Analysis of frequencies and parameters for monitoring surface water quality using multivariate statistical approaches

Nguyen Thanh Giao, Vo Quang Minh*

(College of Environment and Natural Resources, Can Tho University, Can Tho, 900000X, Vietnam)

Abstract: The main objective of this study is to assess the monitoring system of surface water quality in Can Tho using principal component analysis (PCA) and cluster analysis (CA). Water samples were monthly collected in 2020 at 35 sampling sites and then analyzed for 18 parameters including temperature, pH, turbidity, dissolved oxygen (DO), biological oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), ammonium (NH_4^+-N), nitrate (NO_3^--N), nitrite (NO_2^--N), orthophosphate (PO_4^{3-}-P), chromium (Cr), fluoride (F^-), arsenic (As), lead (Pb), mercury (Hg) and coliform. These values were compared with the national technical regulation on surface water quality and calculated water quality index (WQI). The findings indicated that surface water quality in Can Tho was contaminated by organic matters, suspended solids, and microorganisms. TSS, BOD and COD were high in the dry season, while high level of coliform was found in the wet season. The results of computed WQI showed that the surface water quality was from good to excellent. The PCA results show that 77.2% of water quality fluctuations in the study area were potentially affected by domestic, agricultural, industrial activities and hydrological factors (rainwater runoff, riverbank erosion). Moreover, the results of PCA revealed that water quality parameters could be diminished from 18 to 14 parameters without the loss of important for evaluating surface water quality. CA suggested that total sampling sites and frequencies in the current monitoring program are likely to reduce from 35 to 27 sites and 12 to 6-7 times/year, respectively. It is suggested that these findings can provide essential scientific information for local environmental managers in adjusting the surface water quality monitoring network in Can Tho city, Vietnam.

Keywords: CA, Can Tho, Coliform, PCA, surface water; water quality index.

Citation: Giao, N. T., and V. Q. Minh. 2024. Analysis of frequencies and parameters for monitoring surface water quality using multivariate statistical approaches in Can Tho city, Vietnam. Agricultural Engineering International: CIGR Journal, 26(1):1-14.

1 Introduction

In recent years, surface water quality in Can Tho is being increasingly degraded by industrial production, business, and daily activities (Department of Natural Resources and Environmental [DoNRE], 2020). Polluted water quality also leads to many serious consequences for human health, the environment, and ecology (Su et al., 2010). Therefore, water quality monitoring is essential in detecting the impact of socio-economic development on water quality. In Vietnam, environmental monitoring includes waste monitoring and environmental monitoring, which is carried out through automatic and continuous monitoring, periodical monitoring, and monitoring at the request of competent state agencies.

National environmental monitoring includes a network of stations for monitoring the background environment and the impact environment to serve the monitoring, providing information on the quality of the background environment and impact in areas with

Received date: 2022-02-25 **Accepted date:** 2024-01-03

*Corresponding author: Vo Quang Minh, Professor, College of Environment and Natural Resources, Can Tho University, Can Tho, Vietnam. Tel: +84 0913604101. Email: vqminh@ctu.edu.vn.

inter-regional, inter-provincial, and cross-border regions. Provincial environmental monitoring is a network of stations and locations for monitoring the background environment and the impact environment to serve the monitoring and provide information about the quality of the background and the impact environment in the areas of a province.

Investment projects, establishments, concentrated production, business and service zones, industrial clusters that generate waste into the environment must be monitored (Vietnam National Assembly [VNA], 2020). Regular monitoring creates a large data set with many monitoring indicators, locations, and frequencies. Currently, monitoring data is only evaluated through comparison with national technical regulations on surface water quality (Ministry of Natural Resource and Environmental [MoNRE], 2015) and water quality index (WQI) (Vietnam Environment Administration [VEA], 2019). Meanwhile, multivariate analysis approaches such as cluster analysis (CA), principal component analysis (PCA), discriminant analysis (DA) are commonly used in analyzing spatial and temporal fluctuations of surface water quality.

These methods can assist in proposing analytical criteria, sampling location and frequency, and potential water pollution sources (Muangthong and Shrestha, 2015; de Andrade Costa et al., 2020; Ma et al., 2020). Therefore, multivariate analyzes could help optimize the monitoring network (Pinto et al., 2019; Pratama et al., 2020; Ioele et al., 2020; Liu et al., 2021). Although research on water quality assessment using multivariate statistics in Vietnam have been very diverse (Du et al., 2019; Giao et al., 2021a; Giao et al., 2021b), it has not been commonly implemented in Can Tho city. This is a central city for agricultural, industrial and service development of the Mekong Delta region. Therefore, this study was conducted to utilize multivariate statistics including CA and PCA to analyze the spatiotemporal variations of surface water quality in Can Tho and thereby determine the location, criteria, and monitoring frequency. The results can provide valuable scientific information supporting local authorities to reconsider the surface water environment monitoring system in the study area.

2 Materials and methods

2.1 Study area description

Can Tho city cover an area of 140,894.9 ha and is the center of the Mekong Delta region (accounting for 3.5% of the total delta area). The city has five urban districts (Ninh Kieu, Binh Thuy, Cai Rang, O Mon, and Thot Not) and four rural districts (Phong Dien, Co Do, Thoi Lai, and Vinh Thanh). Can Tho city border An Giang to the North, Hau Giang to the South, Kien Giang to the West, and Vinh Long and Dong Thap to the East. The gross regional domestic product in 2020 is estimated to increase by 1.02% compared to 2019. It is contributed by 0.17% from the agriculture, forestry and fishery sector (increasing by 1.58%), 0.56% from the industry and construction sector (increasing by 1.69%), 0.25% from the service sector (increasing by 0.51%), and 0.04% from product tax minus product subsidies (increasing by 0.55%) (DoNRE 2020). The hydrological regime of flow on the river and canal system in Can Tho City is governed by the flow of the Mekong River through the Hau River, the tides of the East Sea, intra-regional rain, and the infrastructure system. Especially, the competition of influence between flow regime upstream of the Mekong River and the tidal regime of the East Sea is most dominant. The average density of rivers and canals in Can Tho city is quite large, about 1.8 km km^{-2}. Especially along the Hau River in Ninh Kieu, O Mon, Cai Rang and Thot Not districts, it can be up to over 2 km km^{-2}. The main system of rivers and canals in Can Tho consists of the Hau River and a system of small canals. Hau River is the western branch of the Mekong River in the territory of Vietnam, and the main source of freshwater supply for the Mekong Delta and Can Tho City. It is considered and the natural boundary of Can Tho City with Dong Thap and Vinh Long provinces.

Hau River is also an international waterway for ships going to Cambodia and upstream countries. It is the largest river in the region with a total length of 55 km flowing through Can Tho city, and the total amount of Hau River water flowing into the sea is about 200 billion m^3/year (accounting for 41% of the total water volume of the Mekong River). The average water flow in Can Tho River is 14 800 m^3 s^{-1}. The total alluvium of the Hau River is 35 million m^3/year that accounts for nearly half of the total alluvium in the Mekong River.

The system of three small canals includes the 16 km long Can Tho canal, which empties into the Hau River at Ninh Kieu pier, Binh Thuy canal, Tra Noc canal, O Mon canal, Thot Not canal, Cai San canal, etc. Large capacity leads water from Hau River to flow inland areas and connects to canals of neighbouring provinces, with freshwater all year round, both for irrigation in the dry season and for drainage in the flood season. Can Tho be located in the tropical monsoon climate with steamy weather having two distinct seasons: the rainy season (from May to November) and the dry season (from December to April). According to the environmental report, the centralized wastewater treatment system in Can Tho has not been constructed, except in the Ninh Kieu district under testing operation. Surface water in Can Tho city is currently affected by domestic, aquaculture, agricultural and industrial activities. Solid waste is also a great concern for surface water quality.

2.2 Water sampling and analysis

Surface water quality monitoring data were collected from the Department of Natural Resources and Environment of Can Tho in 2020. With the goal of assessing the characteristics of waste sources, the surface water quality monitoring program of Can Tho city has been built with a sampling frequency of 12 times/year at 35 monitoring locations. Sampling locations were determined in areas affected by domestic, industrial, agricultural waste sources, and the junction area between canals and large rivers.

Descriptions of the 35 sampling locations are shown in Figure 1, including two locations in Rach Tham Tuong (TT1 and TT2), three locations in Rach Cai Khe (CK1, CK2 and CK3), one location in Xang Thoi Lake (CK4), two locations in Bun Xang Lake (BX1 and BX2), two locations in Cai Son-Hang Bang canal (HB1 and HB2), three location in Ba Lang canal (BL1, BL2 and BL3), one location on Cai Rang river (CR1), three locations in Sang Trang canal (ST1, ST2 and ST3), one location in Cam canal (RC1), two locations in Cay Me canal (CM1 and CM2), three locations in Cai Chom canal (CC1, CC2 and CC3), three locations in Bo Ot canal (BO1, BO2 and BO3), two locations in Tra Nien canal (TN1 and TN2), one position on Xang canal (BD1), one position on Xao Xeo canal (XX1), one position on Thi Doi canal (TD1), one position on KH6 canal (KNT1), one position on canal located at Co Do canal (CD1), one location at Cai San canal (CS1) and one location at Sau Bong canal (SB1).

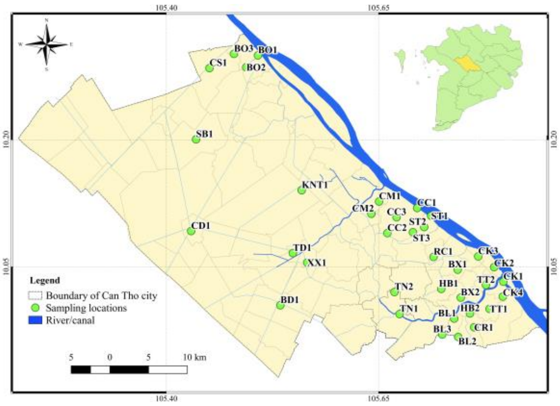

Figure 1 Sampling location in Can Tho city in 2020

Eighteen parameters including temperature, pH, turbidity, dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), ammonium (NH_4^+-N), nitrate (NO_3^--N), nitrite (NO_2^--N), orthophosphate (PO_4^{3-}-P), iron (Fe), chromium (Cr), fluorine (F^-), arsenic (As), lead (Pb), mercury (Hg) and coliform were determined to assess water quality and provide data for multivariate statistical analysis. Water samples collected and stored according to the standards of national. pH, temperature, DO, and turbidity of collected samples were measured in the field, while the remaining parameters were analyzed in the laboratory by standard methods (American Public Health Association [APHA], 2012). The parameters, units, analytical methods and allowable limits are presented in Table 1.

Table 1 Variables, analytical methods, and limits of surface water quality

Variables	Unit	Analytical methods	Standard values[1]
pH	-	TCVN 6492:2011	6-8.5
Temp	ºC	SMEWW 2550B:2012	-
Turb	NTU	Turbid meter	-
DO	mg L^{-1}	TCVN 7325:2004	≥ 6
BOD	mg L^{-1}	TCVN 6001-1:2008	4
COD	mg L^{-1}	TCVN 6491:1999	10
TSS	mg L^{-1}	TCVN 6625:2000	20
NH_4^+-N	mg L^{-1}	TCVN 6179:1996	0.3
NO_2^--N	mg L^{-1}	SMEWW 4500-NO_2^-.B:2012	0.05
NO_3^--N	mg L^{-1}	TCVN 6180:1996	2
PO_4^{3-}-P	mg L^{-1}	SMEWW 4500-PO_4^{3-}.E:2012	0.1
Fe	mg L^{-1}	TCVN 6177:1996	0.5
Cr	mg L^{-1}	SMEWW 3500-Cr.B:2012	0.01
F^-	mg L^{-1}	SMEWW 4500-F-.B&D:2012	1
As	mg L^{-1}	TCVN 6626:2000	0.01
Pb	mg L^{-1}	SMEWW 3113B:2017	0.02
Hg	mg L^{-1}	TCVN 7877:2008	0.001
Coliform	MPN/ 100 mL	TCVN 6187-2:1996	2,500

Note: [1] MoNRE (2015)

2.3 Data processing

Surface water quality data of 35 sampling sites from January to December in 2020 were averaged before inclusion in multivariate statistical analysis. CA was used to group water quality at the sampling sites according to the Ward method (Arora et al., 2014). CA results were presented in the form of a dendrogram showing locations with similar water quality characteristics that will be grouped into a cluster. Accordingly, appropriate monitoring locations can be figured out. PCA was employed to determine the main pollutant parameters and sources through the Eigenvalues. The larger this coefficient is, the more significant it is in explaining the variability of the data and should preferably be greater than 1 (Boyacioglu and Boyacioglu, 2008). Data variables are described by load factor correlation that is classified into three levels as “strong”, “moderate” and “weak” with absolute values of load factor greater than 0.75, from 0.5 to 0.7, and from 0.3 to 0.5, respectively (Liu et al., 2003; Huang et al., 2010). PCA and CA analysis were performed using the Primer V5.2 for Windows licensed software (PRIMER-E Ltd, Plymouth, UK). WQI was calculated according to the guidance of the Vietnam Environment Administration (VEA, 2019), as shown in the Equation 1:

$$WQI = \frac{WQI_I}{100} \times \frac{(\prod_i WQI_{II})^{\frac{1}{m}}}{100} \times [(\frac{1}{k}\sum_{i=1}^{k} WQI_{III})^2 \times \frac{1}{l}\sum_{i=1}^{l} WQI_{IV}]^{\frac{1}{3}} \quad (1)$$

where, WQI_I is the calculated WQI value for pH parameter; WQI_{II} is the calculated WQI value for heavy metals (i.e., Cr, As, Pb and Hg); WQI_{III} for calculated WQI value for 7 parameters (i.e., DO, BOD, COD, NH_4^+- N, NO_3^--N, NO_2^--N and PO_4^{3-}-P); WQI_{IV} for the calculated WQI value for coliform. Water quality is classified into 6 levels, as follows: Level 1 “excellent” (WQI = 91-100) is used for domestic water supply purposes, level 2 “good” (WQI = 76-90) for domestic water supply purposes but needs appropriate treatment measures, level 3 “medium” (WQI = 51-75) for irrigation and other equivalent purposes, level 4 “bad” (*WQI* = 26-50) for navigation and other equivalent purposes, level 5 “poor” (*WQI* = 10-25) water is heavily polluted and needs future treatment measures, and level 6 “Very poor” (*WQI* < 10) contaminated water and need to take remedial and treatment measures.

3 Results and discussion

3.1 Summary of surface water quality in Can Tho city in 2020

The average pH value ranged from 7.04±0.27 to 7.64±0.24 (Figure 2a), and the temperature was varied from 26.80 ℃ -29.93 ℃. Although temperature and pH had slight variation, these values were within the tolerance limits of aquatic organisms (Kale, 2016). Turbidity and TSS ranged from 11.69±3.55 to 31.11±16.33 NTU (Figure 2b) and 25.07±5.54 to 45.53±22.39 mg L^{-1} (Figure 2c), respectively. Turbidity and TSS had seasonal variations, namely that these values in January tended to be higher than in November. Turbidity is not specified in Vietnamese standards. TSS in the water bodies of Can Tho exceeded from 1.30-2.27 times the allowable limit (20 mg L^{-1}) (MoNRE, 2015). Previous studies also showed that TSS in most water bodies in the Mekong Delta is higher than the allowable limit (Tuan et al., 2019; Giao, 2020a) due to riverbank erosion and rainwater runoff (Lim et al., 2020; Giao, 2020b). On the other hand, high TSS degrades water quality and is costly for domestic water treatment (Giao et al., 2021b).

The average DO concentration from January to December tended to decrease (Figure 2d). The highest DO concentration recorded in April and May was 6.07 mg L^{-1}, and the lowest in January was 5.06±0.52 mg L^{-1}. The average BOD in 2020 ranged from 5.06±1.21 to 13.03±3.25 mg L^{-1} (Figure 2e). Meanwhile, the COD value over the months in the study area varied significantly from 11.12 to 23.50 mg L^{-1}. DO, BOD and COD also had seasonal fluctuations; that is, the values of these parameters in March, April and May tended to be higher than in September and October. Compared with standard value, the DO concentration was below the allowable threshold (≥ 6 mg L^{-1}), while the BOD and COD levels both exceeded several times the allowable thresholds (4 mg L^{-1} and 10 mg L^{-1}, respectively). The low DO concentration may be due to heavy rain leaching dissolved organic matters into surface water and reducing dissolved oxygen concentration through biodegradation (Kannel et al., 2008). In Figure 2f, the values of COD had a higher than the BOD in most of the months. It can be seen that the study area was mainly contaminated with biodegradable organics by domestic activities (Giao et al., 2021b), wastewater from industrial and agricultural activities (Zheng et al., 2008).

The concentrations of NO_2^--N and NO_3^--N in surface water were relatively low, ranging from 0.01±0.00 to 0.04±0.02 mg L^{-1} (Figure 2g) and 0.35±0.26 to 1.18±0.47 mg L^{-1} (Figure 2i), respectively. NO_2^--N and NO_3^--N values in 2020 were within the allowable limits (0.05 mg L^{-1} and 2 mg L^{-1}, respectively). Besides, the average NH_4^+-N content in 12 months ranged from 0.17±0.09 to 0.37±0.28 mg L^{-1}, and the average dissolved PO_4^{3-}-P concentration was in the range of 0.03±0.02 – 0.14±0.07 mg L^{-1}. NH_4^+-N and PO_4^{3-}-P values were below the allowable thresholds (0.3 mg L^{-1} and 0.1 mg L^{-1}, respectively) in most of the months.

According to Lomoljo et al. (2009), agricultural activities are one of the primary sources of nitrate, ammonium and orthophosphate present in surface water. The average coliform value had a great difference between months, from 2360±633.1 to 4634±4761.7 MPN/100 mL. Except for February and October, the remaining months of the year had coliform content exceeding the allowable limit (2500 MPN/100 mL) by 1.25-1.85 times. This result is also consistent with the study of Giao et al. (2021b) that detected coliform in May, August and November greater the allowable limit from 1.3 to 1.4 times. Divya and Solomon (2016) suggested that coliform in water originates from human and animal wastes.

The concentrations of heavy metals (Fe, Cr, As, Pb and Hg) in water in 2020 were very low. Fe and F^-

concentrations were highest in November and December at 0.47±0.49 mg L^{-1} and 0.41±0.26 mg L^{-1}, respectively, which was below the allowable thresholds (Fe≤0.5 mg L^{-1}, F^-≤1 mg L^{-1}). The highest concentration recorded in January for the Cr parameter was 0.02±0.02 mg L^{-1}, exceeding column A1 of QCVN 08-MT: 2015/BTNMT (0.01 mg L^{-1}). Particularly, As, Pb, and Hg tended to decrease from January to December, and some months were in the undetectable range. The research results showed that the surface water quality in Can Tho city, Vietnam has seasonal fluctuations. Turbidity, TSS, DO, BOD, COD, NO_2^-- N, PO_4^{3-}- P, and Cr contents in the dry season months were higher than in the rainy season. In contrast, NH_4^+- N, NO_3^--N, coliform, and Fe values in the rainy season were higher than in the dry season. In addition, three months of the dry season (January, March and April) and two months of the rainy season (November and December) had BOD, COD, NH_4^+-N, NO_2^-- N, NO_3^--N, PO_4^{3-}-P, Fe, Cr and F^- exceeding the permissible thresholds in Table 1.

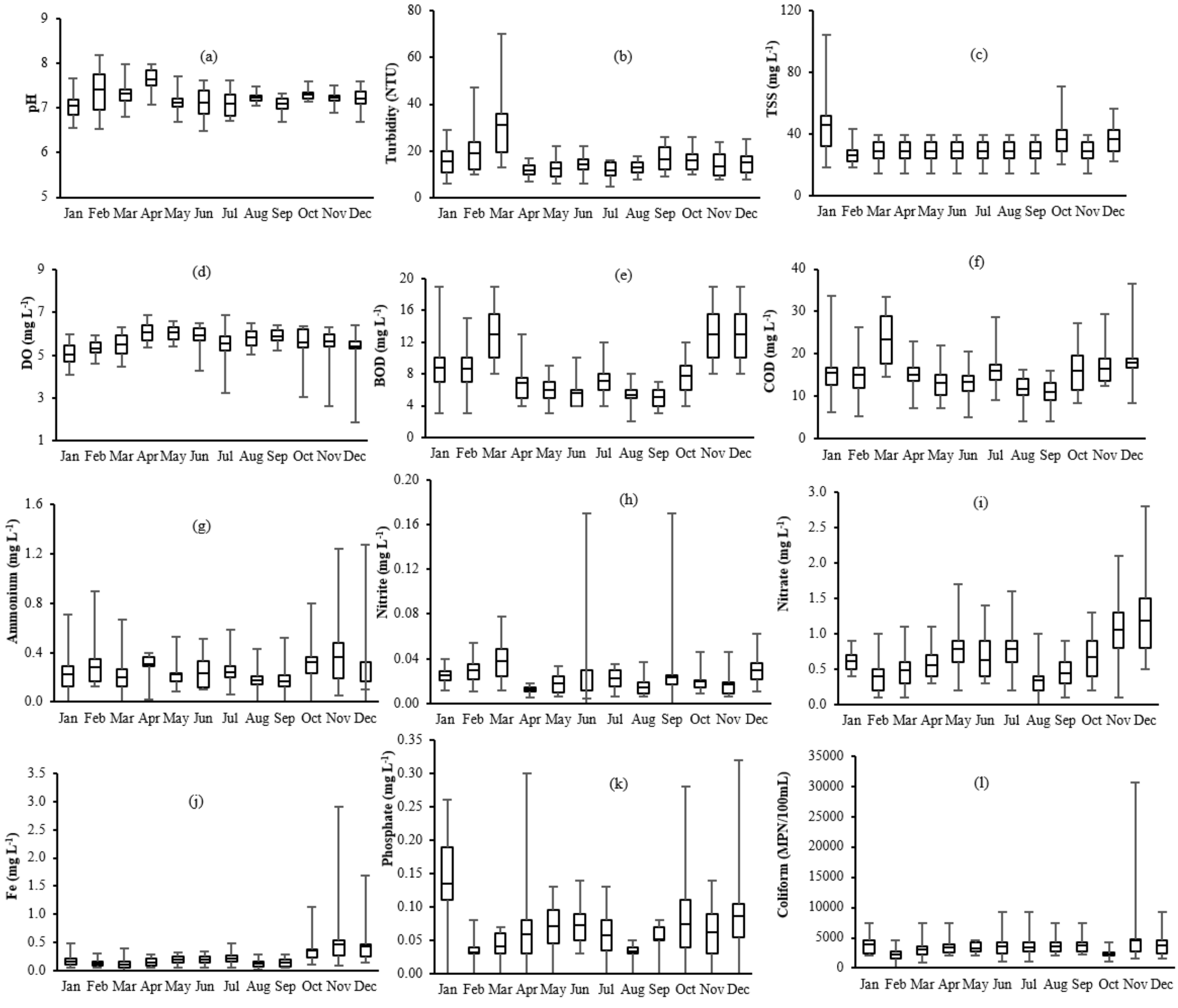

Figure 2 Boxplot showing surface water quality variations in Can Tho

The average values of 18 physicochemical parameters were used to calculate the WQI. WQI values at 35 monitoring points in Can Tho's rivers and canals are shown in Figure 3. The results showed that the overall water quality in Can Tho ranged from good to excellent. WQI values from 76 to 90 accounted for 57.14% of locations were assessed as areas with good water quality, and 42.86% of survey sites with WQI values from 91 to 100 classified excellent water quality. Good water quality was found mainly in the areas affected by industrial activities and domestic activities from residential areas. This partially indicates the impact of social-economic development on surface water quality in Can Tho city. Previous research showed that WQI values in water bodies of An Giang province ranged from poor to unsuitable for drinking (Minh et al., 2020). On the other hand, Giao et al. (2021b) reported that WQI values indicated surface water quality in Dong Thap province ranging from moderate to poor. Thus, it can be seen that the water quality of Can Tho is evaluated better than the two provinces of An Giang and Dong Thap.

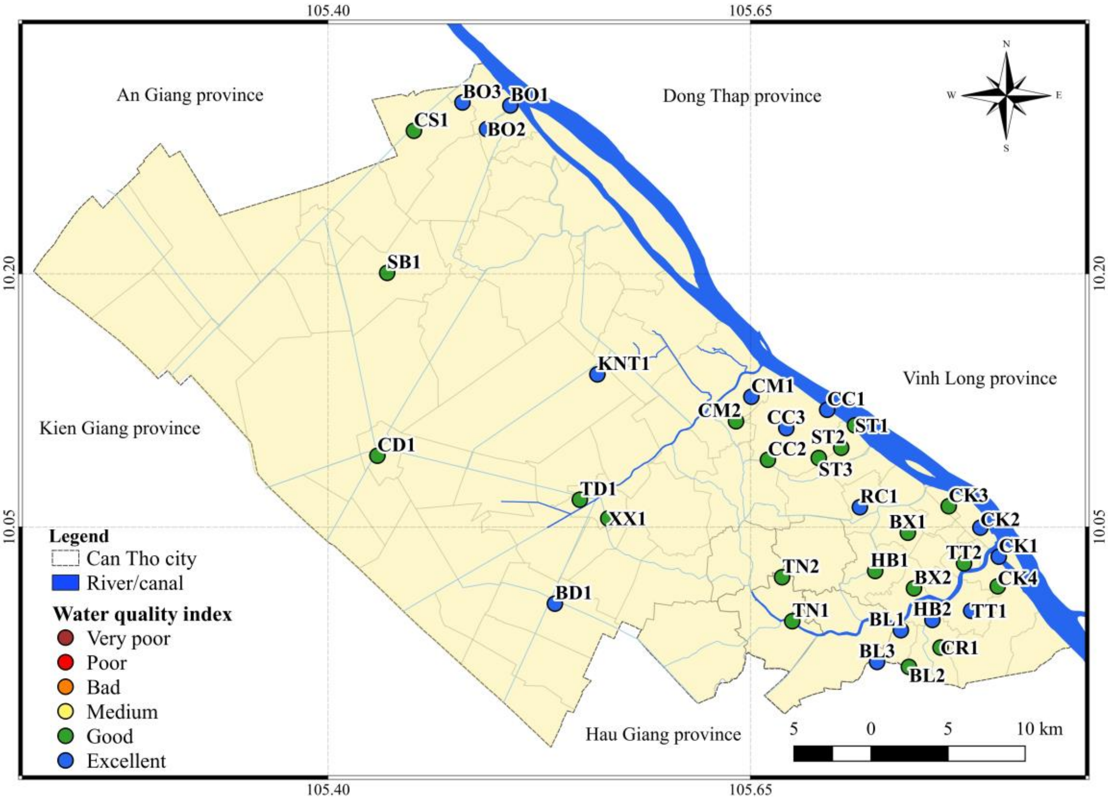

Figure 3 Spatial distribution of overall water quality in Can Tho in 2020

3.2 Identification of pollution sources and monitoring parameters

The principal component analysis result is presented in Table 2. Variation in surface water quality in the study area were explained by 11 factors (Table 2). In particular, PC1, PC2, PC3, PC4, PC5 and PC6 were considered the main sources of pollution as Eigenvalues coefficients greater than 1 (Boyacioglu and Boyacioglu, 2008; Howladar et al., 2018; Kale et al., 2020). These six PCs explained 77.2% of water quality variations in the study area.

PC1 was the most important factor, explaining 36.6% of water quality variations. This source of pollution can be derived from domestic and agricultural activities with the significant contribution of DO, BOD, COD, NH_4^+-N v à PO_4^{3-}-P (Edokpayi et al., 2017; Howladar et al., 2018; Njimou et al., 2021). Similar to previous studies, the DO was inversely correlated with other parameters (Rodrigues et al., 2020; Arora and Keshari, 2021). PC2 also explained 12.1% of the variation with the moderate correlation of coliform (0.51), and the weak correlation of Cr (0.47) and F^- (0.40). Colifrom originated from human and animal feces (Divya and Solomon, 2016; Niyoyitungiye et al., 2020), whereas Cr and F^- were derived primarily from industrial activities (Qin and Tao, 2022). PC3, PC4, PC5 and PC6 accounted for 9.0%, 7.9%, 6.0% and 5.7% of water quality variations in the study area, respectively. The negative correlation between NO_2^--N and Cr, Pb, Hg were recored in PC3. PC4 had moderate contribution of As (-0.52), and weak contribution of NO_3^--N (0.31), coliform (0.32), Hg (-0.43) and pH (-0.46).

In this study, pH had a positive correlation with As and Hg. In the previous study of Niu et al. (2023) also showed a positive correlation between pH and metals such as Cr, Cu, Pb and Zn. In addition, PC6 also recorded a positive correlation between Pb (0.45) and temperature (0.72). The orgin of Cd, Hg, Cd, Pb, As, Cr was industrial waste (Liu et al., 2019; Hoang et al., 2020; Astatkie et al., 2021; Qin and Tao, 2022). PC5 showed the contribution of pH, turbidity, TSS, NO_2^--N, PO_4^{3-}-P at a weak correlation. The occurrence of turbidity and TSS comes from variations in hydrological conditions such as rainwater runoff and riverbank erosion (Lim et al., 2020; Giao, 2020b). It can be implied that domestic, agricultural, and industrial activities, and rainwater runoff, riverbank erosion were the main sources of pollution to water quality in the study area. All parameters being evaluated in this monitoring program (DO, BOD, COD, NH_4^+-N, NO_3^--N, NO_2^--N, PO_4^{3-}-P, coliform, Fe, Cr, F^-, As, Pb, Hg, pH, TSS, turbidity and temperature) had significant contributions to pollution fluctuations. Therefore, all these parameters need to continue to be monitored and evaluated (Liu et al., 2003).

Table 2 Key indicators affecting surface water quality in Can Tho in 2020

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
pH	0.10	0.23	0.16	-0.46	-0.34	-0.09	0.07	-0.56	0.38	-0.19	-0.06
T	0.01	-0.27	-0.26	-0.07	0.13	0.72	0.20	-0.27	0.09	-0.37	0.16
Turb	0.28	-0.13	0.09	-0.19	0.39	-0.24	-0.20	-0.02	0.02	-0.43	0.09
DO	-0.32	0.10	0.22	0.09	0.02	0.10	-0.32	-0.20	-0.23	0.10	-0.12
BOD	0.35	0.10	-0.08	-0.06	0.16	0.12	-0.09	0.00	-0.04	0.12	-0.38
COD	0.35	0.10	-0.15	-0.11	0.03	0.13	0.10	0.13	0.00	0.04	-0.22
TSS	0.26	-0.11	0.20	0.17	0.34	-0.11	-0.41	-0.34	-0.03	-0.08	-0.10
NH4+-N	0.36	-0.01	-0.06	0.14	-0.14	0.02	0.01	0.12	0.02	0.07	-0.12
Fe	0.33	-0.01	-0.09	0.11	-0.11	-0.15	0.25	0.22	-0.11	-0.25	-0.08
NO2--N	-0.12	0.10	-0.35	-0.11	-0.39	0.12	-0.65	0.27	0.09	-0.34	-0.06
NO3--N	0.28	-0.11	0.03	0.31	-0.17	-0.04	-0.22	-0.01	0.02	-0.01	0.64
PO43--P	0.30	-0.14	0.17	0.00	-0.37	0.06	0.01	0.00	0.22	0.19	0.11
Cr	-0.02	0.47	0.31	-0.06	0.28	0.21	0.09	0.27	0.29	-0.06	0.38
F	0.17	0.40	-0.03	-0.01	-0.23	-0.04	0.06	-0.28	-0.70	-0.01	0.21
As	0.17	0.16	-0.20	-0.52	0.22	0.14	-0.24	0.12	-0.11	0.44	0.20
Pb	0.14	-0.23	0.50	0.02	-0.18	0.45	-0.16	0.05	-0.05	0.20	-0.14
Hg	-0.06	-0.23	0.44	-0.43	-0.12	-0.09	0.07	0.36	-0.36	-0.28	-0.01
Coliform	0.06	0.51	0.20	0.32	0.00	0.22	-0.02	0.12	0.00	-0.29	-0.22
Eigen.V	6.58	2.17	1.63	1.41	1.08	1.03	0.88	0.70	0.64	0.50	0.44
%Var.	36.6	12.1	9.0	7.9	6.0	5.7	4.9	3.9	3.5	2.8	2.4
%C.Var.	36.6	48.6	57.7	65.5	71.5	77.2	82.1	86.0	89.5	92.3	94.7

3.3 Assessment of frequency and location of current surface water monitoring in Can Tho

The results of cluster analysis based on the frequency of monitoring are presented in Figure 4. The surface water quality in 12 months was divided into 06 distinct clusters, namely Cluster I (January), Cluster II (February and March), Cluster III (June), Cluster IV (April, May, July, October), Cluster V (November and December) and Cluster VI (August and September). The characteristics of the climate in the study area are divided into the rainy season (May - October) and the dry season (November - April next year). Therefore, the water quality in Clusters I, II and V were considered representative of the dry season. Meanwhile, Cluster III, Cluster IV and Cluster VI represent the rainy season.

In general, the concentrations of nutritional parameters (NH_4^+-N, NO_3^--N, NO_2^--N and PO_4^{3-}-P), and metals and nonmetals (Fe, Cr, Pb, Hg, As and F^-) in all clusters were within the permissible limits of QCVN 08-MT: 2015/BTNMT (Table 3). Meanwhile, DO, BOD and COD at Cluster I, Cluster II, Cluster IV and Cluster V did not meet national standard. Most notably, the concentration of TSS exceeded the limit value in all clusters. Cluster I had a higher concentration of TSS than the other clusters. Besides, Cluster I was differentiated by the highest concentration of Cr, and the lowest DO and pH. This can be a source of industrial waste (Raji et al., 2016; Suteja and Purwiyanto, 2018). Cluster II had value of turbidity 4 times higher than Clusters III, Cluster IV and Cluster V. Although not exceeding the limits of national standard, the concentration of nutritional parameters in Cluster IV was higher than the other clusters. It can be a source of pollution from agricultural activities (Elisante and Muzuka, 2016; Chen et al., 2020). The water quality of Cluster IV and Cluster V had very high concentrations of coliform, BOD and COD.

The CA results shown that the water quality in the study area was mainly polluted by suspended solids, organic matter and microorganisms. The concentrations of turbidity, TSS, BOD, COD, NO_3^--N and PO_4^{3-}-P in the dry season higher than the rainy season. The seasonal fluctuation trend of these parameters had been determined by Abd Wahab et al. (2018) and Woldeab et al. (2018). The main origin can be derived from hydrological conditions, domestic, agricultural and industrial activities. Due to similar characteristics in the same cluster, monitoring frequency can be considered to reduce to save costs and resources to perform monitoring. Specifically, time that can be reduce are February or March (Cluster II), April or October or May or July (Cluster IV), November or December (Cluster V), and August or September (Cluster VI). To ensure that the monitoring times are evenly distributed and fully reflect the nature of all clusters, the study proposes monitoring timelines including, January, March, May, June, October and December. Therefore, the frequency of monitoring can be done about 7 times/year instead of 12 times/year, helping to save 41.67% of monitoring costs per year.

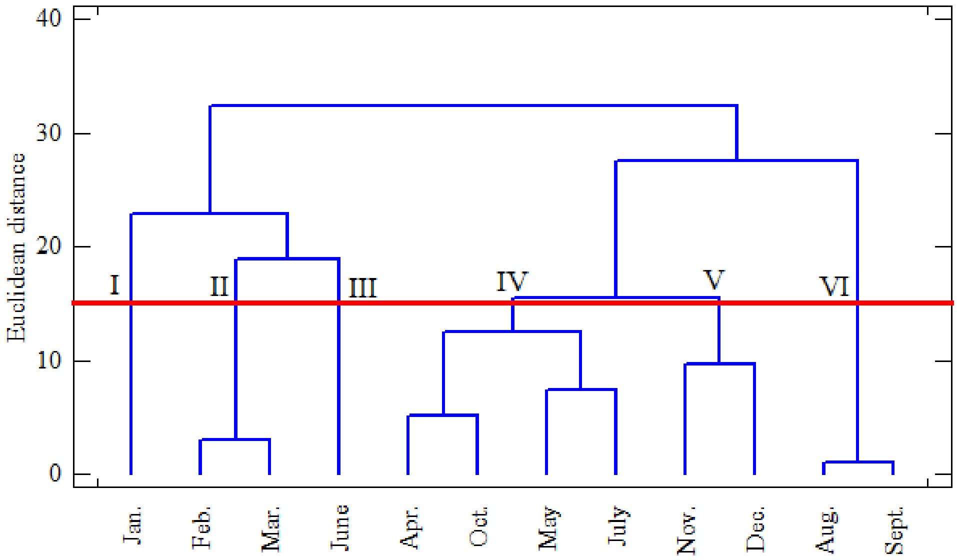

Figure 4 Clustering the frequency of surface water quality

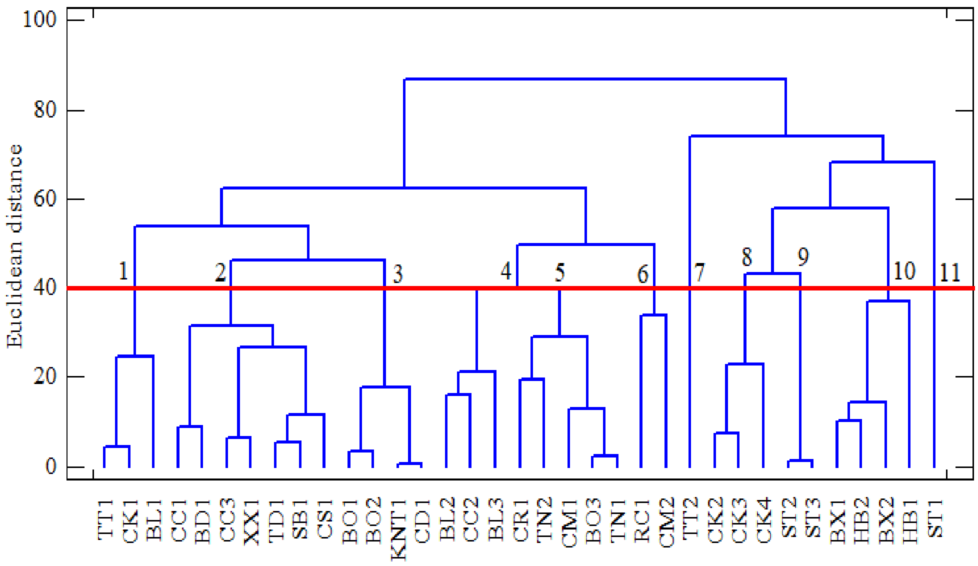

Figure 5 Clustering locations of surface water quality

Table 3 Temporal variation of surface water quality in the identified clusters

Variables	Clus. I	Clus. II	Clus. III	Clus. IV	Clus. V	Clus. VI	Standard values[1]
pH	6.94	7.61	7.21	7.28	7.15	7.20	6-8.5
Temp.	27.30	28.75	28.20	27.13	27.00	26.80	-
Turb	20.00	47.50	12.00	11.25	10.50	10.50	-
DO	5.23	5.45	6.36	5.92	5.75	6.47	≥ 6
BOD	6.00	8.50	4.00	7.50	9.00	2.50	4
COD	13.60	15.75	10.30	16.43	10.95	5.05	10
TSS	40.00	28.25	26.50	26.13	24.25	26.50	20
NH_4^+-N	0.00	0.07	0.26	0.27	0.09	0.03	0.3
Fe	0.10	0.15	0.06	0.21	0.20	0.09	0.5
NO_2^- -N	0.02	0.03	0.04	0.01	0.01	0.00	0.05
NO_3^- -N	0.40	0.10	0.40	0.53	1.35	0.40	2
PO_4^{3-}-P	0.11	0.05	0.03	0.08	0.09	0.03	0.1
Cr	0.04	0.00	0.00	0.00	0.00	0.00	0.01
F^-	0.00	0.05	0.00	0.18	0.10	0.00	1
As	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Pb	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Hg	0.00	0.00	0.00	0.00	0.00	0.00	0.001
Coliform	2300	1950	2100	2350	3150	2600	2500

Note: [1] MoNRE (2015)

Table 4 Spatial variation of surface water quality in the identified clusters

Variables	Clus. 1	Clus. 2	Clus. 3	Clus. 4	Clus. 5	Clus. 6	Clus. 7	Clus. 8	Clus. 9	Clus. 10	Clus. 11	Standard values[1]
pH	7.26	7.24	7.13	7.17	7.23	7.31	7.24	7.30	7.24	7.26	7.30	6-8.5
Temp.	27.44	27.57	27.47	27.51	27.41	27.62	27.53	27.43	27.45	27.40	27.33	-
Turb	17.94	16.29	15.73	13.19	13.20	11.34	22.67	20.55	18.79	14.94	15.58	-
DO	5.87	5.64	5.73	5.87	5.77	5.60	4.47	5.30	5.60	5.79	5.90	≥ 6
BOD	6.64	8.35	7.94	6.83	7.63	8.54	11.00	10.50	10.04	8.85	8.50	4
COD	11.39	15.46	14.80	12.81	14.11	16.87	22.78	19.45	18.22	15.31	15.93	10
TSS	31.46	32.21	31.82	26.59	25.18	20.19	38.71	37.99	37.50	37.35	28.50	20
NH_4^+-N	0.15	0.23	0.24	0.20	0.22	0.25	0.61	0.35	0.34	0.32	0.20	0.3
Fe	0.18	0.19	0.21	0.18	0.21	0.23	0.62	0.29	0.36	0.22	0.16	0.5
NO_2^--N	0.02	0.03	0.02	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.05
NO_3^--N	0.53	0.71	0.68	0.57	0.60	0.54	1.01	0.68	0.68	0.81	0.55	2
PO_4^{3-}-P	0.05	0.06	0.06	0.04	0.06	0.07	0.14	0.08	0.06	0.09	0.05	0.1
Cr	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01
F	0.08	0.10	0.12	0.13	0.11	0.23	0.18	0.18	0.30	0.19	0.22	1
As	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Pb	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Hg	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.001
Coliform	2747	3147	3100	3428	3701	3363	3600	2967	4425	3736	5586	2500

Note: [1] MoNRE (2015)

The result of cluster analysis based on the monitoring locations is shown in Figure 5. As can be seen that the 35 sampling sites were classified into 11 clusters. Monitoring sites in urban areas were grouped in Clusters 1 (TT1, CK1, BL1), Cluster 4 (BL2, BL3, CC2), Cluster 7 (TT2), Cluster 8 (CK2-CK4) and Cluster 10 (HB1, HB2, BX1, BX2). Meanwhile, Cluster 2 (CC1, BD1, CC3, XX1, TD1, SB1, CS1), Cluster 3 (BO1, BO2, KNT1, CD1), Cluster 5 (CR1, TN1, CM1, BO3), Cluster 6 (RC1, CM2), Cluster 9 (ST2, ST3) and Cluster 11 (ST1) were monitoring locations in rural areas.

The CA results showed that all clusters had concentrations of BOD, COD, TSS and Coliform higher than the limits of QCVN 08-MT: 2015/BTNMT (Table 4). Cluster 1 was distinguished from other clusters by the low concentration of BOD, COD, NH_4^+-N, NO_3^- -N, PO_4^{3-}-P, F^- and coliform. Meanwhile, Cluster 7 has the most severe concentrations of organic, nutrient, suspended matter and Fe contamination. Domestic and agricultural activities had contributed greatly to nutrient pollution

(NH_4^+-N và PO_4^{3-}-P) (Truc et al., 2019). The Fe concentration originates from the soil leaching (Phu et al., 2021). Water quality of Cluster 9 had a higher concentration of F^- than the other clusters, especially 3 times higher than Cluster 1–Cluster 5. This cluster had sampling locations collected around the industrial zones. The highest concentrations of coliform were recorded in Cluster 11 originating from human and animal fecal (Divya and Solomon, 2016; Niyoyitungiye et al., 2020). It can be seen that water quality in the study area is influenced by natural activities (rainwater runoff, leaching, erosion) and human activities (living, agriculture, industry). Therefore, water quality needs to be monitored and evaluated. CA analysis can help in optimizing monitoring costs by considering reducing the number of monitoring sites with similar water quality in the same cluster (Chounlamany et al., 2017; Giao, 2020a). This study suggest that the sampling locations that need to be considered for reduction were CC1 or CC3 (Cluster 2), BO1 or BO2 (Cluster 3), BL2 or BL3 (Cluster 4), TN1 or TN2 (Cluster 5), CK3 or CK4 (Cluster 8), ST2 or ST3 (Cluster 9), BX1 or BS2 and HB1 and HB2 (Cluster 10). Therefore, the current monitoring network can reduce 08 sampling locations, saving 22.86% of total monitoring costs.

4 Conclusion

The research results showed that the water quality in the water bodies of Can Tho was contaminated with organic matters, suspended solids, and coliform. Surface water quality in the study area was seasonally varied. The parameters of turbidity, TSS, DO, BOD, COD, NO_2^--N, PO_4^{3-}-P and Cr in the dry season were higher than those in the rainy season, while the parameters of NH_4^+-N, NO_3^--N, coliform and Fe in the rainy season were higher than those in the dry season. pH and temperature were not subjected to seasonal variation. Heavy metals were either not detected or still within the allowable limits. The WQI suggested that water quality in Can Tho city ranges from good to excellent. The PCA results identified six main sources (PC1-PC6), explaining 77.2% of water quality variations. Potential water pollution sources are mainly natural factors (hydrological factors, riverbank erosion) and anthropogenic impacts (agricultural, domestic and industrial activities). CA results presented that water quality was classified into eight clusters reducing the sampling locations from 35 to 27 and sampling frequency from 12 to 6-7 times per year. The key variables influencing surface water quality are pH, temperature, DO, COD, TSS, NH_4^+-N, PO_4^{3-}-P, Fe, Cr, As, F^-, Pb, Hg, and coliform. The findings provide scientific information for local environmental managers reconsidering the monitoring sites, frequencies, and parameters in building the surface water monitoring program. Further studies should focus on investigating the contribution of each water pollution source to propose appropriate solutions for better water quality management.

References

Abd Wahab, N., M. Khairul Amri Kamarudin, M. Ekhwan Toriman, F. Marcus Ata, H. Juahir, A. Ghazali, and A. Anuar. 2018. The evaluation of dissolved oxygen (DO), total suspended solids (TSS) and suspended sediment concentration (SSC) in Terengganu River, Malaysia. *International Journal of Engineering & Technology*, 7: 44-48.

APHA (American Public Health Association). 2012. *Standard Methods for the Examination of Water and Wastewater*. 20th ed. Washington, D.C.: American Public Health Association.

Arora, N., A. S. Arora, S. Sharma, and A. S. Reddy. 2014. Use of cluster analysis – a data mining tool for improved water quality monitoring of River Satluj. *International Journal of Advanced Networking Applications*, 6(2): 63-68.

Arora, S., and A. K. Keshari. 2021. Dissolved oxygen modelling of the Yamuna River using different ANFIS models. *Water Science and Technology*, 84 (10-11): 3359–3371.

Astatkie, H., A. Ambelu, and E. Mengistie. 2021. Contamination of stream sediment with heavy metals in the Awetu watershed of Southwestern Ethiopia. *Frontiers in Earth Science*, 9: 658737.

Boyacioglu, H., and H. Boyacioglu. 2008. Water pollution sources assessment by multivariate statistical methods in the Tahtali Basin, Turkey. *Environmental Geology*, 54: 275-282.

Chen, F., Q. Lao, Z. Li, P. Bian, Q. Zhu, C. Chen, and Z. Song. 2020. Monthly variations of nitrogen isotope of ammonium in wet deposition in a tropical city, South China. *Aerosol an Air Quality Research*, 20(5): 1062-1069.

Chounlamany, V., M. A. Tanchuling, and T. Inoue. 2017. Spatial and temporal variation of water quality of a segment of Marikina River using multivariate statistical methods. *Water Science and Technology*, 76(6): 1510-1522.

De Andrade Costa, D., J. P. Soares de Azevedo, M. A. dos Santos, and R. dos Santos Facchetti Vinhaes Assumpção. 2020. Water quality assessment based on multivariate statistic and water quality index of a strategic river in the Brazilian Atlantic Forest. *Scientific Reports*, 10(1): 22038.

DoNRE (Department of Natural Resources and Environment). 2020. Report on the current state of the environment in Can Tho city for the period 2015 - 2020. Hanoi, Vietnam: DoNRE.

Divya, A. H., and P. A. Solomon. 2016. Effects of some water quality parameters especially total coliform and fecal coliform in surface water of Chalakudy river. *Procedia Technology*, 24: 631-638.

Du, L. V., T. H. Dan, N. T. Giao, N. T. T. Anh, P. Q. Thai, T. V. Son, and L. T. H. Nga. 2019. Multivariate statistical application in the assessment of surface water quality in U Minh Ha - Ca Mau national park. *Science Journal of Can Tho University*, 2: 70-76.

Edokpayi, J. N., J. O. Odiyo, and O. S. Durowoju. 2017. Impact of wastewater on surface water quality in developing countries: a case study of South Africa. In *Water Quality*, ed. H. Tutu, ch. 18, 401-416. London, UK: Intech Open.

Elisante, E., and A. N. N. Muzuka. 2016. Sources and seasonal variation of coliform bacteria abundance in groundwater around the slopes of Mount Meru, Arusha, Tanzania. *Environmental Monitoring Assessment*, 188: 395.

Giao, N. T. 2020b. Spatial variations of surface water quality in Hau Giang province, Vietnam using multivariate statistical techniques. *Environment and Natural Resources Journal*, 18(4): 400-410.

Giao, N. T. 2020a. Evaluating current water quality monitoring system on Hau River, Mekong delta, Vietnam using multivariate statistical techniques. *Applied Environmental Research*, 42(1): 14-25.

Giao, N. T., P. K. Anh, and H. T. H. Nhien. 2021b. Spatiotemporal analysis of surface water quality in Dong Thap province, Vietnam using water quality index and statistical approaches. *Water*, 13(3): 336.

Giao, N. T., P. T. Hang, D. V. Ni, L. T. D. Mi, and H. B. Loc. 2021a. Evaluation of surface water quality in Ke Sach district, Soc Trang province using multivariate analysis. *Journal of Natural Resources and Environment Science*, 35(2021): 68-79.

Hoang, H. G., C. Lin, H. T. Tran, C. F. Chiang, X. T. Bui, N. K. Cheruiyot, C. C. Shern, and C. W. Lee. 2020. Heavy metal contamination trends in surface water and sediments of a river in a highly-industrialized region. *Environmental Technology & Innovation*, 20: 101043.

Howladar, M. F., M. A. Al Numanbakth and M. O. Faruque. 2018. An application of Water Quality Index (WQI) and multivariate statistics to evaluate the water quality around Maddhapara Granite Mining Industrial Area, Dinajpur, Bangladesh. *Environmental Systems Research*, 6: 13.

Huang, F., Xiaoquan Wang, X., Liping Lou, L., Zhiqing Zhou, Z, and J. Wu. 2010. Spatial variation and source apportionment of water pollution in Qiantang River (China) using statistical techniques. *Water Research*, 44(5): 1562-1572.

Ioele, G., M. D. Luca, F. Grande, G. Durante, R. Trozzo, C. Crupi, and G. Ragno. 2020. Assessment of Surface water quality using multivariate analysis: case study of the Crati River, Italy. *Water*, 12(8): 2214.

Kale, V. S. 2016. Consequence of temperature, pH, turbidity and dissolved oxygen water quality parameters. *International Advanced Research Journal in Science, Engineering and Technology*, 3(8): 186-190.

Kale, A., Bandela, N., Kulkarni, J, and K. Raut. 2020. Factor analysis and spatial distribution of water quality parameters of Aurangabad District, India. *Groundwater for Sustainable Development*, 10: 100345.

Kannel, P. R., S. Lee, and Y. S. Lee. 2008. Assessment of spatial–temporal patterns of surface and ground water qualities and factors influencing management strategy of groundwater system in an urban river corridor of Nepal. *Journal of Environmental Management*, 86(4): 595–604.

Lim, K. Y., N. A. Zakaria, and K. Y. Foo. 2020. Water quality assessment of urban catchment after the large-scale flood event: The worst natural tragedy at Pahang River, Malaysia. *Desalination and Water Treatment*, 175: 32-42.

Liu, C. W., K. H. Lin, and Y. M. Kuo. 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.

Liu, J., D. Zhang, Q. Tang, H. Xu, S. Huang, D. Shang, and R. Liu. 2021. Water quality assessment and source identification of the Shuangji River (China) using multivariate statistical methods. *PloS ONE*, 16(1): e0245525.

Liu, R., L. Gou, C. Men, Q. Wang, Y. Miao, and Z. Shen. 2019. Spatial-temporal variation of heavy metal's sources in the surface sediments of the Yangtze River Estuary. *Marine Pollution Bulletin*, 138: 526-533.

Lomoljo, R. M., A. Ismail, and C. K. Yap. 2009. Nitrate, ammonia and phosphate concentrations in the surface water of Kuala Gula Bird Sanctuary, west coast of Peninsular Malaysia. *Pertanika Journal of Tropical Agricultural Science*, 32(1): 1-5.

Ma, X., L. Wang, H. Yang, N. Li, and C. Gong. 2020. Spatiotemporal analysis of water quality using multivariate statistical techniques and the water quality identification index for the Qinhuai River Basin, East China. *Water*, 12(10): 2764.

Minh, H. V. T., R. Avtar, P. Kumar, K. N. Le, M. Kurasaki, and T. V. Ty. 2020. Impact of rice intensification and urbanization on surface water quality in An Giang using a statistical approach. *Water*, 12(6): 1710.

MoNRE (Ministry of Environment and Natural Resources). 2015. QCVN 08-MT: 2015/BTNMT National technical regulation on surface water quality. Hanoi, Vietnam: MoNRE.

Muangthong, S., and S. Shrestha. 2015. Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. *Environmental Monitoring Assessment*, 187: 548.

Niu, S., Y. Xia, C. Yang, and C. Liu. 2023. Impacts of the steel industry on sediment pollution by heavy metals in urban water system. *Environmental Pollution*, 335: 122364.

Niyoyitungiye, L., A. Giri, and M. Ndayisenga. 2020. Assessment of coliforms bacteria contamination in Lake Tanganyika as bioindicators of recreational and drinking water quality. *South Asian Journal of Research in Microbiology*, 6(3): 9-16.

Njimou, J. R., D. B. Tambe, A. S. Nana, A. S. Maicaneanu, A. A. Noujep, A. Tallab, and N. G. Elambo. 2021. Sequential treatment of municipal wastewater and valorization of filtering media. *Chemical Engineering Transactions*, 86: 661-666.

Pinto, C. C., G. M. Calazans, and S. C. Oliveira. 2019. Assessment of spatial variations in the surface water quality of the Velhas River Basin, Brazil, using multivariate statistical analysis and nonparametric statistic. *Environmental Monitoring and Assessment*, 191: 164.

Phu, H., N. L. N. Thao, and H. T. N. Han. 2021. Assessment of water quality in Long Xuyen quadrangle and measures for the protection of local water resources. *Vietnam Journal of Hydrometeorology*, 723: 13-22.

Pratama, M. A., Y. D. Immanuel, and D. R. Marthanty. 2020. A multivariate and spatiotemporal analysis of water quality in Code River, Indonesia. *The Scientific World Journal*, 2020: 8897029.

Qin, Y., and Y. Tao. 2022. Pollution status of heavy metals and metalloids in Chinese lakes: Distribution, bioaccumulation and risk assessment. *Ecotoxicology and Environmental Safety*, 248: 114293.

Raji, M. I. O., Y. K. E. Ibrahim, B. A. Tytler, and J. O. Ehinmidu. 2016. Assessment and seasonal variations of heavy metals and mineral elements in river Sokoto, North-western Nigeria. *Nigerian Journal of Basic and Applied Sciences*, 24(2): 9-14.

Rodrigues, M., A. Cravo, P. Freire, A. Rosa, and D. Santos. 2020. Temporal assessment of the water quality along an urban estuary (Tagus estuary, Portugal). *Marine Chemistry*, 223: 103824.

Su, S., D. Li, Q. Zhang, R. Xiao, F. Huang, and J. Wu. 2010. Temporal trend and source apportionment of water pollution in different functional zones of Qiantang River, China. *Water Research*, 45(4): 1781-1795.

Suteja, Y., and A. I. S. Purwiyanto. 2018. Nitrate and phosphate from rivers as mitigation of eutrophication in Benoa bay, Bali-Indonesia. *IOP Conference Series: Earth and Environmental Science*, 162: 012021.

Truc, D. T., P. H. Phat, N. D. G. Nam, P. V. Toan, and V. P. D. Tri. 2019. The water surface quality of Tien River in the area of Tan Chau district, An Giang province. *Journal of Science Can Tho University*, 55: 53-60.

Tuan, D. D. A., B. A. Thu, and N. H. Trung. 2019. Assessment of the current state of surface water for water exploitation and supply for Soc Trang city. *Science Journal of Can Tho University*, 55: 61-70.

VEA(Vietnam Environment Administration). 2019. Decision No. 1460/QD-TCMT, dated 12/11/2019, on promulgating Technical Guidelines for calculation and publication of Vietnam Water Quality Index (VN_WQI). Hanoi, Vietnam: Vietnam Environment Administration.

VNA (Vietnam National Assembly). 2020. Law on Environmental Protection No. 77/2020/QH14 (dated November 17, 2020, effective from January 1, 2022). Hanoi, Vietnam: Vietnam National Assembly.

Woldeab, B., A. Beyene, A. Ambelu, I. Buffam, and S. T. Mereta. 2018. Seasonal and spatial variation of reservoir water quality in the southwest of Ethiopia. *Environmental Monitoring Assessment*, 190: 163.

Zheng, B., C. Cao, Y. Qin, and M. Huang. 2008. Analysis of nitrogen distribution characters and their sources of the major input rivers of Three Gorges Reservoir. *Environmental Science*, 29(1): 1-6.

Abbreviations

PCA	Principal Component Analysis
CA	Cluster Analysis
WQI	Water Quality Index
DoNRE	Department of Natural Resources and Environmental
VEA	Vietnam National Assembly
MoNRE	Ministry of Natural Resource and Environmental