International Journal of Environmental Science and Technology (2023) 20:2487–2502
https://doi.org/10.1007/s13762-023-04756-5

ORIGINAL PAPER

Water quality assessment of the Nam River, Korea, using multivariate statistical analysis and WQI

H. G. Kwon[1] · C. D. Jo[1]

Received: 27 January 2022 / Revised: 17 July 2022 / Accepted: 2 January 2023 / Published online: 7 January 2023

Abstract

Water quality assessment using water quality index (WQI) is performed based on major variables reflecting the river characteristics. We evaluated the water quality of tributaries of the Nam River in South Korea. We analyzed the tributaries spatial characteristics using cluster analyses and selected the main water quality variables (Cluster 1: chemical oxygen demand [COD], total organic carbon [TOC], total nitrogen, and total phosphorus; Cluster 2: water temperature [WT], dissolved oxygen [DO], COD, and TOC; Cluster 3: WT, DO, electrical conductance, COD, and TOC) of the clustered rivers. The WQI for each tributary was calculated using variables selected post statistical analysis. We verified that the WQI calculated in this study was similar to the annual change in water quality of the target river. Finally, it was analyzed that performing river water quality evaluation using the major variables selected using statistical analysis reflects the current water quality status of the river in more detail. (WQI grade was S1 Good (63.0), S2 Poor (53.3), S3 Excellent (98.4), S4 Poor (48.4), S5 (Excellent (100.0), S6 Good (77.6), S7 Good (76.2), S8 Good (76.5), S9 Good (69.9), S10 Excellent (81.5), S11 Good (71.2), S12 Good (63.1), and S13 Good (63.5).) Our method effectively reduced the number of variables required for index calculation compared with WQI methods of the MOE. Further, the reduced number of variables simplified the analysis process, reduced analysis time, and enabled water quality assessment that reflected the water quality characteristics of the river to be evaluated.

Keywords Cluster analysis · Major variable · Principal component analysis · River tributaries · Spatial classification · Water quality management

Abbreviations

WQI	Water quality index
WT	Water temperature

Introduction

Rivers are a major source of drinking water and are also used for irrigation and industrial purposes; therefore, the prevention and management of river water pollution is crucial. Globally, climate change has led to a reduction in water resources, which is compounded by the increasing demand for water because of rising urbanization (Haque et al. 2014; Karakuş 2019). The increase in demand for water resources and the imbalance in availability can pose a threat to human sustainable development. For this reason, each country is evaluating the water footprint for water resource management; this is a concept that includes all water required from product consumption to disposal, considering the entire process of material production and service provision (Park et al. 2020). In Korea, the life cycle assessment (LCA) method was used to calculate the water footprint and evaluation data of the rainwater exclusion system in consideration of the environment. (Ahn et al. 2019). LCA is a tool to implement an environmental management system, and it is a technique to evaluate potential environmental impacts that may occur in the manufacturing process of a product (Chen et al. 2020). To achieve common performance for the circular economy from the management point of view of scarce water resources, a strategy to increase economic advantage and reduce environmental water loss by performing LCA is needed (Silvestri et al. 2021).

Reliable water quality data are required to prevent and manage water pollution in rivers, and continuous water

Editorial responsibility: Samareh Mirkia.

✉ H. G. Kwon
hun7082@korea.kr

[1] Nakdong River Environment Research Center, National Institute of Environmental Research, Dalseong-Gun, Daegu 43008, Republic of Korea

quality assessment must be conducted to monitor pollutants and water resource management based on spatial-scale water quality analyses (Wu et al. 2018). To assess water quality, several researchers have used methods, such as classification, correlation analysis, and time-series numerical analysis of measured data (Singh et al. 2008; Sharma et al. 2012). These methods have the advantage of not requiring expert knowledge about water quality or the environment. However, water quality assessment by time-series numerical analysis on long-term measured data requires a cumbersome and time-consuming data arrangement process.

Principle component analysis (PCA), used in regional assessment to analyze major water quality variables, enables an in-depth analysis of river characteristics through complex data matrix analysis (Tripathi et al. 2019). This method provides a quick solution for the proper management of water resources (Singh et al. 2004; Liu et al. 2011). Barakat et al. (2016) evaluated water quality variability of the Oum Er Rbia and El Abid rivers in Morocco according to spatial (cluster) classification by using PCA and cluster analysis (CA). These techniques are valuable for water resource management as they can identify and help investigate the origin of critical factors affecting rivers (Barakat et al. 2016; Varol 2020). Multivariate statistical analysis techniques, such as PCA and CA, reflect temporal and spatial changes in water quality and are used widely when evaluating water quality for river management by assisting in the selection of components that affect river water quality (Gamble and Babbar-sebens 2012; Bora and Goswami 2016).

The combined impact of many different factors that characterize the water quality is complex; and so are the challenges of classifying the significant parameters used to measure the status of water resources quantitatively. Therefore, the water quality index (WQI) is considered a mathematical tool that significantly simplifies these data sets and provides a single classifying value that describes the water quality status of water bodies or degree of pollution (Naseem et al. 2021). WQI is a performance measurement that combines information from significant physical, chemical, and biological parameters into a functional form, reducing large amounts of data to a single number (Gradilla-Hernández et al. 2019).

Using WQI, water quality is converted into a single score for comprehensive assessment, helping the public and policy makers to understand water quality (Terrado et al. 2010; Upadhyay et al. 2011; Gupta et al. 2017). In WQI, a water body is allocated a number between 0 and 100, which indicates its drinking water quality (Gradilla-Hernández et al. 2019). The WQI is the most effective way to express the suitability of a water body, such as a river, as a water source for human use (Ewaid et al. 2017; Khangembam et al. 2019). Further, WQI is used as a policy-making tool for water quality assessment by environmental monitoring organizations (Bharti and Katyal 2011), as the public can easily obtain and understand the relevant in-depth information (Naseem et al. 2021).

When performing water quality evaluation of the survey site using WQI, the same water quality variables were applied to each stream in the previous studies (Shin et al. 2018; Cho et al. 2021), whereas in this study, the water quality evaluation method that more quickly reflected the characteristics of rivers using major water quality variables for each stream was used. The applicability and feasibility of the results of this study were evaluated by selecting major variables reflecting the quality of rivers using statistical techniques. We also compared the WQI calculated in this study with the WQI currently presented by the domestic Ministry of Environment. These results may serve as basis for simplifying the number of variables for water quality evaluation of rivers in the future and for efficiently derive pollutants that need to be preferentially managed. In addition, this study may be used by policy makers for water quality improvement by region or by river.

Materials and methods

Study progress flow

Figure 1 shows the series of analysis procedures performed in this study. It was divided into three stages. First, cluster analysis by survey sites, and major water quality variables were selected through multivariate statistical analysis.

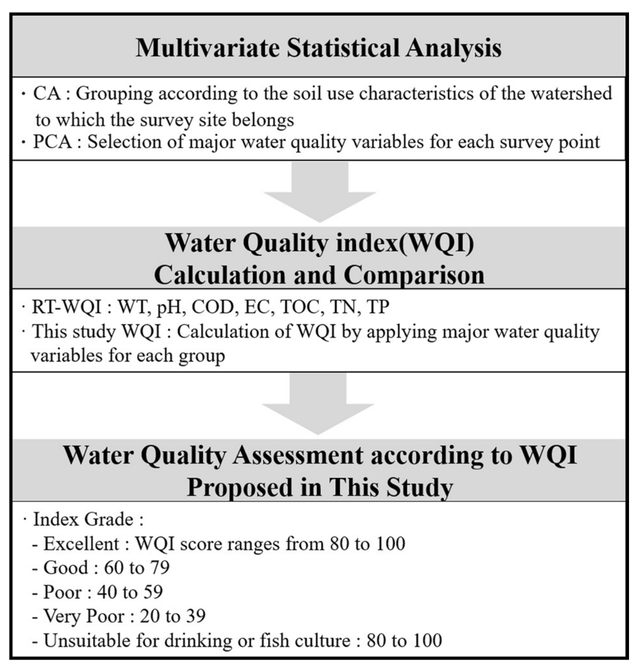


Fig. 1 Flowchart to illustrate how this study

Second, the WQI (according to the difference in the application of water quality variables) was calculated and compared with RT-WQI. Third, water quality evaluation was performed using the WQI calculation method proposed in this study.

Survey point and item

The Nam River is located in the southwest of South Korea. It has a total flow length of 144.59 km and a basin area of 3 467.5 km^2. From the upper stream to the lower stream, the major tributaries of the Nam River include the Namgang, Ram, Obong, and Im streams, the Dongcheon River, Sindeung, Yang, and Migok streams, Youngcheon River, and the Euiryung and Haman streams. The survey points and variables of this study are shown in Table 1 and Fig. 2. The survey points include the 13 water quality measurement points at the main tributaries of the Nam River operated by the Ministry of Environment. Water quality data (2012–2020) obtained from the Water Environment Information System (http://water.nier.go.kr) were used. From the acquired data, nine variables used in the RT-WQI were selected. In 2012, an artificial structure called a weir was installed in the Nakdong River, changing the water environment (Jo et al. 2022). Therefore, in this study, the monitoring results collected from 2012 to 2020, after the artificial structures were installed, were applied to the water quality evaluation.

Statistical analysis

CA and PCA were applied for the spatial classification of the measured data and selection of major water quality variables. Therefore, various rivers can be classified according to their characteristics, and the main water quality variables representing the water quality characteristics of the river could be determined (Pekey et al. 2004). CA and PCA were standardized to z-scores to prevent errors between data owing to different units of measurement. An statistical software (IBM SPSS 24.0 Inc.) was used for data analysis.

Cluster analysis

In this study, CA was performed for the spatial classification of nine water quality variables (water temperature [WT], dissolved oxygen [DO], electrical conductivity [EC], pH, biochemical oxygen demand [BOD], total organic carbon [TOC], chemical oxygen demand [COD], suspended Solid [SS], total nitrogen [TN], and total phosphorus [TP]) of the 13 tributaries of the Nam River (Fig. 2). For that, we used monthly survey results from 2012 to 2020.

CA collects and classifies variables with similar characteristics among multiple variables; it is a statistical analysis method that can identify the structure of data by grouping them according to the homogeneous properties of the variables. It classifies the data into clusters by using the distance or the similarity between the variables of the population, and it can be classified as a hierarchical or a non-hierarchical method, according to the method used to analyze the clusters. To classify the clusters as temporally and spatially homogeneous based on the correlation of the data (Vega et al. 1998), hierarchical CA, in which sequential clustering occurs, is used and is represented by a dendrogram (McKenna 2003). The hierarchical clustering method is used for classification; the Euclidean distance (Otto et al. 1998) is used for measuring the distance between clustered variables; and Ward's method (Fan et al.

Table 1 Survey point status and survey items

Point #	Measurement point name	Area (km^2)	Survey count and items
S1	Namgang stream	160.0	Amount of data used to calculate the WQI: 108 per variable at each site
S2	Ram stream	264.0	Survey interval: one per month
S3	Obong stream	218.2	Survey items: nine items
S4	Im stream		WT (Water temperature), EC (Electrical conductivity), DO (Dissolved oxygen), pH, BOD (Biochemical oxygen demand), COD (Chemical oxygen demand), TOC (Total organic carbon), TN (Total nitrogen), TP (Total phosphorus)
S5	Dongcheon river	354.5	
S6	Sindeung stream	172.1	
S7	Yang stream upstream	252.5	
S8	Migok stream		
S9	Yang stream downstream		
S10	Youngcheon river upstream	205.5	
S11	Youngcheon river downstream		
S12	Euiryung stream	114.4	
S13	Haman stream	155.2	

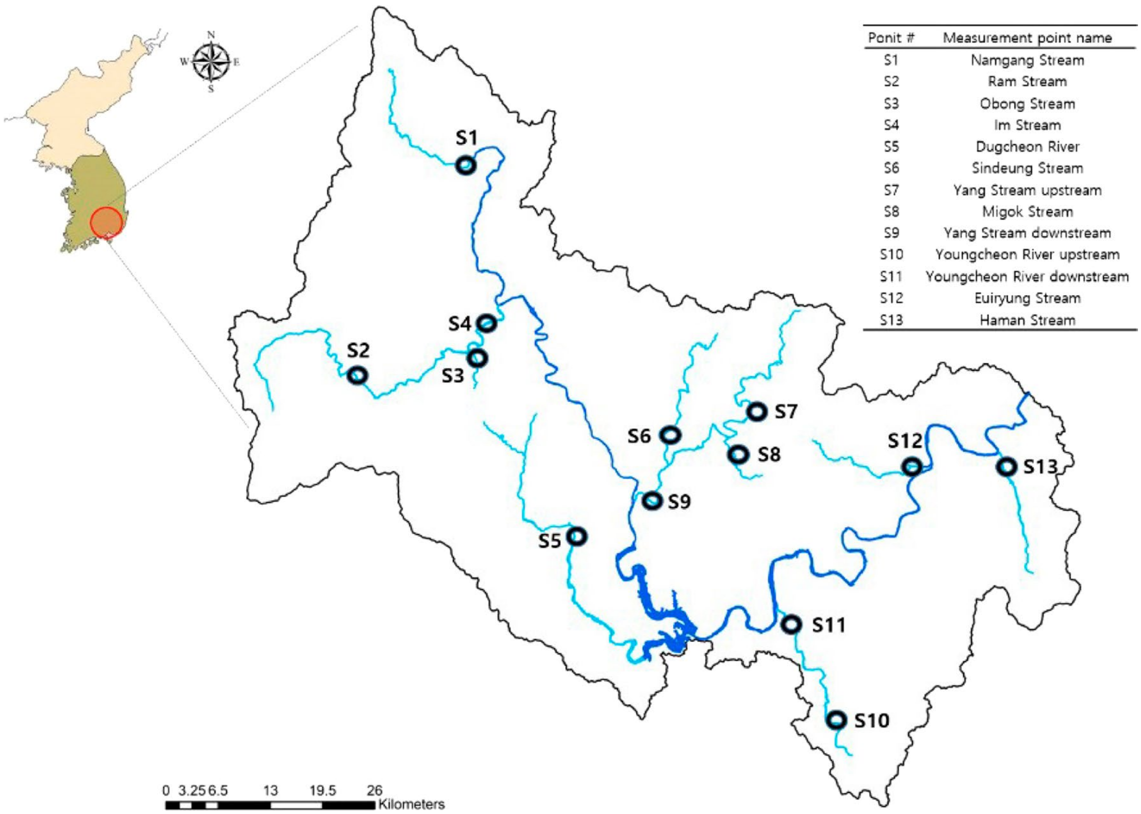


Fig. 2 Nam River watershed with 13 water quality measurement points. The first letter of all the words in the name of the measurement locations has been capitalized in this study

2010), which can minimize information loss between clusters, is used for merging.

Principal component analysis

PCA is used to find the direction of scattered data, re-express the data to facilitate understanding of the information in the data, and reduce the variability of variables. In other words, it is a statistical analysis method to find a new variable (principal component), expressed as a linear combination of variables by using the correlation between multiple variables (Sârbu and Pop 2005; Barakat et al. 2016). In PCA, the eigenvalue represents the magnitude of the variance that can be explained by the principal component. If the eigenvalue is greater than 1.0, it means that one principal component can explain more than one variable. Therefore, the principal component is extracted based on principal components with an eigenvalue of 1.0 or greater (Varol et al. 2012; Park et al. 2019).

In this study, PCA was performed to select the main water quality variables (WT, EC, DO, pH, BOD, COD, TOC, TN, and TP) by CA that could represent the characteristics of each stream and PCA was used to extract the components that could have the most significant influence on the water quality characteristics of the basin surveyed. Based on the CA results, PCA was performed on the water quality components of each cluster. The eigenvalue represents the magnitude of the variance that can be explained by the principal component, and as mentioned, an eigenvalue greater than 1.0 means that one principal component (PC) could explain more than one variable. We used the only PC with an eigenvalue of 1 in our analyses. With regard to the components of the ten water quality variables, the eigenvalue obviously represents the eigenvalue, and variability (%) is the ratio that explains the variability of the PC with respect to the original variable. Cumulative (%) is the cumulative ratio used to explain the PC.

Water quality index

WQI converts important water quality variables into a single index, which can be used as an easy-to-understand scale when evaluating the water quality of a target stream. Many researchers have used multivariate statistical techniques, such as PCA and CA, for selecting key variables for calculating WQI (Sutadian et al. 2016, 2018; Tripathi and Singal 2019; Gradilla-Hernández et al. 2019). With regard to the assessment criteria for the calculated WQI, we used five

Table 2 WQI assessment criteria provided by the Korean Ministry of Environment

*WQI		
Assessment details	WQI	Index grade
Clean water with almost no pollutants, suitable for recreational water activities at all times	80 to 100	Excellent
Relatively good water quality maintained, suitable for recreational water activities	60 to 79	Good
Generally, good water quality maintained, but pollutants are introduced from time to time that could affect recreational water activities	40 to 59	Poor
Contaminated water owing to frequent inflow of pollutants, raising a caution for recreational water activities	20 to 39	Very Poor
Severely contaminated water, not suitable for recreational water activities	0 to 19	Unsuitable for drinking or fish culture

**WQI* water quality index

grades suggested by the Ministry of Environment (Table 2). The five grades were Excellent (80 to 100), Good (60 to 79), Poor (40 to 59), Very Poor (20 to 39), and Unsuitable for Drinking Use and Fish Culture (0 to 19). With regard to the water quality variables for calculating WQI, the water quality data (WT, EC, DO, pH, BOD, COD, TOC, TN, and TP) measured at 13 measurement points from the Water Environment Information System (http://water.nier.go.kr), were used.

For quality assessment of the water of the Namgang inflow tributaries, we calculated the annual WQI from 2012 to 2020. The WQI was calculated as in Eq. 1 (Seo et al. 2021a, b) for seven factors, including WT, pH, DO, EC, TOC, TN, and TP, and by varying the main water quality variable for each group. The main water quality variables applied in the calculation of WQI for each cluster were selected by PCA.

$$\mathrm{WQI} = 100 - \sqrt{\left[\left(F_1^2 + F_2^2 + F_3^2\right)/3\right]} \tag{1}$$

where F_1 is the fraction of the number of water quality variables that violate the criteria, F_2 is the fraction of the total number of times the criterion is violated in all water quality variable, and F_3 is calculated by the sum of the values obtained by fractionating the degree of violation of the criteria for the water quality variable. Details are available from the Real-time Water Quality Information System (http://www.koreawqi.go.kr). The assessment for each survey point was conducted by varying the WQI calculation factors.

The water quality index calculated using the RT-WQI provided by the Ministry of Environment was compared with the index calculated based on the method proposed in this study using the PCs for each Cluster (Table 3 and Fig. 3). Seven water quality variables were applied to the RT-WQI analysis provided by the Ministry of Environment, including WT, pH, DO, EC, TOC, TN, and TP. However, the water quality variables used in our method of calculating the WQI (i.e., using the principal variables of each Cluster) were four in Cluster 1 (COD, TOC, TN, and TP), four in Cluster 2 (WT, DO, COD, and TOC), and five in Cluster 3 (WT, DO, EC, COD, and TOC).

Results and discussion

Results of water quality analysis according to spatial classification

CA was performed based on the area according to the land-use type in the river basin for the classification of the water quality concentration and spatial characteristics of the 13 survey points (Fig. 2). The CA results were classified into three clusters, and the survey points for each cluster were Cluster 1 (S1 to S5), Cluster 2 (S6 to S11), and Cluster 3 (S12 and S13). The analyses result of the characteristics of the basin for each survey point indicated that the clusters were formed by the influence of urbanization and human activities (Fig. 4). In terms of the land-use ratio, the average urbanization/drying ratio was 2.2% for Cluster 1, 3.6% for Cluster 2, and 5.7% for Cluster 3. The average ratio of agricultural land was 11.7% for Cluster 1, 17.9% for Cluster 2 17.9%, and 18.4% for Cluster 3, demonstrating a trend similar to the urbanization/drying ratio. As the ratio of the forests in the basin (Cluster 1: 78.8%, Cluster 2: 68.3%, and Cluster 3: 63.8%) that represented the natural characteristics of the basin showed an opposite trend from the previous results, apparently urbanization and human activities were the crucial factors defining the characteristics of the water basin. Seo et al. (2021a) identified water quality variables that affect the spatial difference and concentration of water quality by river through a study on the effect of spatial variability by river basin on the water quality of the Nakdong River in Korea.

The water quality characteristics were explained by dividing the water quality variables into clusters and expressing the average and range of concentrations for each variable

Table 3 WQI scores of the stations in the Nam River watershed during the study period

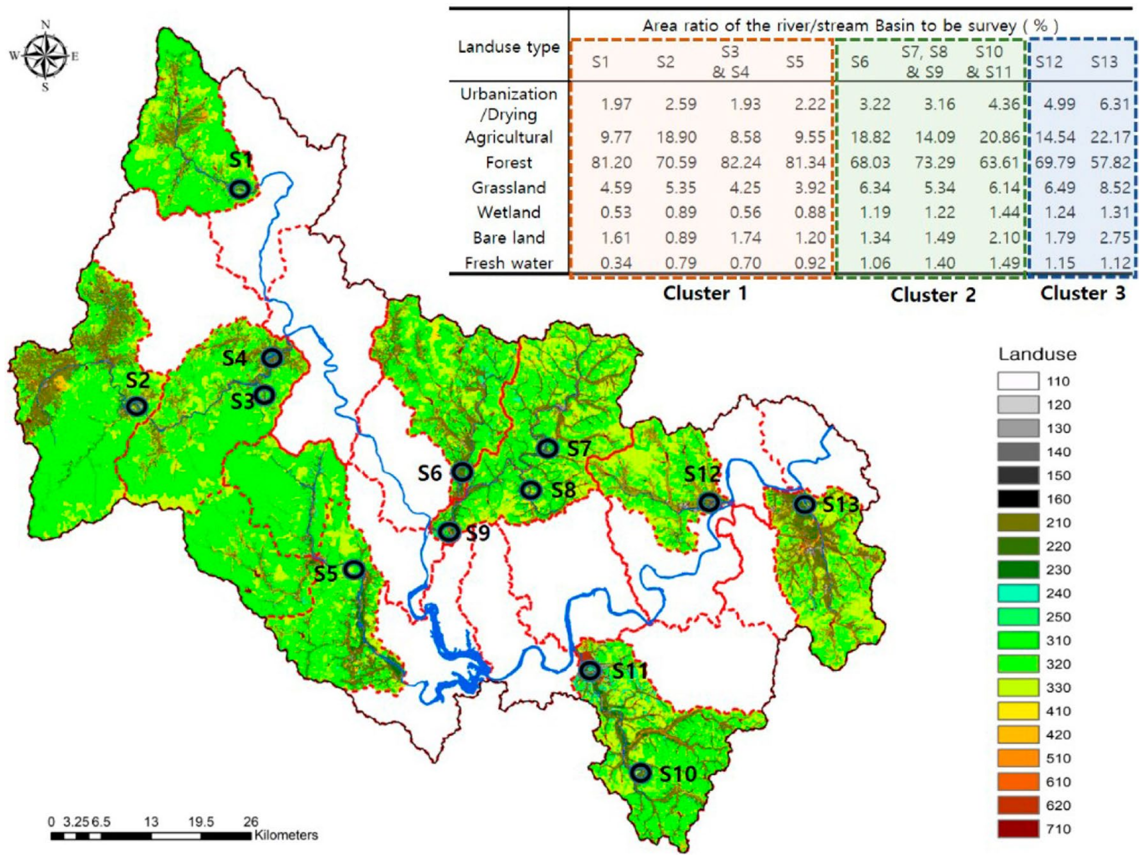
Year	Method	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
2012	RT-WQI*	82.51	76.37	100.00	100.00	100.00	87.52	79.22	75.16	94.04	92.46	76.19	68.99	59.52
	This study**	71.08	41.99	100.00	71.00	100.00	79.75	77.94	70.42	89.77	90.28	71.00	57.53	53.61
2013	RT-WQI	94.17	88.36	100.00	88.37	100.00	94.05	86.52	87.94	81.52	87.70	79.75	55.98	58.61
	This study	56.63	70.89	100.00	42.11	100.00	80.04	78.96	79.98	69.92	90.28	70.30	64.82	61.47
2014	RT-WQI	93.94	82.56	100.00	88.37	100.00	93.80	83.80	87.41	74.52	75.85	69.40	56.82	57.88
	This study	85.55	42.16	100.00	42.15	100.00	80.09	76.51	80.10	69.19	70.93	70.45	61.16	61.80
2015	RT-WQI	94.19	99.31	100.00	94.06	100.00	93.83	86.11	87.46	73.59	92.89	75.13	56.98	64.25
	This study	100.00	71.10	100.00	41.76	100.00	79.90	77.62	80.10	68.36	90.16	69.72	60.74	69.21
2016	RT-WQI	94.19	94.17	94.19	100.00	100.00	88.26	81.31	81.81	74.23	87.72	81.75	64.53	66.26
	This study	70.99	42.14	100.00	70.92	100.00	70.56	68.63	70.27	68.37	80.44	70.12	68.82	71.87
2017	RT-WQI	88.11	88.13	100.00	94.06	100.00	81.98	77.11	84.24	71.38	80.05	78.23	54.06	49.96
	This study	41.84	41.56	100.00	41.92	100.00	70.23	75.86	79.68	58.93	70.83	70.20	57.62	59.75
2018	RT-WQI	94.06	94.06	94.19	94.03	100.00	87.42	83.18	78.15	73.53	80.98	84.51	57.41	60.14
	This study	56.50	42.11	85.55	41.62	100.00	79.17	76.80	69.27	77.13	70.60	78.96	63.51	68.92
2019	RT-WQI	94.06	88.37	100.00	94.19	100.00	87.33	83.04	83.34	78.21	92.33	73.91	60.77	52.34
	This study	42.12	42.16	100.00	41.93	100.00	78.81	76.55	78.90	67.55	89.99	69.71	66.31	61.21
2020	RT-WQI	88.23	94.19	100.00	94.19	100.00	93.52	79.90	86.05	75.49	82.02	81.82	67.85	54.58
	This study	41.92	85.49	100.00	41.93	100.00	79.65	76.99	79.46	60.04	80.16	70.72	67.78	63.39
M ± SD	RT-WQI	91.5 ± 4.2	89.5 ± 6.9	98.7 ± 2.6	94.1 ± 4.1	100 ± 0	89.7 ± 4.3	82.2 ± 3.1	83.5 ± 4.5	77.4 ± 6.9	85.8 ± 6.3	77.9 ± 4.7	60.4 ± 5.5	58.2 ± 5.3
	This study	63 ± 20.7	53.3 ± 17.4	98.4 ± 4.8	48.4 ± 12.8	100 ± 0	77.6 ± 4.1	76.2 ± 3	76.5 ± 4.9	69.9 ± 9.2	81.5 ± 9	71.2 ± 2.9	63.1 ± 4.2	63.5 ± 5.7
t-test	*p*-value	0.003 (○)	0.000 (○)	0.865 (●)	0.000 (○)	– (●)	0.000 (○)	0.001 (○)	0.006 (●)	0.069 (●)	0.261 (●)	0.002 (○)	0.245 (●)	0.05 7(●)

The average water quality index (WQI) was calculated for nine years using two methods and conducting a t-test. ○: $p<0.05$, ●: $p>0.05$; *M*: mean; *SD*: standard deviation

*Variables used to calculate the RT-Water Quality Index (RT-WQI) are WT, pH, DO, EC, TOC, TN, and TP

**Variables used to calculate WQI in this study are (a) COD, TOC, TN, and TP (S1–S5), (b) WT, DO, COD, and TOC (S6–S11), and (c) WT, DO, EC, COD, and TOC (S12, S13

Fig. 3 Comparison of the assessment results for each method used to calculate the water quality index. Point S5 is not shown because the assessment results are the same



Landuse type	S1	S2	S3 & S4	S5	S6	S7, S8 & S9	S10 & S11	S12	S13
	Area ratio of the river/stream Basin to be survey (%)								
	Cluster 1				Cluster 2			Cluster 3	
Urbanization /Drying	1.97	2.59	1.93	2.22	3.22	3.16	4.36	4.99	6.31
Agricultural	9.77	18.90	8.58	9.55	18.82	14.09	20.86	14.54	22.17
Forest	81.20	70.59	82.24	81.34	68.03	73.29	63.61	69.79	57.82
Grassland	4.59	5.35	4.25	3.92	6.34	5.34	6.14	6.49	8.52
Wetland	0.53	0.89	0.56	0.88	1.19	1.22	1.44	1.24	1.31
Bare land	1.61	0.89	1.74	1.20	1.34	1.49	2.10	1.79	2.75
Fresh water	0.34	0.79	0.70	0.92	1.06	1.40	1.49	1.15	1.12

Fig. 4 Ratio of the land-use area in the basin to which the river belongs. The upper right table shows the ratio of area according to land use in the watershed at the survey sites, with the results being expressed using the cluster analysis. The land-use types include urbanization/drying land, agricultural land, forest, grassland, wetland, bare land, and fresh water

(Table 4 and Fig. 5). Interquartile range (IQR) was calculated by subtracting the third quartile by the first quartile, and the text (A, B, and C) on the box plot was used to characterize the homogeneity of the Scheffe test (analysis of variance, ANOVA; $p < 0.05$).

In the results shown in the box plots for each group, ANOVA results according to the Scheffe method also appeared to confirm the homogeneity between the groups. The WT was analyzed as 16.4 °C and 16.9 °C on average for Clusters 1 and 3, and 14.4 °C on average for Cluster 2. Among the investigated water quality variables, the mean concentration, and the range of change in concentration of electrical conductivity (EC), BOD, COD, TOC, TN and TP were the highest in Cluster 3 with a wide range, followed by Clusters 2 and 1 (Fig. 6). With regard to DO, Clusters 1 and 2 had similar mean values (11.0, 11.1 mg/L, respectively) and ranges (6.7 to 17.0 mg/L, 5.7 to 17.5 mg/L, respectively), whereas Cluster 3 had a lower mean concentration (10.1 mg/L). With regard to pH, similar concentrations and trends were observed for all clusters. With regard to the classification of the basin area by land use as shown in Fig. 4, the results show that the water quality characteristics of each watershed to which a stream belongs can be distinguished by the influence of human activities, such as urbanization and agriculture, similar to the results of CA. Compared with the other clusters, the water quality concentration of the basins in Cluster 1 was superior because of its strong natural characteristics, whereas the water quality concentration of the basins in Cluster 3 was inferior because of more urbanized and agricultural lands. Other studies have suggested that human activities affecting agricultural and urbanization of land have a significant positive correlation with water quality variables of rivers (Giri and Qiu 2016; Lintern et al. 2018). The results of analyzing clusters and water quality characteristics appeared to suggest the need for the differentiation of the method and target item for water quality assessment according to the land-use characteristics of the basin.

Results of principal component analysis

The results of the PCA indicated the cumulative explanatory power of Clusters 1 and 2 as 70% and 65%, respectively, with only three principal components, and the cumulative explanatory power of Cluster 3 as 82%, which could be

Table 4 Average concentration and concentration range of each water quality variable at survey points

Variables		Cluster 1					Cluster 2						Cluster 3	
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
WT (°C)	R	0–29.3	0.1–30.3	2.1–27.4	1.4–30.4	3.5–29.7	1.9–30.9	3.1–30.6	1.9–29.8	1.4–30.4	3.5–31	2.7–30.4	4.5–30	0.8–31.9
	M ± SD	13 ± 8.5	13.6 ± 8.6	14 ± 7.7	14.6 ± 8.2	16.7 ± 7.5	15.6 ± 8.5	16.8 ± 8.6	15.7 ± 8.5	16.4 ± 8.8	17.3 ± 7.9	16.5 ± 8.7	17.1 ± 7.5	16.7 ± 8.9
EC (μs/cm)	R	39–285	8–214	49–110	52–171	35–94	69–246	84–344	79–342	47–765	93–431	41–460	77–1852	103–637
	M ± SD	108.3 ± 40.8	120.4 ± 35.6	75 ± 13.3	102.9 ± 28.2	59.9 ± 13.9	151.5 ± 36.4	220.9 ± 44.7	232.8 ± 48.7	210.6 ± 108	220.4 ± 67.9	224 ± 86.3	635.7 ± 382.8	319.6 ± 117.7
DO (mg/L)	R	7.3–17	6.7–14.5	7.4–15.4	7.4–15.7	7.6–15.4	7.8–17.5	7–16.1	7.4–17.5	5.7–16.2	7.6–15.9	6.5–16.5	5.7–13.7	4.9–18.2
	M ± SD	11.5 ± 2.5	10.8 ± 1.9	10.8 ± 2.2	11.2 ± 2.3	10.9 ± 2.2	11.7 ± 2.5	11.3 ± 2.4	11.6 ± 2.1	10 ± 2.8	11.3 ± 2	10.7 ± 2.6	8.9 ± 1.9	11.3 ± 3.8
pH	R	6.6–8.7	6.2–9.2	7.1–8.1	7–8.6	7.1–8.5	7.1–8.9	7.1–9.1	7.1–8.7	6.7–9.1	7.2–8.9	0.8–9	6.4–9	7–9.4
	M ± SD	7.9 ± 0.5	8.4 ± 0.3	7.5 ± 0.2	7.7 ± 0.3	7.8 ± 0.3	8 ± 0.4	8.1 ± 0.5	7.9 ± 0.4	7.9 ± 0.6	8.1 ± 0.4	7.9 ± 1.1	7.6 ± 0.4	8.1 ± 0.6
BOD (mg/L)	R	0.1–2.7	0.1–4.3	0.3–0.8	0.4–1.8	0.3–1.2	0.4–2	0.5–2.7	0.4–2	0.3–3.2	0.3–1.7	0.4–4.8	0.3–3.9	0.7–8.8
	M ± SD	0.9 ± 0.6	0.8 ± 0.7	0.5 ± 0.1	0.8 ± 0.3	0.7 ± 0.2	1 ± 0.4	1.3 ± 0.5	1 ± 0.4	1.4 ± 0.6	1 ± 0.3	1.6 ± 0.9	1.7 ± 0.8	3.5 ± 1.9
COD (mg/L)	R	1.2–5.7	1.3–7.7	1–4.4	1.7–7.6	1.2–3.6	2.1–8.6	3.1–9.5	2.2–8.3	1.6–7.6	1.8–6.1	1.8–7.7	2.6–15.5	2.2–11.8
	M ± SD	2.7 ± 1.1	3.4 ± 1.3	2.2 ± 0.7	3.4 ± 1	2.3 ± 0.6	4.4 ± 1.6	5.5 ± 1.4	4.4 ± 1.3	4.5 ± 1.5	3.4 ± 1	4 ± 1.3	7.1 ± 2.7	6.6 ± 2
TOC (mg/L)	R	0.3–3.9	0.5–3.5	0.4–2.7	1–5.6	0.4–2.3	0.9–6	1.8–7.4	1.3–6	0.8–8.3	0.7–4.2	0.5–6.4	1.1–10.1	1.6–8.7
	M ± SD	2.1 ± 0.8	1.8 ± 0.7	1.1 ± 0.4	2.1 ± 0.8	1.2 ± 0.4	2.9 ± 1.3	4 ± 1.2	3 ± 1.1	4.1 ± 1.5	2.1 ± 0.8	3.2 ± 1.3	5.5 ± 2.2	4.8 ± 1.5
TN (mg/L)	R	0.921–2.675	0.87–2.679	0.377–1.266	0.693–2.752	0.706–2.186	0.56–1.996	0.397–3.284	0.256–2.851	0.329–4.277	1.117–3.263	0.309–2.914	0.737–7.811	0.797–5.243
	M ± SD	1.802 ± 0.436	1.735 ± 0.457	0.756 ± 0.192	1.442 ± 0.39	1.14 ± 0.332	1.2 ± 0.31	1.187 ± 0.572	1.142 ± 0.488	1.133 ± 0.645	2.085 ± 0.4	1.35 ± 0.606	2.996 ± 1.323	2.637 ± 1.048
TP (mg/L)	R	0.003–0.196	0.006–0.131	0.003–0.048	0.006–0.073	0.004–0.057	0.009–0.137	0.008–0.061	0.004–0.078	0.005–0.059	0.012–0.09	0.003–0.089	0.013–0.169	0.015–0.438
	M ± SD	0.025 ± 0.03	0.036 ± 0.025	0.014 ± 0.009	0.026 ± 0.015	0.019 ± 0.013	0.037 ± 0.026	0.024 ± 0.014	0.027 ± 0.018	0.023 ± 0.011	0.043 ± 0.019	0.029 ± 0.017	0.047 ± 0.028	0.089 ± 0.069

The average concentration and the concentration range are tabulated for 108 water quality data measured monthly by survey point from 2012 to 2020

R: range; *M*: mean; *SD*: standard deviation. Range is indicated by using the standard deviation and the minimum and maximum concentrations of the water quality variables at each survey point, with the average and standard deviation shown in the table

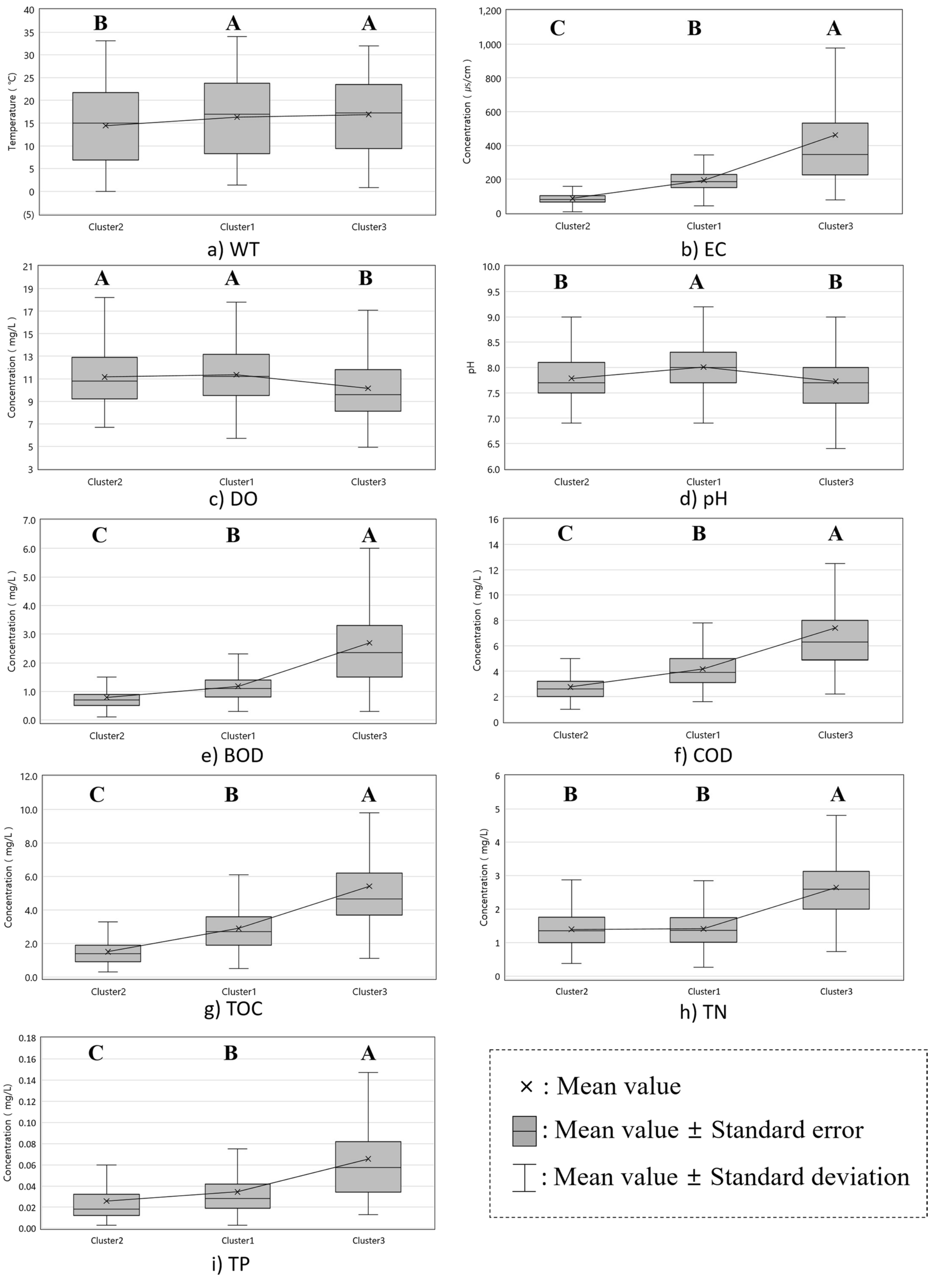


Fig. 5 Boxplot of the water quality measurement data from 2012 to 2020 for each group, confirming homogeneity according to spatial classification

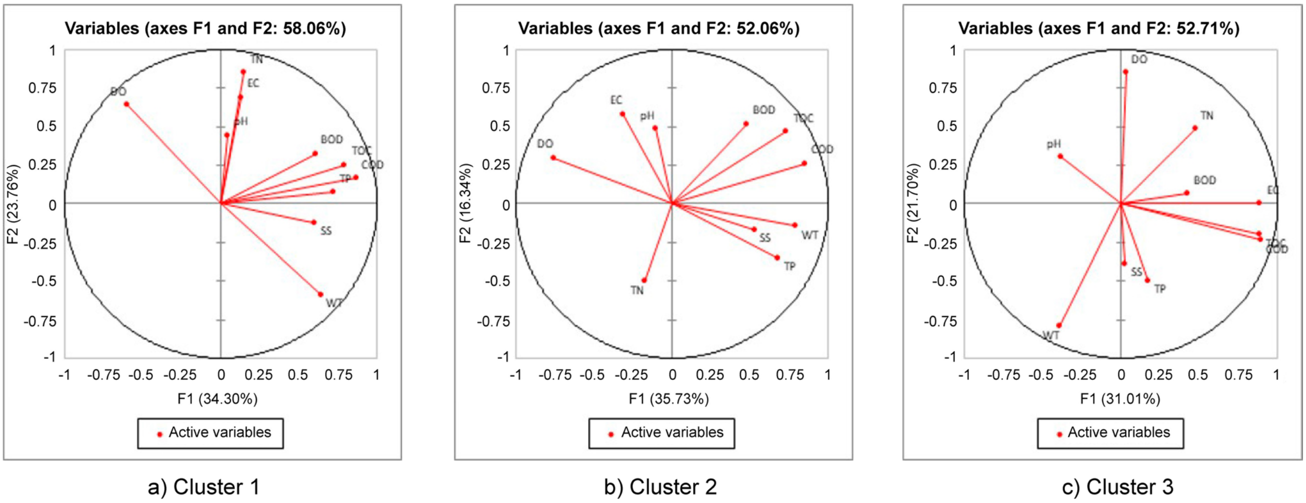


Fig. 6 Comparison of key water quality variables according to the cluster analysis. The larger the factor-loading value of the component, the closer it is to the circle, thereby indicating a stronger influence that could explain all variables (WT, pH, EC, DO, BOD, COD, TOC, SS, TN, and TP)

Table 5 Results of analyses to determine variability and eigenvalue of principal components (PC) by cluster

Division		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Cluster 1	Eigenvalue	**3.430**	**2.376**	**1.243**	0.769	0.687	0.486	0.380	0.319	0.239	0.071
	Variability (%)	34.297	23.763	12.433	7.686	6.874	4.863	3.796	3.186	2.394	0.709
	Cumulative (%)	34.297	58.060	70.492	78.179	85.052	89.915	93.711	96.897	99.291	100.000
Cluster 2	Eigenvalue	**3.573**	**1.634**	**1.333**	0.967	0.794	0.595	0.498	0.313	0.159	0.134
	Variability (%)	35.725	16.336	13.328	9.675	7.939	5.954	4.978	3.128	1.593	1.344
	Cumulative (%)	35.725	52.061	65.389	75.064	83.003	88.957	93.936	97.064	98.656	100.000
Cluster 3	Eigenvalue	**3.101**	**2.170**	**1.658**	**1.241**	0.631	0.440	0.323	0.197	0.160	0.079
	Variability (%)	31.012	21.696	16.579	12.412	6.310	4.396	3.235	1.967	1.600	0.793
	Cumulative (%)	31.012	52.707	69.287	81.699	88.009	92.405	95.640	97.607	99.207	100.000

Bold values indicate eigenvalues > 1.0

Eigenvalue: eigenvalue of the principal component; *Variability*: total variance, which is the ratio that describes the variability of the original variable; *Cumulative*: cumulative value of the total variance to describe the principal component

reduced to two principal components. As the number of variables decreases, the efficiency of PCA increases, indicating the value of the method proposed in this study, i.e., reducing a total of ten components to 3 and 4 PCs per Cluster (Table 5 and Fig. 6).

With regard to PCA, the factor-loading value for each component was considered "strong" when it exceeded 0.7, "moderate" when it was between 0.7 and 0.5, and "weak" when it was < 0.5. The results of PCA for each cluster and component are shown in Table 6. To select the principal component, the component corresponding to "strong" was selected among the factor-loading values of PC1 and PC2, which had the highest explanatory power.

Among the PC1 variables of Cluster 1, COD (0.864), TOC (0.794), and TP (0.718) were selected as the factor-loading values corresponding to "strong", that is, which have the strongest influence in evaluating the water quality characteristics of the Cluster. In the case of PC2, TN (0.852) was selected. For PC1 of Cluster 2, WT (0.794), DO (0.756), COD (0.854), and TOC (0.732) were selected, and there was no variable corresponding to "strong" in PC2. For PC1 of Cluster 3, EC (0.889), COD (0.895), and TOC (0.882) were selected; for PC2, WT (0.790) and DO (0.849) were selected as principal variables. Cho et al. (2021) selected four out of ten major water quality variables using PCA to identify major variables contributing to the water quality fluctuation characteristics of the South Han River in Korea. Kim et al. (2007) reported that in determining the number of major water quality variables, the determined variable was sufficient to explain the whole when the cumulative % was 60% to 80% or more.

Table 6 Principal component analysis by cluster

Cluster	Variable	Component			
		PC 1	PC 2	PC 3	PC 4
Cluster 1	WT	*0.638*	*−0.588*	−0.417	
	pH	0.045	0.446	−0.521	
	DO	*−0.603*	*0.640*	0.382	
	EC	0.130	*0.690*	−0.420	
	BOD	*0.603*	0.327	0.164	
	COD	**0.864**	0.167	−0.019	
	TOC	**0.794**	0.255	−0.132	
	SS	*0.601*	−0.121	0.544	
	TN	0.148	**0.852**	0.063	
	TP	**0.718**	0.075	0.362	
	Eigenvalue	3.430	2.376	1.243	
	Variability (%)	34.297	23.763	12.433	
	Cumulative (%)	34.297	58.060	70.492	
Cluster 2	WT	**0.794**	−0.146	−0.412	
	pH	−0.108	0.484	0.012	
	DO	**−0.756**	0.299	0.406	
	EC	−0.312	0.580	0.310	
	BOD	0.479	0.517	0.238	
	COD	**0.854**	0.263	0.173	
	TOC	**0.732**	0.467	−0.008	
	SS	0.527	−0.165	0.533	
	TN	−0.172	−0.499	0.584	
	TP	*0.679*	−0.350	0.437	
	Eigenvalue	3.573	1.634	1.333	
	Variability (%)	35.725	16.336	13.328	
	Cumulative (%)	35.725	52.061	65.389	
Cluster 3	WT	−0.389	**0.790**	0.114	−0.280
	pH	−0.379	−0.303	*0.661*	−0.332
	DO	0.031	**−0.849**	0.398	−0.040
	EC	**0.889**	−0.002	−0.204	−0.106
	BOD	0.425	−0.065	0.771	−0.101
	COD	**0.895**	0.233	0.117	−0.221
	TOC	**0.882**	0.197	0.023	−0.299
	SS	0.028	0.387	0.39	*0.590*
	TN	0.473	−0.488	−0.172	*0.561*
	TP	0.175	0.496	0.466	0.477
	Eigenvalue	3.101	2.170	1.658	1.241
	Variability (%)	31.012	21.696	16.579	12.412
	Cumulative (%)	31.012	52.707	69.287	81.699

Bold and italic values represent strong and moderate loading, respectively

The selected principal components representing the water quality variables at the survey points selected after performing PCA included COD, TOC, TN, and TP for Cluster 1; WT, DO, COD, and TOC for Cluster 2; and WT, DO, EC, COD, and TOC for Cluster 3.

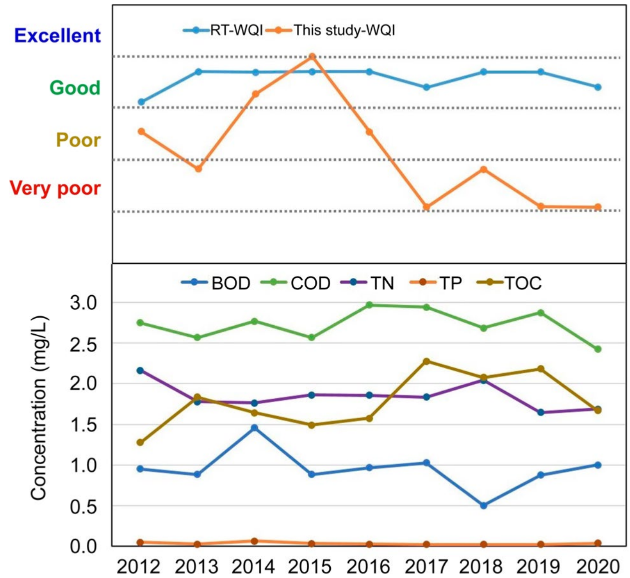


Fig. 7 Comparison of the water quality index assessment results and the annual average water quality at point S1. The difference between the existing method used for calculating the water quality index (RT-WQI) and the calculation method proposed in this study was analyzed

Results of water quality index assessment

After comparing the calculated WQIs (Table 3, Fig. 3), we obtained similar results for most sites despite the differences in the assessment factors. However, the results for S1, S2, and S4 differed. The river water quality using the existing RT-WQI for S1 was evaluated as "Excellent" every year from 2013 to 2016. However, the WQI calculation results applying the principal variables were evaluated differently by year (Very Poor, in 2013; Good, in 2014 and 2015; and Poor, in 2016).

The concentration of each water quality variable in the same year as the WQI assessment is shown in Fig. 7. From 2013 to 2016, which showed a large difference in the WQI assessment, the differences in concentration by year were identified clearly, and the same differences as in the previous results were observed from 2018 to 2019. As the index is calculated based on factors applicable to all the basins in South Korea for RT-WQI, it will be difficult to reflect the detailed characteristics of a specific basin. For S2 and S4, the WQI calculation results were analyzed differently depending on whether principal variables were applied. As a result of comparing the water quality concentration change by year and the calculation result (Fig. 7), it was determined that calculating the WQI by selecting the main variables for each river would reflect the characteristics of the target river in more detail. When the change in water quality by year and the WQI calculation result were compared, the concentration of water quality variables showed (BOD

0.5–1.5 mg/L, COD 2.4–3.0 mg/L, TN 1.647–2.162 mg/L, TP 0.020–0.065 mg/L, and TOC 1.3–2.3 mg/L) a change according to the surveyed year. However, the RT-WQI calculation result for the same year was analyzed as Good, indicating no change in water quality by year. Conversely, the WQI results of this study showed a similar trend from Excellent to Good according to the annual concentration change of water quality variables. Kang et al., (2019a) analyzed the long-term measurement data of the Nam River in Korea. The authors suggested that monthly river water quality variability is large. In addition, a study that evaluated the water quality grade for Namgang reported that the yearly and seasonal evaluation results showed similar trends to water quality variability (Kang et al. 2019b). According to Pejman et al. (2009), when water quality is evaluated in a watershed, the principal water quality variables reportedly appear different depending on the types of rivers distributed in the watershed. For some rivers (S1–S3), the assessment results were analyzed differently depending on the selection conditions of the water quality variables; however, when compared with the concentration of the water quality variables by year, the assessment results by the WQI applied in this study reflect the water quality.

The results of the annual water quality assessment from 2016 to 2020 for each river or stream were incorporated and compared with the nine-year average of water quality assessment for the water quality variables proposed in this study, as shown in Fig. 8. For survey points classified as Cluster 1 (S1, S2, S3, S4, and S5), the average and annual assessment results at the S3 and S5 points were Excellent. The average assessment result was Good at S1, but the annual assessment result was Poor in 2019 and 2020 compared with that in 2016, indicating that the water quality was deteriorating. The assessment result at S4 was the worst (Very Poor) of all the survey points. At S2, upstream of S4, the assessment result was Poor on average, but the water quality improved to Excellent in 2020, indicating that the water quality at S4 could have been affected partially by the upstream environment, but the extent of this influence was not absolute. Based on these results, the pollutants dispersed across the basin at S4 must be monitored to enhance the water quality. With regard to the survey points classified as Cluster 2 (S6, S7, S8, S9, S10, and S11), the water quality was assessed as Good or Excellent, except for that of S9. At S9, the water quality deteriorated from Good in 2018 to Poor in 2020. With regard to the survey points classified as Cluster 3 (S12 and S13), the water quality was assessed as Good on average; however, the WQI assessment indicated a value in the boundary between Poor and Good. This result points to the need for continuous water quality management in the future.

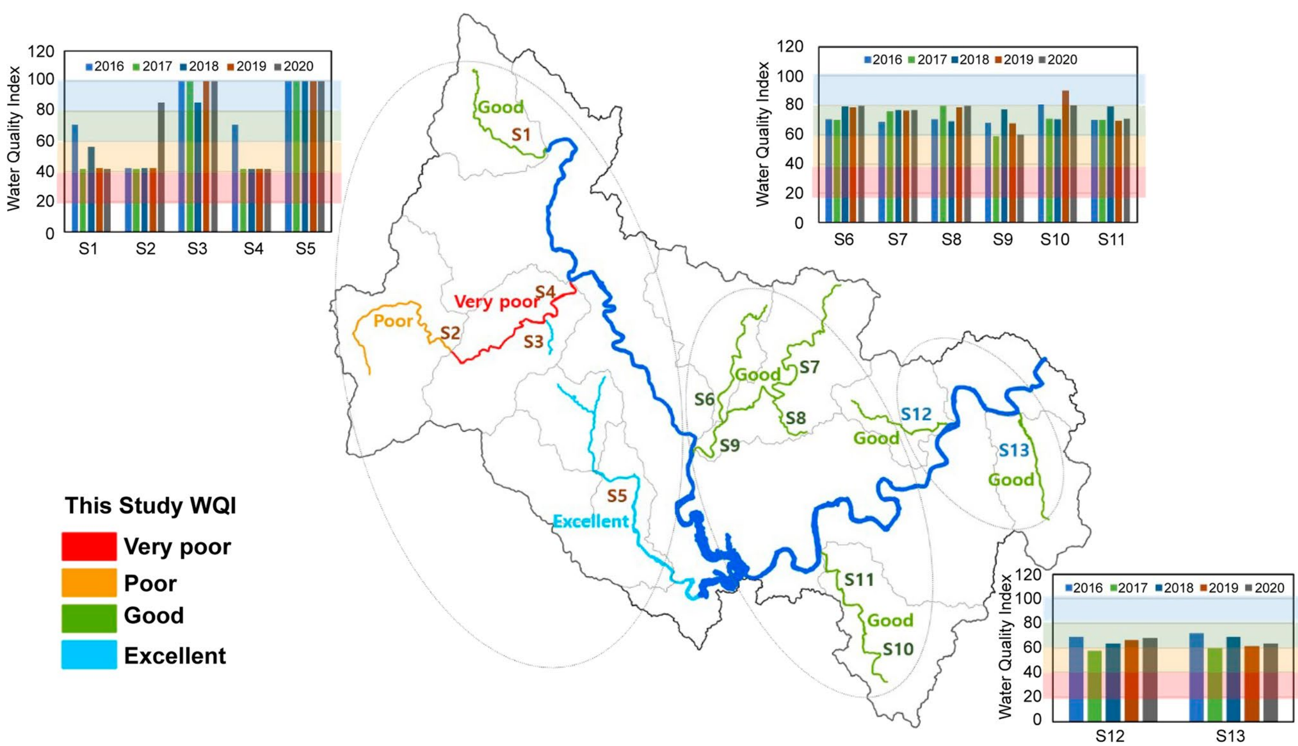


Fig. 8 Water quality assessment results for each river basin and analyses of the change patterns in the river basin compared to the results of the water quality assessment over the last five years. The assessment results over the last five years are grouped into previously analyzed clusters

Conclusion

In this study, statistical techniques were used to select the main water quality variables of streams, and the WQI was calculated by applying the same water quality variables to streams with the same spatial classification. By comparing the results obtained using the RT-WQI and those obtained by the WQI calculation method proposed in this study, the main water quality variables (COD, TOC, TN and TP for Cluster 1; WT, DO, COD and TOC for Cluster 2; WT, DO, EC, COD, and TOC for Cluster 3) for each river were selected to assess water quality. The selected major variables can be used as representative items to evaluate the characteristics of each river. As a result of the assessment, WQI grade was assessment as S1 Good (63.0), S2 Poor (53.3), S3 Excellent (98.4), S4 Poor (48.4), S5 (Excellent (100.0), S6 Good (77.6), S7 Good (76.2), S8 Good (76.5), S9 Good (69.9), S10 Excellent (81.5), S11 Good (71.2), S12 Good (63.1), and S13 Good (63.5). The water quality of a river is highly dependent on the land-use status of the basin, implying that the level of pollution would be high in the basin subjected to extensive human activity. Accordingly, such variables should preferably be considered in the assessment of water quality reflecting the of rivers and basins.

The limitation of the current study is that it only included the rivers with the water quality measurement points operated by the Ministry of Environment. For more accurate analyses than those of this study, a higher number of rivers should be included, and an additional review of ionic substances would be required for the selection of water quality parameters. And in the case of the current analysis result, the method of this study can be applied even to rivers that have undergone the same water quality evaluation as the existing RT-WQI. However, efforts to increase applicability through diverse case studies are needed for rivers that show differences in some evaluation results.

The existing water quality improvement management plan uses a single standard for all rivers and basins in South Korea, implying several limitations in terms of space, time, and expense. River management policies should be established in a direction that can improve the water quality of rivers distributed throughout the country (bottom-up method) by establishing a management plan reflecting the characteristics of the target rivers by watershed and improving the water quality. As can be seen from the results of this study, it is necessary to simplify the variables statistically and scientifically for water quality evaluation and to establish water quality evaluation methods according to land-use and river characteristics. Therefore, in the future, calculating and assessing WQI based on the characteristics of each river will enable priority management of rivers that are more vulnerable to pollution by establishing the direction of water quality and basin management for rivers.

Acknowledgements This work was supported by a grant from the National Institute of Environment Research (NIER), funded by the Ministry of Environment (MOE) of the Republic of Korea [grant number NIER-2022-01-01-043].

Author's contribution All authors contributed to the planning and structuring of the study. HGK wrote the manuscript and analyzed the data. CDJ conducted data collection. HGK reviewed the initial manuscript and provided suggestions. All authors read and approved the final manuscript.

Funding This work was supported by a grant from the National Institute of Environment Research (NIER), funded by the Ministry of Environment (MOE) of the Republic of Korea [grant number NIER-2022–01-01–043]. The funding agency had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

Ahn J, Park SW, Kim YI (2019) Development and application of integrated evaluation index of rainwater drainage system based on the life cycle assessment. J Korea Acad-Ind Coop Soc 20:1–7. https://doi.org/10.5762/KAIS.2019.20.7.1

Barakat A, El Baghdadi M, 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 4:284–292. https://doi.org/10.1016/j.iswcr.2016.11.002
Bharti N, Katyal D (2011) Water quality indices used for surface water vulnerability assessment. Int J Environ Sci 2:154–173
Bora M, Goswami DC (2016) Water quality assessment in terms of water quality index (WQI): Case study of the Kolong River, Assam, India. Appl Water Sci 7:3125–3135. https://doi.org/10.1007/s13201-016-0451-y
Chen D, Kim J, Batagalle V, Kim B-S (2020) Impact evaluation of water footprint on stages of drainage works. KSCE J Civ Environ Eng Res 40:225–231. https://doi.org/10.12652/Ksce.2020.40.2.0225
Cho YC, Choi HM, Ryu IG, Kim SH, Shin DS, Yu SJ (2021) Assessment of water quality in the lower reaches Namhan river by using statistical analysis and water quality index (WQI). J Korean Soc Water Environ 37:114–127. https://doi.org/10.15681/KSWE.2021.37.2.114
Ewaid SH, Abed SA (2017) Water quality index for Al-Gharraf river, southern Iraq. Egypt J Aquat Res 43:117–122. https://doi.org/10.1016/j.ejar.2017.03.001
Fan X, Cui B, Zhao H, Zhang Z, Zhang H (2010) Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. Procedia Environ Sci 2:1220–1234. https://doi.org/10.1016/j.proenv.2010.10.133
Gamble A, Babbar-Sebens M (2012) On the use of multivariate statistical methods for combining in-stream monitoring data and spatial analysis to characterize water quality conditions in the White River basin, Indiana, USA. Environ Monit Assess 184:845–875. https://doi.org/10.1007/s10661-011-2005-y
Giri S, Qiu Z (2016) Understanding the relationship of land uses and water quality in twenty first century: a review. J Environ Manag 173:41–48. https://doi.org/10.1016/j.jenvman.2016.02.029
Gradilla-Hernández MS, de Anda J, García-González A, Meza-Rodríguez D, Montes CY, Perfecto-Avalos Y (2019) Multivariate water quality analysis of Lake Cajititlán. Mexico Environ Monit Assess 192:5. https://doi.org/10.1007/s10661-019-7972-4
Gupta N, Pandey P, Hussain J (2017) Effect of physicochemical and biological parameters on the quality of river water of Narmada, Madhya Pradesh, India. Water Sci 31:11–23. https://doi.org/10.1016/j.wsj.2017.03.002
Haque MM, Rahman A, Hagare D, Kibria G (2014) Probabilistic water demand forecasting using projected climatic data for Blue Mountains water supply system in Australia. Water Resour Manag 28:1959–1971. https://doi.org/10.1007/s11269-014-0587-1
Jo CD, Lee CG, Kwon HG (2022) Effects of multifunctional weir construction on key water quality indicators: a case study in Nakdong River, Korea. Int J Environ Sci Technol. https://doi.org/10.1007/s13762-022-03973-8
Karakuş CB (2019) Evaluation of water quality of Kızılırmak River (Sivas/Turkey) using geo-statistical and multivariable statistical approaches. Environ Dev Sustain 1–35
Kang DH, So YH, Park KD, Kim IK, Kim BW (2019a) Monthly variation analysis of BOD, COD and TOC using long-term observation data in the downstream of the Nam river. J Environ Sci Int 28:571–579. https://doi.org/10.5322/JESI.2019.28.6.571
Kang DH, So YH, Park KD, Kim IK, Kim BW (2019b) Water quality level of the living environmental standards and regression analysis in the major tributaries of the Nakdong river. J Environ Sci Int 28:357–369. https://doi.org/10.5322/JESI.2019.28.3.357
Khangembam S, Kshetrimayum KS (2019) Evaluation of hydrogeochemical controlling factors and water quality index of water resources of the Barak valley of Assam, Northeast India. Groundwater Sustain Dev 8:541–553. https://doi.org/10.1016/j.gsd.2019.02.001
Kim JH, Choi CM, Kim WI, Lee JS, Jung GB, Han KH, Ryu JS, Lee JT, Kwun SK (2007) Multi-variate statistical analysis for evaluation of water quality properties in Korean rural watershed. Korean J Environ Agric 26:17–24
Lintern A, Webb JA, Ryu D, Liu S, Waters D, Leahy P, Bende-Michl U, Western AW (2018) What are the key catchment characteristics affecting spatial differences in riverine water quality? Water Resour Res 54:7252–7272. https://doi.org/10.1029/2017WR022172
Liu WC, Yu HL, Chung CE (2011) Assessment of water quality in a subtropical alpine lake using multivariate statistical techniques and geostatistical mapping: a case study. Int J Environ Res Public Health 8:1126–1140. https://doi.org/10.3390/ijerph8041126
McKenna JE (2003) An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. Environ Model Softw 18:205–220. https://doi.org/10.1016/S1364-8152(02)00094-4
Naseem A, Ishak MIS, Ahmad MI, Umar K, Yusuff MSMd, Anees MT, Qadir A, Almanasir YKA (2021) Modification of the water quality index (WQI) process for simple calculation using the multi-criteria decision-making (MCDM) method: a review. Water 13:905. https://doi.org/10.3390/w13070905
Otto M (1998) Multivariate methods. In: Kellner R, Mermet JM, Otto M, Widmer HM (eds.) Analytical Chemistry Wiley/VCH, Weinheim, Germany, p 916
Park DY, Kim SJ, Park HS, Chung SW (2019) Simulations of the effect of control and phosphate loading on the reduction of algae biomass in Gangjeong-Goryong Weir. J Environ Impact Assess 28:507–524
Park S, Lee M, Park K, An Y (2020) Calculating virtual water for international water transactions: deriving water footprints in South Korea. J Korea Water Resour Assoc 53:765–772. https://doi.org/10.3741/JKWRA.2020.53.S-1.765
Pejman AH, Bidhendi GRN, Karbassi AR, Mehrdadi N, Bidhendi ME (2009) Evaluation of spatial and seasonal variations in surface water quality using multivariate statistical techniques. Int J Environ Sci Technol 6:467–476. https://doi.org/10.1007/BF03326086
Pekey H, Karakaş D, Bakoğlu M (2004) Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses. Mar Pollut Bull 49:809–818. https://doi.org/10.1016/j.marpolbul.2004.06.029
Sârbu C, Pop HF (2005) Principal component analysis versus fuzzy principal component analysis a case study: the quality of Danube water (1985–1996). Talanta 65:1215–1220. https://doi.org/10.1016/j.talanta.2004.08.047
Seo MJ, Heo JH, Kim YS (2021a) Present and potential future critical source areas of nonpoint source pollution: a case of the Nakdong river watershed, South Korea. J Environ Sci Pollut Res Int 28:45676–45692. https://doi.org/10.1007/s11356-021-12976-w
Seo JY, Won JG, Choi JH, Kim SD (2021) Major watershed characteristics influencing spatial variability of stream TP concentration in the Nakdong river basin. J Korean Soc Water Environ 37:204–216. https://doi.org/10.15681/KSWE.2021.37.3.204
Sharma A, Singh AK, Kumar K (2012) Environmental geochemistry and quality assessment of surface and subsurface water of Mahi river basin, western India. Environ Earth Sci 65:1231–1250
Shin SH, Im TH, Kim SH, Shin DS, Kwon HG (2018) Consideration in applying korean-comprehensive water quality index(K-CWQI) to assessment of water quality in the major tributaries of Gumho-river basin. J Korean Soc Hazard Mitig 18:455–465. https://doi.org/10.9798/KOSHAM.2018.18.2.455
Silvestri L, Forcina A, Di Bona G, Silvestri C (2021) Circular economy strategy of reusing olive mill wastewater in the ceramic industry: how the plant location can benefit environmental and economic

performance. J Cleaner Prod 326:129388. https://doi.org/10.1016/j.jclepro.2021.129388

Singh AK, Mondal GC, Kumar S, Singh TB, Tewary BK, Sinha A (2008) Major ion chemistry, weathering processes and water quality assessment in upper catchment of Damodar River basin, India. Environ Geol 54:745–758. https://doi.org/10.1007/s00254-007-0860-1

Singh KP, Malik A, Mohan D, Sinha S (2004) Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti river (India)—A case study. Water Res 38:3980–3992. https://doi.org/10.1016/j.watres.2004.06.011

Sutadian AD, Muttil N, Yilmaz AG, Perera BJC (2016) Development of river water quality indices—A review. Environ Monit Assess 188:58

Sutadian AD, Muttil N, Yilmaz AG, Perera BJC (2018) Development of a water quality index for rivers in west Java Province, Indonesia. Ecol 85:966–982. https://doi.org/10.1016/j.ecolind.2017.11.049

Terrado M, Barcel D, Tauler R, Borrell E, de Campos S (2010) Surface water quality indices for the analysis of data generated by automated sampling networks. Trends Anal Chem 29:40–52

Tripathi M, Singal SK (2019) Use of principal component analysis for parameter selection for development of a novel water quality index: a case study of river Ganga India. Ecol 96:430–436. https://doi.org/10.1016/j.ecolind.2018.09.025

Upadhyay R, Dasgupta N, Hasan A, Upadhyay SK (2011) Managing water quality of River Yamuna in NCR Delhi. Phys Chem Earth 36:372–378. https://doi.org/10.1016/j.pce.2010.03.018

Varol M (2020) Use of water quality index and multivariate statistical methods for the evaluation of water quality of a stream affected by multiple stressors: a case study. Environ Pollut 266:115417. https://doi.org/10.1016/j.envpol.2020.115417

Varol M, Gökot B, Bekleyen A, Şen B (2012) Spatial and temporal variations in surface water quality of the dam reservoirs in the Tigris river basin, Turkey. CATENA 92:11–21. https://doi.org/10.1016/j.catena.2011.11.013

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 Res 32:3581–3592. https://doi.org/10.1016/S0043-1354(98)00138-9

Wu Z, Wang X, Chen Y, Cai Y, Deng J (2018) Assessing river water quality using water quality index in Lake Taihu Basin, China. Sci Total Environ 612:914–922. https://doi.org/10.1016/j.scitotenv.2017.08.293