of BBT reservoir for the nearby community.

REFERENCES

[1]. American Public Health Association (1998). *Standard Methods for the Examination of Water and Wastewater*. 20^{th} edition, Washington DC, USA.

[2]. Barakat A, Baghdadi ME, Rais J, Aghezzaf B, Slassi M., (2016). *Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques*. International Soil and Water Conservation Research, 4(4): 284 - 292.

[3]. Bolstad, PV, Swank WT (1997). *Cumulative impacts of land-use on water quality in a southern Appalachian watershed.* Journal of the American Water Resources Association, 33(3): 519 - 533.

[4]. Boyacioglu H, Boyacioglu H., (2008). *Water pollution sources assessment by multivariate statistical methods in the Tahtali Basin, Turkey*. Environment Geology, 54: 275 - 282.

[5]. Cho KH, Park Y, Kang JH, Ki SJ, Cha S, Lee SW, Kim JH., (2009). *Interpretation of seasonal water quality variation in the Yeongsan Reservoir, Korea using multivariate statistical analyses.* Water Science and Technology, 59(11): 2219 - 2226.

[6]. Chounlamany V, Tanchuling MA, Inoue T., (2017). *Spatial and temporal variation of water quality of a segment of Marikina River using multivariate statistical methods.* Water Science and Technology, 66(6): 1510 - 1522.

[7]. Feher IC, Zaharie M, Oprean I., (2016). *Spatial and seasonal variation of organic pollutants in surface water using multivariate statistical techniques.* Water Science and Technology, 74: 1726 - 1735.

[8]. Galal-Gorchev H, Ozolins G, Bonnefoy X., (1993). *Revision of the WHO guidelines for drinking water quality.* Annalidell'IstitutoSuperiore di Sanità, 29: 335 - 345.

[9]. Giao NT (2020a). *Evaluating current water quality monitoring system on Hau River, Mekong delta, Vietnam using multivariate statistical technique.* Journal of Applied Environmental Research, 42(1): 14 - 25.

[10]. Giao NT (2020b). *Surface water quality in aquacultural areas in An Giang province, Vietnam*. International Journal of Environment, Agriculture and Biotechnology, 5(4), 1054 - 1061.

[11]. Kazi TG, Arain MB, Jamali MK, Jalbani N, Afridi HI, Sarfraz RA, Shah AQ (2009). *Assessment of water quality of polluted reservoir using multivariate statistical techniques: A case study.* Ecotoxicology and Environmental Safety, 72(20): 301 - 309.

[12]. Li JX, Liao WG., (2003). *An analysis on the possibilities of eutrophication in the Three Gorges Reservoir.* Science and Technology Review, 9, 49 - 52.

[13]. Liu CW, Lin KH, Kuo YM., (2003). *Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan*. Science of the Total Environment, 313, 77 - 89.

[14]. Ly NHT, Giao NT., (2018). *Surface water quality in canals in An Giang province, Viet Nam, from 2009 to 2016.* Journal of Vietnamese Environment, 10(2): 113 - 119.

[15]. Minh HVT, Kurasaki M, Ty TV, Tran DQ, Le KN, Avtar R, Rahman M and Osaki M., (2019). *Effects of multi-dike protection systems on surface water quality in the Vietnamese Mekong Delta.* Water, 11, 1010.

[16]. Ministry of Environment and Natural Resources (2015). *QCVN 08-MT: 2015/BTNMT National technical regulation on surface water quality*.

[17]. Salah EAM, Turki AM, Othman EMA., (2012). *Assessment of water quality of Euphrates River using cluster analysis.* Journal of Environmental Protection, 3: 1629 - 1633.

[18]. Shaw PJA (2003). *Multivariate Statistics for the Environmental Sciences*. Hodder-Arnold, UK.

[19]. Singh KP, Malik A, Sinha S (2005). *Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques – a case study*. Analytica Chimica Acta, 538: 355 - 374.

[20]. Tri TN, Dat HD and Sang NV (2012). *Researching the biodiversity of fish fauna in Bung Binh Thien wetland, An Giang province.* Journal of Biology, 34: 21 - 29.

[21]. Ty DV, Huy NH, Da CT, Ut VN and Viet TV (2017). *Evaluation of water quality fluctuations in Bung Binh Thien, An Giang province.* Science Journal of Can Tho University, 125 - 131.

[22]. UNICEF (2008). *Handbook on Water Quality*. United Nations Children's Fund (UNICEF), New York.

[23]. Vega M, Pardo R, Barrado E, Debán L., (1998). *Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis*. Water Research 1998, 32: 3581 - 3592.

[24]. Vietnam National Environmental Agency (2011). *Decision No. 879 /QĐ-TCMT on guideline for water quality index calculation*.

[25]. Wang P, Wang Y, Bai Y, Tian Z, Li J, Shao X, Mustavic, LF, Li BL., (2013). *Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China.* Journal of Hydro-Environment Research, 7: 30 - 40.

[26]. Zeinalzadeh K, Rezaei E., (2017). *Determining spatial and temporal changes of surface water quality using principal component analysis*. Journal of Hydrology: Regional Studies, 13: 1 - 10.