JEE Journal of Ecological Engineering

Journal of Ecological Engineering, 2025, 26(11), 100–113
https://doi.org/10.12911/22998993/207690
ISSN 2299–8993, License CC-BY 4.0

Received: 2025.06.08
Accepted: 2025.08.01
Published: 2025.08.15

Pollution-based clustering of surface water quality: Case study in Hau Giang province, Vietnam

Mong Tuyen Thi Nguyen[1,2], Giao Thanh Nguyen[1], Thu Minh Huynh Vuong[1], Ngan Nguyen Vo Chau[1*]

[1] College of Environment and Natural Resources, Can Tho University, Can Tho City, Vietnam
[2] Hau Giang Center for Natural Resources and Environment Monitoring, Hau Giang Province, Vietnam
* Corresponding author's e-mail: nvcngan@ctu.edu.vn

ABSTRACT

The study was carried out to assess the surface water quality in Hau Giang province by applying the organic pollution index (OPI), water quality index (WQI), and comprehensive pollution index (CPI). Water quality data from 48 monitoring locations were obtained from the Hau Giang's Department of Natural Resources and Environment in 2023. CA was employed to classify water quality in the study area based on the OPI, WQI, and CPI. Pearson correlation analysis was applied to examine the relationships among these water quality indices. Additionally, principal component analysis (PCA) was conducted to identify potential sources influencing surface water quality in the region. The three indices classified water quality into three levels and generally characterized water quality at a medium pollution level. Water quality indices exhibited spatial and temporal variability, with water quality generally deteriorating during the rainy season compared to the dry season. The results of OPI and WQI showed that water quality at locations S16-S20, S35, S37, S38 was in good level. All indices of OPI, WQI and CPI showed good water quality at the locations S35 and S38. Meanwhile, the OPI and CPI indices were similar in assessing the heavily polluted locations including S31, S42 and S47. WQI values only reflected poor water quality at locations S24 and S42. The WQI had a strong negative correlation with the CPI and OPI. Meanwhile, the CPI index had a positive and strong correlation with OPI. The results found the differences in using various indices to reflect water quality in the same study area. PCA revealed that the water quality was affected by two main sources (PC1, PC2) and 4 secondary sources (PC3-PC6). Primary sources explained 60% of the variation while secondary sources explained only 26.6% of the variation. The results of this study showed the necessity of using a combination of WQI, CPI and OPI indices while assessing surface water quality.

Keywords: correlation, multivariate statistical analysis, pollution index, surface water quality.

INTRODUCTION

The assessment of water quality plays a critical role in identifying and preventing water pollution, as well as providing essential information to support the management and sustainable use of water resources (Syeed et al., 2023; Murei et al., 2024). In almost all countries around the world, water quality monitoring and assessment is regulated by law (Syeed et al., 2023; Naderian et al., 2024). Water quality monitoring programs exist at different levels such as international, regional, national, provincial/city levels (VNA, 2020), and at emission sources monitoring indicators based on physical, chemical, and biological factors (MONRE, 2023; Syeed et al., 2023; Murei et al., 2024). The selection of assessment indicators depends on the characteristics of emission sources and resources allocated to water resource management (Murei et al., 2024). Different emission sources have different input characteristics and will therefore be expressed in different water quality assessment indicators (MONRE, 2023; Syeed et al., 2023; Murei et al., 2024). There are many methods used to evaluate water quality including comparing the value of each individual parameter with the allowable limit value of that parameter (MONRE, 2023; VEA, 2019; Syeed et al., 2023). In this way, the pollution characteristics of the water source will be determined, for example, the water source is polluted by organic substances, nutrients, microorganisms, heavy metals, pesticides, antibiotics or surfactants (Syeed et al., 2023). In addition, the pollution index method is also quite commonly used. The OPI, WQI, and CPI have been widely employed in previous studies to assess overall water quality (Son et al., 2020; Giao et al., 2022; Giao and Nhien, 2023; Syeed et al., 2023; Wieczorek et al., 2024). In addition, multivariate statistical methods, such as CA and PCA, are also commonly applied to classify sampling sites and identify potential sources influencing water quality (Chounlamany et al., 2017; Islam et al., 2022; Zhou et al., 2007; Syeed et al., 2023).

In Vietnam, the assessment of surface water quality (SWQ) monitoring programs primarily relies on the national technical standards for the WQI (VNA, 2020). Some scientists have also used OPI, CPI, CA and PCA methods in assessing SWQ (Son et al., 2020; Giao et al., 2022; Giao and Nhien, 2023). However, each index has its own calculation method and provides different information about SWQ. The combined use of these indices for SWQ assessment can provide comprehensive and valuable information to effectively supporting SWQ management. This study was conducted to simultaneously use of OPI, CPI, and WQI indices in reflecting various characteristics of SWQ. CA is employed to classify water quality by grouping sampling sites or observations based on the similarity of index values. PCA is utilized to identify potential sources influencing water quality in the study area. The current results provide scientific insights that support water quality assessment and contribute to the effective management of water resources.

MATERIALS AND METHODS

Description of water quality data

SWQ data used in this study were obtained from the Department of Natural Resources and Environment of Hau Giang province, Vietnam. The dataset comprises measurements from 48 monitoring locations (S1-S48) distributed across all major flowing water bodies in the province. At each site, water samples were collected six times per year (in March, April, June, July, September, and November). The analyzed water quality parameters included temperature, pH, dissolved oxygen (DO), total suspended solids (TSS), nitrite (NO_2^--N), nitrate (NO_3^--N), ammonium (NH_4^+-N), orthophosphate (PO_4^{3-}-P), biological oxygen demand (BOD), chemical oxygen demand (COD), iron (Fe), and Coliform. The procedures for sample collection, transportation, and analysis followed the guidelines outlined in the Standard Methods for the Examination of Water and Wastewater (APHA, AWWA, WEF, 2017) (Figure 1).

Calculation of water quality indices, correlation and principal component analysis

The CPI was calculated using Equation 1.

$$CPI = \frac{1}{n}\sum_{i=1}^{n} PI_i \quad (1)$$

$$PI_i = \frac{C_i}{S_i}$$

where: n – number of parameters; C_i – concentration of the monitored pollutants; S_i – corresponds to the permissible limit values of pollutants according to QCVN 08:2023/BTNMT, Class A.

CPI value classifies water quality into five categories including excellent (0.0–0.2), good (0.21–0.4), slightly polluted (0.41–1.0), moderately polluted (1.01–2.0) and heavily polluted (> 2.01). The OPI was calculated using Equation 2.

$$OPI = \frac{COD}{COD_S} + \frac{DIN}{DIN_S} + \frac{DIP}{DIP_S} - \frac{DO}{DO_S} \quad (2)$$

where: COD, DIN ($N\text{-}NH_4^+$, $N\text{-}NO_3^-$, $N\text{-}NO_2^-$), DIP ($P\text{-}PO_4^{3-}$), DO – the concentrations measured in the surface water samples of the study area.

COD_s, DIN_s, DIP_s, DO_s: the regulated SWQ values stated in QCVN 08:2023/BTNMT, Class A. OPI classifies SWQ into six categories including excellent (OPI < 0), good (0 ≤ OPI < 1), initially polluted (1 ≤ OPI < 2), slightly polluted (2 ≤ OPI < 3), moderately polluted (3 ≤ OPI < 4) and heavily polluted (OPI > 4).

The WQI was calculated using Equation 3. The resulting WQI values offer general information regarding the suitability of water for various uses at the monitoring sites.

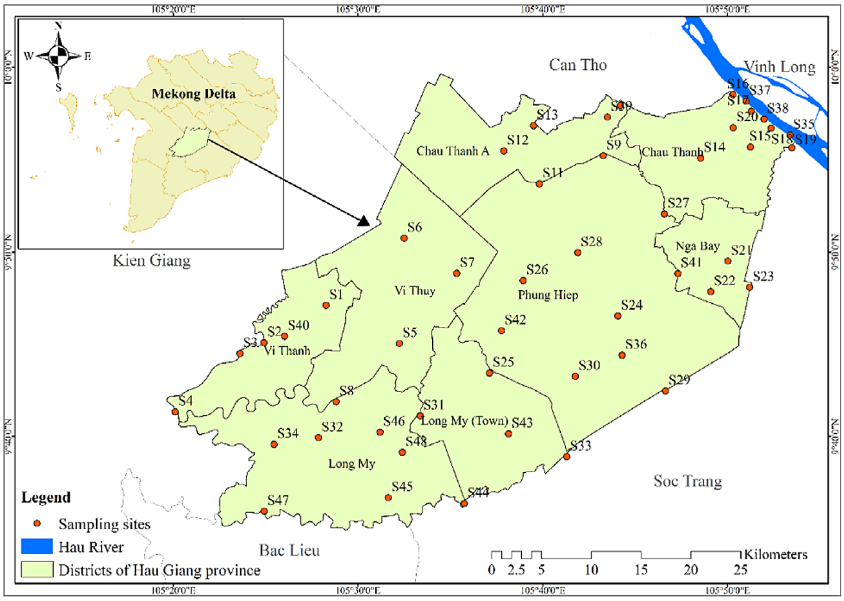

Figure 1. Map of the sampling locations

$$WQI = \frac{WQI_{pH}}{100} \times \left[\frac{1}{2} \times \sum_{a=1}^{7} WQI_a \times \sum_{b=1}^{2} WQI_b\right]^{\frac{1}{2}} WQI = \frac{WQI_{pH}}{100} \times \left[\frac{1}{2} \times \sum_{a=1}^{7} WQI_a \times \sum_{b=1}^{2} WQI_b\right]^{1/2} WQI = \frac{WQI_{pH}}{100} \left[\frac{1}{2} \sum_{a=1}^{2} WQI_a . WQI_b\right]^{1/2} \quad (3)$$

where: WQI_a – the calculated WQI value based on the parameters DO, BOD, COD, $N\text{-}NH_4^+$, $N\text{-}NO_2^-$, $N\text{-}NO_3^-$, and $P\text{-}PO_4^{3-}$, WQI_b – the WQI value calculated for Coliforms and *E. Coli*, WQI_{pH} – the calculated WQI value for pH, WQI value divides SWQ into six classes including excellent (91–100), good (76–90), moderate (51–75), poor (26–50), very poor (10–25) and heavily polluted (< 10) (VEA, 2019).

The Pearson correlation coefficient was calculated using Equation 4. Pearson correlation analysis serves as a preliminary descriptive method for estimating the strength and direction of associations among multiple variables in the study.

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) \times (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \times \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} r = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) . (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} . \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (4)$$

where: r – the correlation coefficient between parameters X and Y, n – the number of observations, X_i – the value of parameter X for the i^{th} observation, Y_i – the value of parameter Y for the i^{th} observation.

A coefficient approaching −1 or 1 signifies a strong inverse or direct correlation, respectively. Correlations are considered moderate when the absolute value of the coefficient lies between |0.3| and |0.5|; values greater than 0.5 indicate a strong correlation, while coefficients with an absolute value less than |0.3| represent a weak correlation (Prathumratana et al., 2008; Heale and Twycross, 2015).

PCA was employed to identify potential pollution sources and determine the key variables influencing SWQ in the study area. Principal components with eigenvalues greater than 1 were retained for interpretation (Chounlamany et al., 2017; Islam et al., 2022). The loading factors was used to identify key water quality parameters. The values of 0.3–0.5, > 0.5–0.75, > 0.75 was considered weak, medium, strong, respectively (Chounlamany et al., 2017; Islam et al., 2022). CA was applied to group SWQ indices based on Euclidean distance metrics (Chounlamany et al., 2017). In addition, CA was performed to group the sampling locations based on the similarity of water quality indices. This analysis makes no prior assumptions regarding the similarity among sites; instead, clusters are formed statistically when the linkage distance criterion $D_{link}/D_{max} \times 100$ is less than 60, where D_{link} is the linkage distance for a specific case and D_{max} is the maximum linkage distance. The number of clusters was determined based on the characteristics of the study. The Ward's method and Euclidean range were used as measures of similarity (Zhou et al., 2007). CA and PCA was performed using PRIMER V.5.2.9 tool. The difference of values of OPI, CPI and WQI among the sampling sites and sampling times was examined using SPSS analysis of variance.

RESULTS AND DISCUSSION

Organic pollution in water bodies

Figure 2 presented the variation of organic pollution index in water bodies of Hau Giang province. The OPI index fluctuated greatly depending on the sampling location. Organic pollution at the locations S31, S40, S42, and S47 was the highest and statistically significantly different ($p < 0.05$) compared to other locations (Figure 2). These locations are affected by polluted water from the Lung Ngoc Hoang nature reserve, the activities of the Vi Thanh industrial cluster, the activities of the household waste landfilling and the impact of intersection water between three provinces Kien Giang, Bac Lieu and Hau Giang. Previous studies have also shown that water from conservation areas, industrial production and landfilling activities generated large amounts of pollutants, including organic substances (Nhien and Giao, 2019; Hong and Giao, 2021; Giao, 2022a; Hong and Giao, 2022). The locations with the lowest organic pollution values include S35, S27, and S38 (Figure 2), which are just starting to be contaminated with organic matters. These locations are located on Hau river, which is affected by Song Hau 1 Thermal Power Plant and Song Hau Industrial Park. The large water flow and volume in the Hau river can reduce organic pollution in these two locations. Previous studies also showed that large rivers often have lower pollution levels due to pollutant dilution (Giao, 2020a; Wieczorek et al., 2024). The locations S16-S20 had slight organic pollution while the remaining locations had severe organic pollution (Figure 2). The previous research used the OPI index to evaluate SWQ in Can Tho city and the results showed that surface water in the study area was organically polluted (Giao et al., 2022). The results from former and current studies indicated that surface water in Mekong Delta is contaminated by organic matters.

Figure 3 illustrates that the levels of organic matter in surface water bodies of Hau Giang province exhibited temporal fluctuations. In addition, the results showed that the OPI values were always greater than 4, indicating a water environment with heavy organic pollution. This result is also similar to previous research results that the water bodies of Hau Giang province were polluted with organics (BOD_5 and COD), nutrients (NH_4^+-N, NO_2^--N, PO_4^{3-}-P) and microorganisms (Coliforms) (Giao, 2020a). The highest OPI value was observed in July, whereas the lowest values were recorded in March and April. Giao et al. (2022) investigated the variation of OPI values in surface water in Can Tho city and found that OPI values also exhibited spatial fluctuations. In their study, OPI was found to be highest in March and lowest in September,

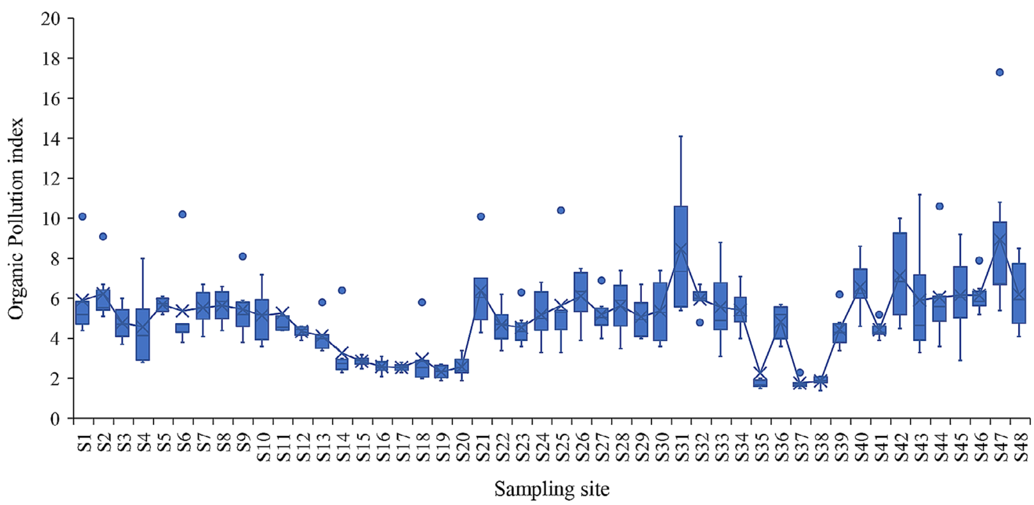

Figure 2. Spatial variation of OPI in the water bodies

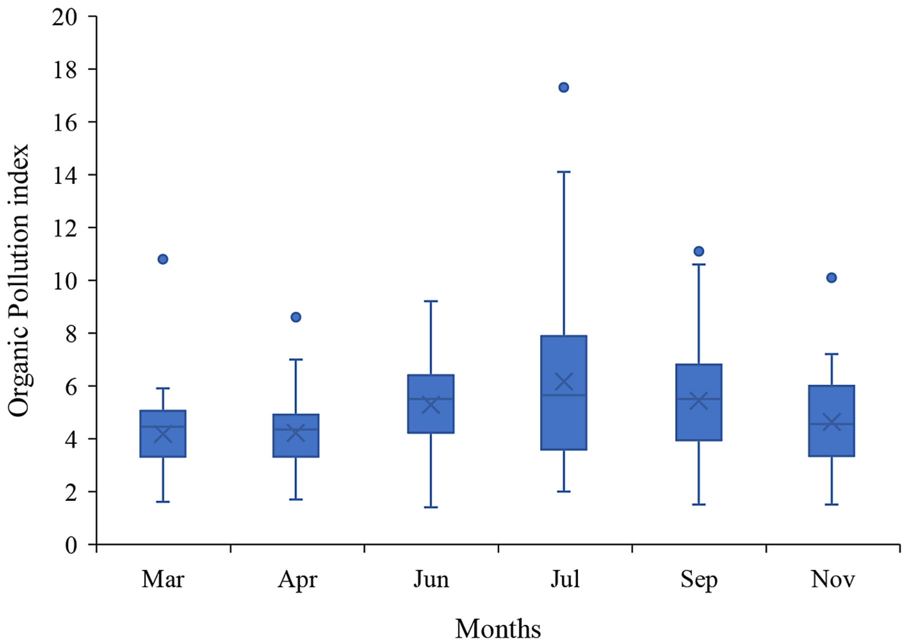

Figure 3. Temporal variation of OPI in the water bodies

with values showing an increasing trend again in December. These findings suggest that the level of organic matter in surface water bodies of Hau Giang province tends to be higher compared to those in Can Tho - a city nearby Hau Giang. The time of high organic matter in these two study areas is also very different. This shows that OPI fluctuates greatly over space and time. High levels of organic matter in surface water can limited water use purposes because the presence of organic matter causes a number of serious environmental and health problems when it is used as water for domestic use (Rapukdi et al., 2019; Benítez et al., 2021; Mahato and Gupta, 2022; Wu et al., 2022).

Spatial and temporal variation of overall water quality in water bodies

The variation of WQI index in surface water in water bodies of Hau Giang province is presented in Figure 4. Water quality according to WQI value ranges from 31 (bad) to 93 (very good), reaching an average value at 68 (average water quality). The former study by Giao et al. (2022) showed that water quality in Can Tho city water bodies ranged from average to very good. The studies on SWQ in the Mekong Delta show that water quality also ranged from very poor to very good (Giao, 2020b; Giao et al., 2022). The

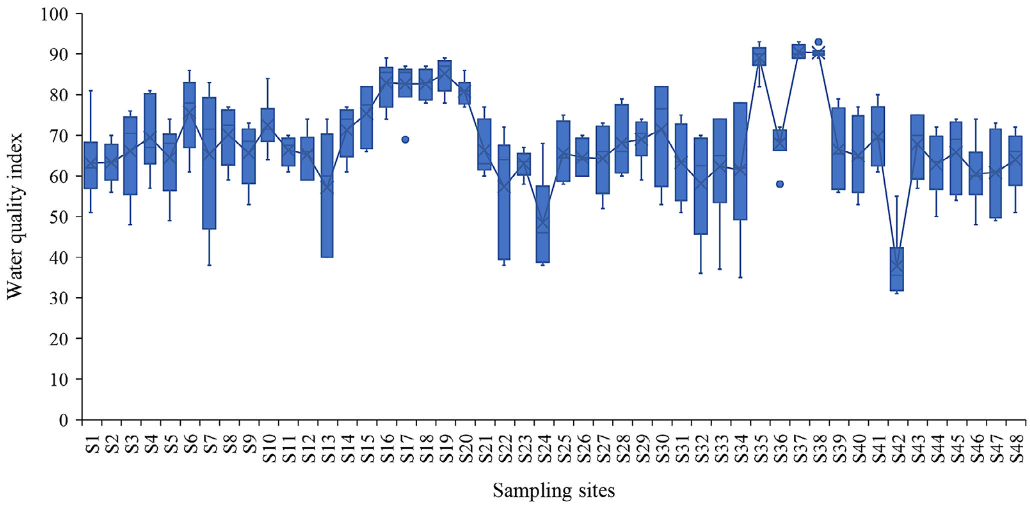

Figure 4. Spatial variation of WQI in the water bodies

main influencing indicators for low overall water quality are often due to the presence of organic matter, microorganisms and heavy metals (Giao et al., 2022; Wieczorek etal., 2024). Overall water quality fluctuated greatly over space, especially at locations S7, S13, S22, S24, S32-S34. These locations with large fluctuations in WQI values were all affected by the discharge sources. For example, surface water source at location S7 was affected by waste from the market area and wastewater from agricultural production areas. At location S13, surface water source was affected by waste sources from industrial production activities in the Nhon Nghia industrial cluster. Surface water at location S22 was affected by waste from activities of the Nga Bay industrial cluster, while surface water at location S24 was affected by waste from markets, residential areas and hospitals. Surface water from locations S32-S34 was affected by markets, residential areas, and rice growing areas. WQI values at locations such as S17-S20, S35, S37-S38 showed that good water quality was significantly higher ($p < 0.05$) than that in other locations. These locations showed signs of new beginnings of organic pollution or light organic pollution according to the assessment results of the OPI index in the previous discussion. The lowest WQI values (bad water quality) were found at locations S42 and S24 where they were affected by landfills and where they were affected by mixed sources such as markets, hospitals, and residential areas. Statistical results of WQI values showed that most locations had average water quality. Previous research by Giao and Nhien (2023) showed that WQI values also varied greatly across space. The overall water quality at some monitoring locations in Binh Duong, Binh Phuoc, An Giang, Tien Giang, Ben Tre and Dong Thap provinces was very good while the overall water quality at monitoring locations in Tay Ninh province was good. Former study by Giao (2022a) reported that water bodies in Ca Mau province had that overall water quality ranging from very bad to good, of which good water quality accounted for 1.9%, average accounted for 9.6%, bad water quality accounted for 48.1% and very good water quality accounted for 40.3%. Meanwhile, water quality in An Giang had WQI values ranging from 15 to 71, classified from very bad to moderate (Giao and Nhien, 2020).

The results revealed that the overall SWQ of water bodies in Hau Giang province exhibited significant seasonal variations ($p < 0.05$). The amplitude of WQI fluctuations was smaller during the dry season (March-June) compared to the rainy season (July-November). Although SWQ in both seasons was generally classified as average, WQI values during the dry season tended to be higher than those observed in the rainy season (Figure 5). A similar pattern of seasonal WQI variation was reported by Giao et al. (2022); however, in

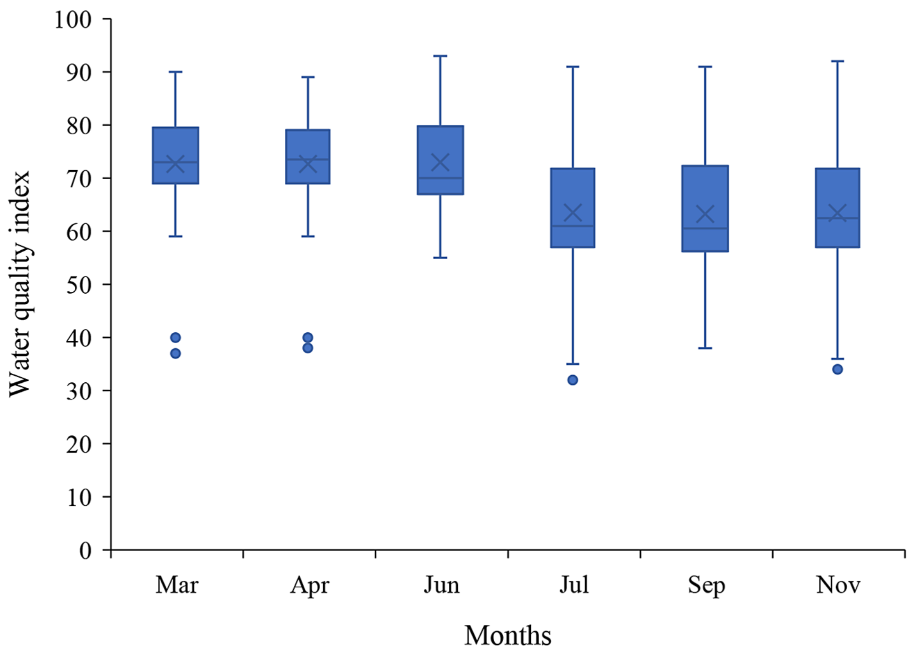

Figure 5. Temporal variation of WQI in the water bodies

contrast to Hau Giang, SWQ in Can Tho city was lower during the dry period (March and December) than in the rainy period (September).

The SWQ in Hau Giang province according to the CPI index ranked SWQ from mildly polluted to heavily polluted (Figure 6). Among that, the locations S17, S19, S35, S37 and S38 were lightly polluted and the locations S31, S42, S47 were heavily polluted. Lightly polluted locations were all located in large river areas while heavily polluted locations were affected by markets, residential areas, landfills and intersections between rivers. This result was also consistent with the organic pollution assessment results discussed in the previous section and was also consistent with previous research results (Nhien and Giao, 2019; Hong and Giao, 2021; Giao, 2022b; Hong and Giao, 2022; Wieczorek et al., 2024). The results also showed that water quality according to the CPI fluctuated greatly across space, meaning that it was affected by different socio-economic activities, which could lead to different levels of pollution (Giao et al., 2022). The level of pollution also depends on the characteristics of the water body, the effectiveness of pollution control by scientific and technical solutions and the enforcement of laws on environmental protection (Wieczorek et al., 2024).

CPI values in sampling months fluctuated greatly, especially in the rainy season of August and September (Figure 7). In all months, CPI values showed that water quality was classified as average pollution, in which CPI values in March,

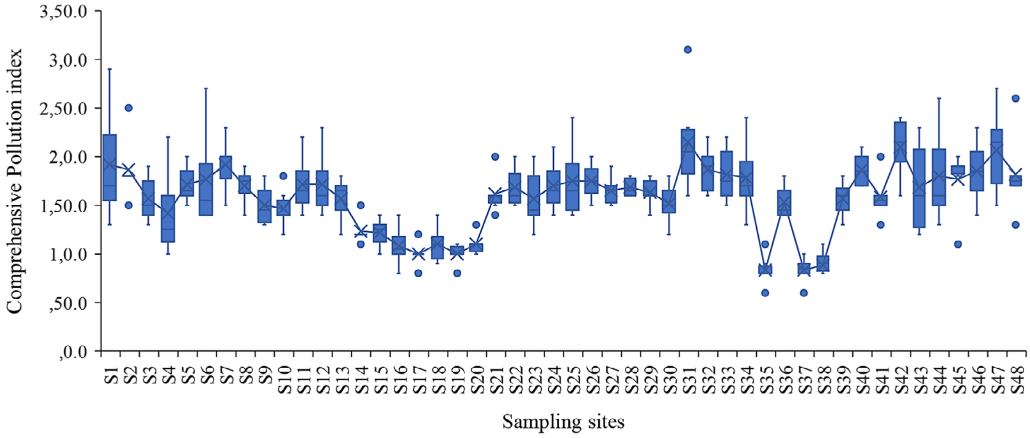

Figure 6. Spatial variation of CPI in the water bodies

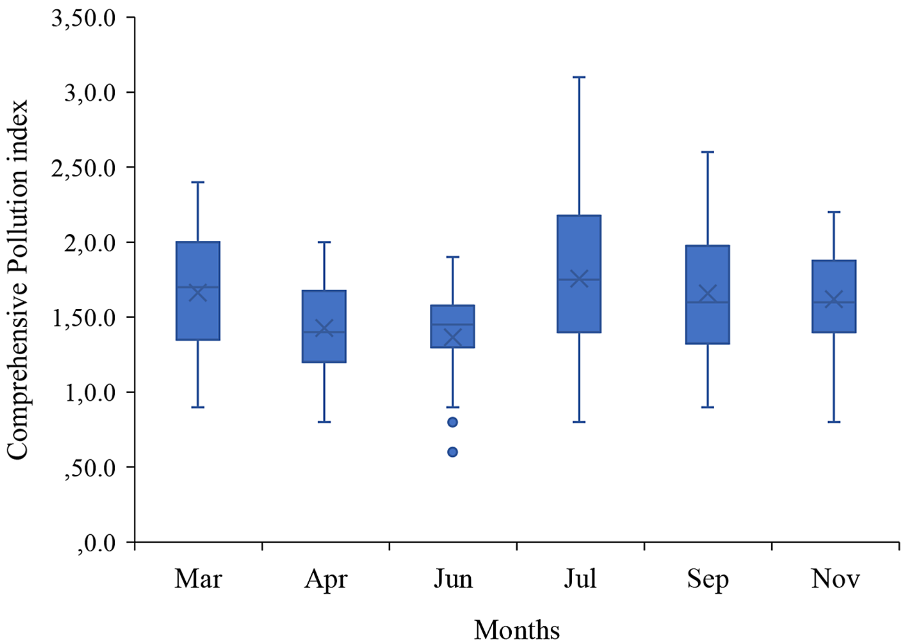

Figure 7. Temporal variation of CPI in the water bodies

July, September, and November were significantly higher ($p < 0.05$) than those in April and June. Previous research results showed that the CPI value of seasonal fluctuations (Giao et al., 2022). SWQ in Can Tho city had a CPI value of mild to moderate pollution; while CPI values in September showed moderate to heavy pollution levels (Giao et al., 2022). The variation of the values of CPI is often associated with the seasonal variation of individual SWQ parameters (Chounlamany et al., 2017; Islam et al., 2022).

Correlation of SWQ indices

The results of analyzing the relationship between water quality assessment indices in the study area are presented in Table 1. The WQI had a strong negative correlation with CPI and OPI index. This could mean that the more the CPI and OPI indexes increase, the more the WQI index decreases. The negative correlation between WQI and CPI showed that these two indices were not completely compatible with each other when classifying water quality. The research results also showed that the CPI index had a positive and very strong correlation with the OPI index (Table 1).

Clustering SWQ using water pollution indices

Water quality grouping according to organic pollution index is presented in Figure 8. The results of water quality grouping into six groups, with 3 levels of water quality: starting organic pollution (Group I), light organic pollution (Group II) and heavy organic pollution (Group III-VI). Group I included 2 locations S37 and S38 on Hau river, which were affected by Song

Table 1. Correlation between water quality indices

Prameter		WQI	CPI	OPI
WQI	r	1	-0.686**	-0.536**
	Sig.	-	0.000	0.000
	N	288	288	288
CPI	r	-0.686**	1	0.805**
	Sig.	0.000	-	0.000
	N	288	288	288
OPI	r	-0.536**	0.805**	1
	Sig.	0.000	0.000	-
	N	288	288	288

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

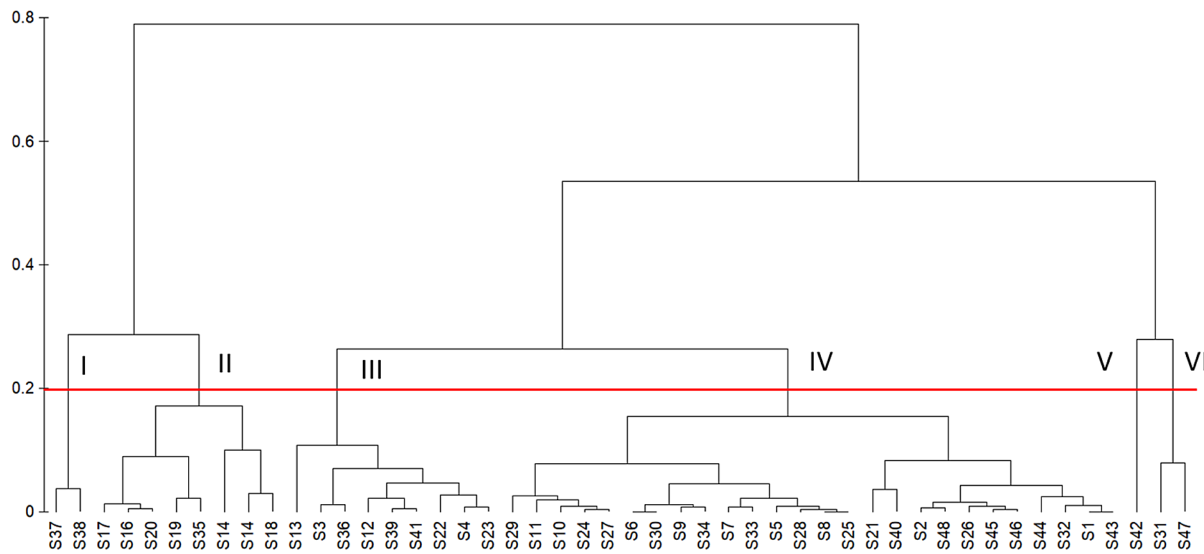

Figure 8. Clustering SWQ using OPI

Hau 1 Thermal Power Plant and Song Hau Industrial Park. Group I had just begun to be contaminated with organic matter (Table 2). Group II included 8 locations, accounting for 16.7% of the total number of monitoring locations classified as light organic pollution. Notably, 38 locations (accounting for 79.1%) were classified as severely organically polluted, including groups III–VI. The results of this study showed that the SWQ in the area was organically polluted. Taking water from these areas for use as domestic water supply poses many potential risks (Benítez et al., 2021; Mahato and Gupta, 2022; Wu et al., 2022). Therefore, solutions to improve SWQ in the study area need to be implemented soon. Domestic water supply in this area needs to be monitored for by-products of the disinfection process such as trihalogenomethanes (Rapukdi et al., 2019).

The study grouped water quality according to WQI and the results are presented in Figure 9. The research results showed that water quality was classified into five groups including groups I–V. Water quality falls into three categories: bad (Group I–II), average (Group III–IV) and good (Group V) (Table 3). Group I with bad quality included two positions S24 and S42, accounting for 4.2% of the total positions. In the study area, average water quality (Group III and Group IV) accounted for a high proportion with 38 locations accounting for 79.2%. Group V was classified as having good water quality, comprising 8 locations and accounting for 16.7% of the total monitoring sites. Overall, the results indicate that surface water quality in the study area is generally characterized as average. This level of water quality may lead to increased costs for domestic water treatment and pose potential health risks associated with poor water quality (Rapukdi et al., 2019; Benítez et al., 2021; Mahato and Gupta, 2022; Wu et al., 2022).

The results of water quality grouping using CPI values are presented in Figure 10. CA classifies water quality according to CPI into three

Table 2. Water quality in the identified clusters using OPI

Cluster	Sites	Number of sites	%	OPI	Water quality classification
I	S37, S38	2	4.2	1.8	Initially polluted
II	S14, S15, S16, S17, S18, S19, S20, S35	8	16.7	2.5	Slightly polluted
III	S3, S4, S12, S13, S22, S23, S36, S39, S41	9	18.8	4.5	Heavily polluted
IV	S1, S2, S5, S6, S7, S8, S9, S10, S11, S12, S21, S24, S26, S27, S28, S29, S30, S32, S33, S34, S40, S43, S44, S45, S46, S48	26	54.2	5.8	Heavily polluted
V	S42	1	2.1	7.2	Heavily polluted
VI	S31, S47	2	4.2	8.7	Heavily polluted

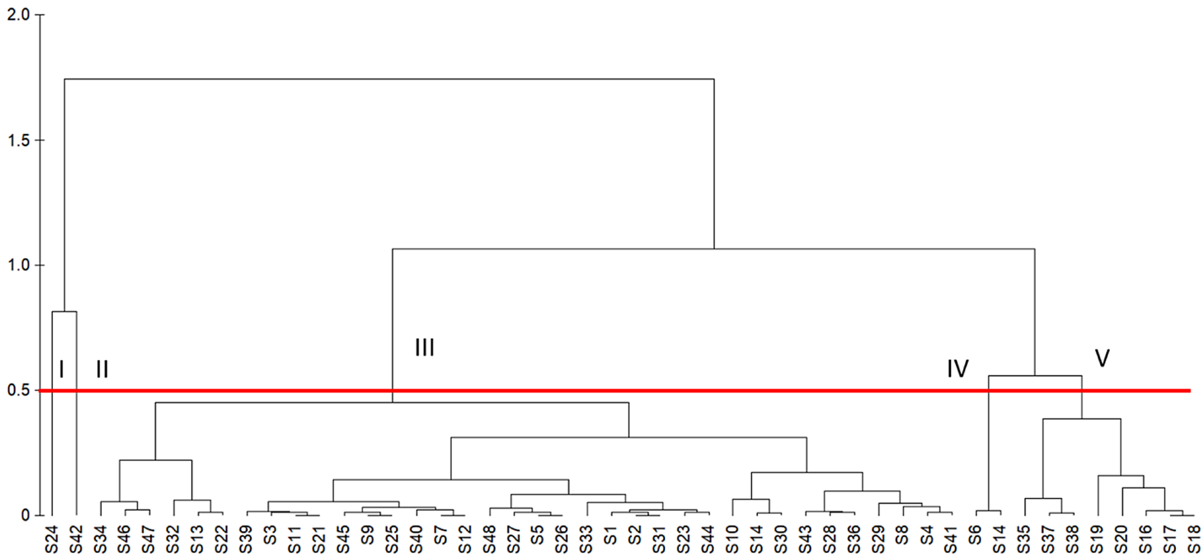

Figure 9. Clustering SWQ using WQI

Table 3. Water quality in the identified clusters using WQI

Cluster	Sites	Number of sites	%	WQI	Water quality classification
I	S24	1	2.1	48.5	Poor
II	S42	1	2.1	38.0	Poor
III	S1, S2, S3, S4, S5, S7, S8, S9, S10, S11, S12, S13, S15, S21, S22, S23, S25, S26, S27, S28, S29, S30, S31, S32, S33, S34, S36, S39, S40, S41, S43, S44, S45, S46, S47, S48	36	75.0	65.5	Moderate
IV	S6, S14	2	4.2	73.5	Moderate
V	S16, S17, S18, S19, S20, S35, S37, S38	8	16.7	88.2	Good

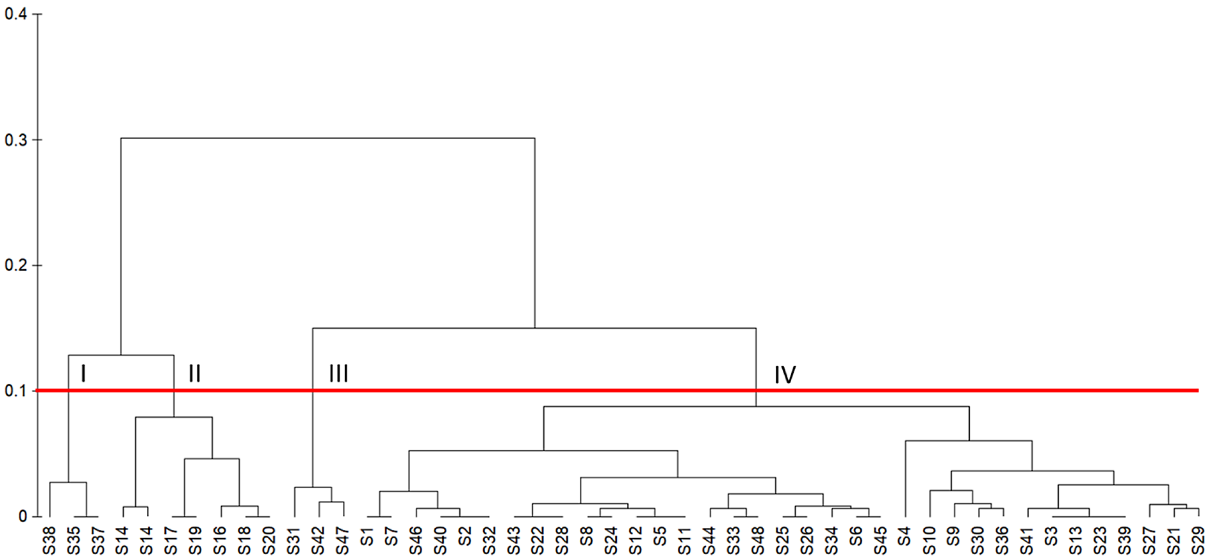

Figure 10. Clustering SWQ using CPI

categories including good (Group I), average (Group II, IV) and heavily polluted (Group III). Water quality was classified as good with two positions (S35 and S38) accounting for 4.2%. The water quality group classified as average had up to 42 locations, accounting for 89.6% (Table 4). The water quality group classified as severely polluted had only three locations, accounting for 6.3%. Thus, the water quality characteristic in the study area according to CPI was average pollution.

The results presented in Table 5 showing that using a combination of SWQ grouping indices could provide more information for management. All three indices of OPI, WQI and CPI divided quality into three groups. The OPI index provides information about organically polluted water sources, serving the management of domestic water supply and calculating WQI values. In Vietnam, if the water is organically polluted, calculating the WQI needs to use a weighted formula (VEA, 2019). The WQI and CPI indices

Table 4. Water quality in the identified clusters using CPI

Cluster	Sites	Number of sites	%	CPI	Water quality classification
I	S35, S38	2	4.2	0.9	Good
II	S14, S15, S16, S17, S18, S19, S20	7	14.6	1.1	Moderately polluted
III	S31, S42, S47	3	6.3	2.1	Heavily polluted
IV	S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S21, S22, S23, S24, S25, S26, S27, S28, S29, S30, S32, S33, S34, S36, S37, S39, S40, S41, S43, S44, S45, S46, S48	36	75.0	1.7	Moderately polluted

Table 5. Comparing classification of water quality by clustering OPI, WQI and CPI

Parameter	Water quality classification	Number of sites	Proportion (%)
OPI	Initially polluted	2	4.2
	Slightly polluted	8	16.7
	Heavily polluted	38	79.1
WQI	Good	8	16.7
	Moderate	38	79.1
	Poor	2	4.2
CPI	Good	2	4.2
	Moderately polluted	42	89.5
	Heavily polluted	3	6.3

classified water quality differently; however, both reflected the general characteristics of average water quality in the study area. This finding is consistent with previous studies that assessed SWQ in the Mekong Delta region (Giao et al., 2021; Giao et al., 2022). Water quality grouping according to CPI was more stringent because the number of moderate to severely polluted locations was greater than that of SWQ grouping using the WQI index. The reason may be because in the WQI calculation according to the regulations of the Vietnam Environment Administration, the TSS index has been removed while the CPI calculation still uses the TSS index (VEA, 2019). The research results also showed that if the water is heavily polluted with organic matters, the CPI and WQI calculation results would give an average water quality value. Therefore, in water quality assessment, it is necessary to consider choosing to use a combination of WQI, CPI along with OPI index. This was also consistent with the results of previous research by Giao et al. (2022) using OPI, WQI and CPI indices in assessing SWQ in Can Tho city.

Potential sources influencing SWQ in the study areas

The PCA results showed that SWQ in the study area was affected by two main sources (PC1, PC2) and four secondary sources (PC3-PC6). Primary sources explained 60% of the variation in SWQ while secondary sources explained only 26.6% of the variation. Temperature was affected by three factors PC2, PC3, PC4. The pH was influenced by PC2 and PC6. DO was

Table 6. Sources influencing SWQ in the study areas

Var.	PC1	PC2	PC3	PC4	PC5	PC6
T	-0.005	0.618	-0.357	0.510	-0.269	-0.173
pH	0.243	-0.446	-0.228	-0.251	-0.146	-0.489
DO	0.137	-0.346	0.370	0.739	0.293	-0.062
TSS	-0.321	0.021	-0.083	0.000	0.528	-0.224
NO_2-N	-0.265	-0.098	-0.507	0.008	0.283	0.439
NO_3-N	-0.250	-0.394	-0.432	0.197	-0.166	-0.166
NH_4-N	-0.305	0.163	0.355	-0.249	0.059	0.104
PO_3-P	-0.349	0.111	0.080	-0.064	-0.170	-0.445
BOD	-0.358	-0.151	0.214	0.130	-0.322	0.026
COD	-0.355	-0.104	0.223	0.054	-0.329	0.001
Coliform	-0.321	-0.227	-0.076	0.079	-0.180	0.338
Fe	-0.338	0.101	-0.003	-0.021	0.396	-0.365
Eigenvalues	5.79	1.40	0.96	0.88	0.86	0.53
%Variation	48.3	11.7	8.0	7.3	7.1	4.4
%Cum.Var.	48.3	60.0	67.9	75.3	82.4	86.6

similar to the temperature that it was impacted by PC2-PC4. DO in water is often controlled by river disturbance, the presence of organic matters and aquatic plants (MRC, 2015). TSS was affected by PC1 and PC5. Factors affecting TSS in canals are usually rainwater runoff and riverbank erosion (MRC, 2015). Nitrite was only affected by PC3 while nitrate and ammonium were influenced by PC2, PC3 and PC1, PC3 respectively. Sources that influence nutrients include the use of fertilizers, detergents, and the decomposition of pollutants (Barakat et al., 2016; Zeinalzadeh and Rezaei, 2017). Orthophosphate was influenced by PC1 and PC6. BOD and COD were both affected by PC1 and PC5. Sources affecting organic matters include domestic wastewater, industrial and agricultural wastewater (Chounlamany et al., 2017; Islam et al., 2022; Zeinalzadeh and Rezaei, 2017; Siwiec et al., 2018). Coliform was influenced by sources PC1 and PC6 while Fe was impacted by three sources including PC1, PC5 and PC6. Coliforms originate from human and animal waste (UNICEF, 2008; WHO, 2008) while temperature is influenced by natural factors. Thus, the research results showed that water quality indicators were affected by from one to three potential polluting sources. Further research needs to survey and calculate the contribution of sources to water quality problems to have effective management solutions that contribute to improving water quality. Previous research showed that there were 12 potential sources explaining the fluctuations in SWQ in Hau Giang province, including three main sources and nine secondary sources (Giao, 2020a). In the study conducted in Can Tho city, 17 potential sources were identified as accounting for 100% of the SWQ fluctuations (Giao et al., 2022). In contrast, the current study shows a reduction in the number of potential sources influencing SWQ, indicating a positive trend and suggesting improved effectiveness in controlling pollution sources in the study area.

CONCLUSIONS

The study employed the OPI, WQI, and CPI to assess SWQ in the water bodies of Hau Giang province. The results indicated that:

- the OPI classified organic pollution levels in SWQ into three categories of just starting, lightly polluted, and heavily polluted. The WQI index classified SWQ into three levels including good, average and bad. The CPI distinguished SWQ into three levels of bad, average and good. All three indices classified SWQ into three levels and characterized SWQ of the study area at an average level.
- the results of SWQ grouping between OPI and WQI showed that there were similar levels of organic pollution and good water quality (locations S16-S20, S35, S37, S38). All three indices of OPI, WQI and CPI showed good water quality at locations S35 and S38. Meanwhile, the OPI and CPI indices were similar in assessing heavily polluted locations including S31, S42 and S47.
- the WQI value only reflected poor water quality at S42 and S24. The results indicated that the use of different indices – OPI, WQI, and CPI – led to variations in the assessment of SWQ within the same study area. The values of these indices were significantly influenced by pollution sources, water body characteristics, and temporal factors.
- the OPI index provides information about organically polluted water resources, while the WQI and CPI indices classify the overall water quality of a water body.

The findings of this study highlight the necessity of employing a combination of WQI, CPI, and OPI indices to achieve a more comprehensive and accurate assessment of SWQ.

REFERENCES

1. APHA, AWWA, WEF. (2017). *Standard methods of for the examination of water and wastewater, 23ed*. American Public Health Association: Washington, DC, USA.
2. Barakat A., Baghdadi M.E., 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. *Int. Soil Water Conserv. Res., 4*(4), 284–292.
3. Benítez J.S., Rodríguez C.M., Casas A.F. (2021). Disinfection byproducts (DBPs) in drinking water supply systems: A systematic review. *Physics and Chemistry of the Earth, Parts A/B/C, 123*, 102987.
4. Chounlamany V., Tanchuling M.A., 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.
5. Giao N.T. (2020a). Spatial variations of surface

water quality in Hau Giang Province, Vietnam using multivariate statistical techniques. *Environment and Natural Resources, 18*(4), 400–410.

6. Giao N.T. (2020b). Surface water quality in tributaries adjacent to Hau River in Can Tho city. *Journal of Agriculture and Rural Development, 15*, 79–86 (In Vietnamese).
7. Giao N.T. (2022a). Analysis of surface water quality using multivariate statistical approach: A case study in Ca Mau peninsula, Vietnam. *Pollution, 8*(2), 463–477.
8. Giao N.T. (2022b). Characteristics of effluents from industrial, hospital, aquacultural and domestic wastewater treatment in Tien Giang province, Vietnam. *Journal of Science and Technology Research, 4*(2), 7–13.
9. Giao N.T., Nhien H.T.H. (2020). Phytoplankton - Water quality relationship in water bodies in the Mekong Delta, Viet Nam. *Appl. Environ. Res., 42*(2), 1–12.
10. Giao N.T., Nhien H.T.H. (2023). Evaluating surface water quality using indexes of water quality and plankton diversity. *Civil Engineering Journal, 9*(5), 1187–1202.
11. Giao N.T., Nhien H.T.H., Anh P.K., Ni D.V. (2021). Classification of water quality in low lying area in Vietnamese Mekong Delta using set pair analysis method and Vietnamese water quality index. *Environ. Monit. Assess., 193*(6), 1–16.
12. Giao N.T., Nhien H.T.H, Anh P.K., Thuptimdang P. (2022). The uses of combined water quality indexes and multivariate statistical techniques in evaluating surface water quality variability in Can Tho, Vietnam. *Environmental Monitoring and Assessment, 194*, 844.
13. Heale R., Twycross A. (2015). Validity and reliability in quantitative studies. *Evid. Based Nurs. 18*, 66–67.
14. Hong T.T.K., Giao N.T. (2021). Characteristics of "black water" in Lung Ngoc Hoang Natural Reserve, Phung Hiep District, Hau Giang Province, Vietnam. *Journal of Energy Technology and Environment, 3*(4), 1–8.
15. Hong T.T.K., Giao N.T. (2022). Effect of industrial activities on surface water quality in the An Giang Province, Vietnam. *Journal of Energy Technology and Environment, 4*(3), 19–29.
16. Islam M.S., Nakagawa K., Abdullah-Al-mamun M., Khan A.S., Goni M.A., Berndtsson R. (2022). Spatial distribution and source identification of water quality parameters of an industrial seaport riverbank area in Bangladesh. *Water, 14*(9), 1356.
17. Mahato J.K., Gupta S.K. (2022). Advanced oxidation of Trihalomethane (THMs) precursors and seasonwise multi-pathway human carcinogenic risk assessment in Indian drinking water supplies. *Process Safety and Environmental Protection, 159*, 996–1007.
18. MONRE. (2023). QCVN 08:2023/BTNMT *National technical regulation on surface water quality*. Vietnamese Ministry of Environment and Natural Resources
19. MRC. (2015). *Lower Mekong regional water quality monitoring report*. MRC Technical Paper No. 51. Mekong River Commission.
20. Murei A., Kamika I., Momba M.N.B. (2024). Selection of a diagnostic tool for microbial water quality monitoring and management of faecal contamination of water sources in rural communities. *Science of The Total Environment, 906*, 167484.
21. Naderian D., Noori R., Heggy E., Bateni S.M., Bhattarai R., Nohegar A., Sharma S. (2024). A water quality database for global lakes. *Resources, Conservation and Recycling, 202*, 107401.
22. Nhien H.T.H, Giao N.T. (2019). Environmental quality of soil, water and sediment at Dong Thang Landfill, Can Tho City, Viet Nam. *Applied Environmental Research, 41*(2), 73–83.
23. Prathumratana L., Sthiannopkao S., Kim K.W. (2008). The relationship of climatic and hydrological parameters to surface water quality in the lower Mekong River. *Environ. Int., 34*, 860–866.
24. Ratpukdi T., Sinorak S., Kiattisaksiri P., Punyapalakul P., Siripattanakul-Ratpukdi S. (2019). Occurrence of trihalomethanes and haloacetonitriles in water distribution networks of Khon Kaen Municipality, Thailand. Water Supply, *19*(6), 1748–1757.
25. Siwiec T., Reczek L., Michel M.M., Gut B., Hawer-Strojek P., Czajkowska J., Jóźwiakowski K., Gajewska M., Bugajski P. (2018). Correlations between organic pollution indicators in municipal wastewater. *Archives of Environmental Protection, 44*(4), 50–57.
26. Son C.T., Giang N.T.H., Thao T.P., Nui N.H., Lam N.T., Cong V.H. (2020). Assessment of Cau River water quality assessment using a combination of water quality and pollution indices. *Journal of Water Supply: Research and Technology, 69*, 160–172.
27. Syeed M., Hossain S., Karim R., Uddin M.F., Hasan M., Khan R.H. (2023). Surface water quality profiling using the water quality index, pollution index and statistical methods: A critical review. *Environmental and Sustainability Indicators, 18*, 100247.
28. UNICEF. (2008). *UNICEF Handbook on water quality.* New York: United Nations Children's Fund.
29. VEA - Vietnam Environment Administration. (2019). Decision 1460/QD-TCMT dated November 12 2019 on Issuing the manual for calculating water quality index. [Online]. Available from: https://thuvienphapluat.vn/van-ban/tai-nguyen-moi-truong/Quyet-dinh-1460-QD-TCMT-2019-ky-thuat-tinh-toan-va-cong-bo-chi-so-chat-luong-nuoc-428277.aspx [Accessed 23 January 2023] (In Vietnamese).
30. VNA - Vietnamese National Assembly. (2020). Law No. 72/2020/QH14 dated November 17 2020

on Environmental Protection. [Online]. Available from: https://thuvienphapluat.vn/van-ban/EN/Tai-nguyen-Moi-truong/Law-72-2020-QH14-on-Environmental-Protection/463512/tieng-anh.aspx [Accessed 2 May 2023] (In Vietnamese).

31. WHO. (2008). *Guidelines for drinking water quality – 1: Recommendations, Third edition, incorporating first and second addenda.* Geneva: World Health Organization.

32. Wieczorek K., Turek A., Szczesio M., Wolf W.M. (2024). A holistic approach to the spatio-temporal variability investigation of the main river water quality – The importance of tributaries. *Science of The Total Environment, 906*, 167588.

33. Wu L., Qiu X.W, Wang T., Tao K., Bao L.J., Zeng E.Y. (2022). Water quality and organic pollution with health risk assessment in China: A short review. *ACS EST Water, 2*, 8, 1279–1288.

34. Zeinalzadeh K., Rezaei E. (2017). Determining spatial and temporal changes of surface water quality using principal component analysis. *J. Hydrol. Reg. Stud., 13*, 1–10.

35. Zhou F., Liu Y., Guo H. (2007). Application of multivariate statistical methods to water quality assessment of the water courses in North Western new territories Hong Kong. *Environ. Monit. Assess., 132*, 1–13.