Membrane and Water Treatment, *Vol. 13, No. 5 (2022) 245-257*
https://doi.org/10.12989/mwt.2022.13.5.245

Assessment and spatial variation of water quality using statistical techniques: Case study of Nakdong river, Korea

Shin Kim*

Nakdong River Environment Research Center, National Institute of Environmental Research, 24-11, Gukgasandan-daero 52-gil, Guji-myeon, Dalseong-gun, Daegu, Republic of Korea

(Received September 15, 2019, Revised August 3, 2022, Accepted August 30, 2022)

Abstract. Water quality characteristics and their spatial variations in the Nakdong River were statistically analyzed by multivariate techniques including correlation analysis, CA, and FA/PCA based on water quality parameters for 17 sites over 2017–2019, yielding PI values for primary factors. Site 10 indicated the highest parameter concentrations, and results of pearson's correlation analysis suggest that non-biodegradable organic matter had been distributed on the site. Five clusters were identified in order of descending pollution levels: I (Ib > Ia) > II (IIa > IIb) > III. Spatial variations started from sub-cluster Ib in which Daegu city and Geumho-river are joined. T-P, PO4-P, SS, COD, and TOC corresponded to VF 1 and 2, which were found to be principal components with strong influence on water quality. Sub-cluster Ib was strongly influenced by NO3-N and T-N compared to other clusters. According to the PIs, water quality pollution deteriorated due to non-biodegradable organic matter, nitrogen- and phosphorus-based nutrient salts in the middle and lower reaches, illustrating worsening water pollution due to inflows of anthropogenic sources on the Geumho-river, i.e., sewage and wastewater, discharged from Site 10, at which there is a concentration of urban, agricultural, and industrial areas.

Keywords: factor analysis/principal component analysis; correlation analysis; cluster analysis; Nakdong river; pollution index; water quality

1. Introduction

Rivers are closely related to human activities, and water quality is largely influenced by a combination of complex anthropogenic activities and natural factors (Ren *et al*. 2003, Xian *et al*. 2007). Water quality are affected by both natural and anthropogenic factors (Ma *et al*. 2020). Anthropogenic influences such as urbanization, industrial, and agricultural practices, chemical spill accidents, dam construction, and natural processes such as erosion and climatic conditions can also affect water quality (Zhang *et al*. 2009). In particular, in urban, agricultural, and industrial areas, anthropogenic pollutants such as wastewater effluent discharge are absorbed or deposited into rivers, functioning as major factors in deteriorating water quality. In order to improve and manage water quality influenced by various factors, it is necessary to understand changes in water quality characteristics and identify the main contributors to pollution, and various statistical analysis techniques have been used to conduct such a study (Fan *et al*. 2010, Yang *et al*. 2010, 2013, Dutta *et al*. 2018).

Multivariate statistical analysis has been applied as a useful method to reduce errors in interpreting uncertain data of water systems with complex water quality characteristics, effectively evaluating and estimating the main factors influencing the water environment (Ahn and Yang 2007, Kim *et al*. 2014). Multivariate statistical techniques, such as cluster analysis (CA), factor analysis, and principal component analysis (FA/PCA), have been widely used to interpret water quality data in the identification of possible sources that influence water systems, offering valuable tools for the reliable management of water resources (Zeng and Rasmussen 2005, Ouyang *et al*. 2006, Hussain *et al*. 2008, Razmkhah *et al*. 2010). The combine use of different multivariate statistical techniques has been increasingly used in the assessment of water quality (Alves *et al*. 2018). In the Nakdong River, various studies using these techniques have been conducted, mostly focused on the middle and lower reaches, tributaries joining the mainstream, and other specific areas where anthropogenic pollution sources are distributed (Kim *et al*. 2010, Choi *et al*. 2011, 2012). To the best our knowledge no studies exist on the spatial variations and water quality characteristics of each area in the Nakdong River basin. In this study, multivariate statistical analysis was performed to more reliably interpret water quality data and extract factors affecting water quality.

Characterization of the spatial variation and source apportionment of water quality parameters can provide an improved understanding of the conditions of the water environment, helping researchers establish priorities for sustainable water management (Kolovos *et al*. 2002, Wang 2002, Chang 2005). The Nakdong river is an important water resource that provides drinking water to approximately 13 million people, it also supplies water to industrial complex and large cities along the river (Seo *et al*. 2019). This water system is affected by multiple sites of point source and non-point source pollution discharged from

---

*Corresponding author, Ph.D.,
E-mail: sinima@korea.kr

http://www.techno-press.org/?journal=mwt&subpage=7

ISSN: 2005-8624 (Print), 2092-7037 (Online)

urban, agricultural, and industrial areas, and is characterized by high spatial variations (Suh *et al*. 2019). Different pollutant types and effluent from sewage and wastewater treatment plants near the Nakdong River have flowed into the mainstream through tributaries, deteriorating the water quality and directly affecting the aquatic ecosystems (Lee *et al*. 2016). Therefore, studies on spatial variations and water quality assessment considering the effects of complicated water quality parameter inflows are necessary to more comprehensively understand the aquatic ecosystems of the Nakdong River water environment, as well as its needs in terms of water quality management.

Accordingly, this study targeted the Nakdong River (17 sites, 2017–2019), aiming to analyze correlations between various water quality parameters, identify water quality characteristics, and estimate the spatial variations of those characteristics. Further, clusters were set based on water quality characteristics through CA, identifying the main factors which heavily influenced water quality in the Nakdong River basin and analyzing the clusters configured through FA/PCA. Water quality pollution assessment was conducted by calculating the pollution index (PI) of the main contributing parameters. Based on these results, this study assessed river water quality characteristics, indicated their spatial variations, and identified target parameters and areas to be managed first for improving water quality. Furthermore, this study can be used as a basis for reference data necessary for understand pollution resource and spatial variations, and for reduce water pollution and improve water quality, not only in the Nakdong River, but also in other river system.

2. Material and methods

2.1 Study area

The Nakdong River study area has a total area of 23,384.21 km^2, a mainstream length of 400.7 km, and a river length of 510.36 km, accounting for nearly 25% of the total area in South Korea. The geographic location is 34°59’41″–37°12’52″ N latitude, and 127°29’19″–129°18’00″ E longitude, and includes three metropolitan cities (Pusan, Daegu, and Ulsan) and certain parts of five provinces (Gyeongsangbuk-do, Gyeongsangnam-do, Jeollanam-do, Jeollabuk-do, and Gangwon-do). In addition, the Nakdong River basin has complex land-use patterns, including agricultural and livestock activities in the upstream part, and the formation of a densely populated metropolis and industrial activities in the middle and lower reaches. This study targeted a total of 17 sites, ranging from Site 1 in Andong city, Gyeongsangbuk-do, in the Nakdong River upstream reaches, to Site 17 in Busan Metropolitan city downstream. The sites are geographically divided into upstream (Sites 1–6), middle stream (Sites 7–12), and downstream (Sites 13–17) (Fig. 1a). Small and large tributaries join the mainstream between each site; among them, the larger tributaries include Wi-cheon between Sites 3 and 4, Gam-cheon between Sites 5 and 6, Geumho-river between Sites 9 and 10, Hwang-river between Sites 12 and 13, Nam-river between Sites 13 and 14 and Milyang-river between Sites 15 and 16. In the Nakdong River, tributaries in the middle and lower reaches tend to be larger than those in upstream. Pollution sources generated from urban, industrial, and agricultural areas, which are located near all 17 sites, flow directly into the river or via the tributaries, affecting overall water quality. Accordingly, current population, and the status of agricultural and industrial areas near the were obtained using the Water Emission Management System (http://wems.nier.go.kr) provided by the National Institute of Environmental Research of South Korea. Among all sites, Site 10 in Daegu city had the highest population of more than 50,000; thus, it was assumed that the site recorded the highest inflow of domestic wastewater due to urbanization. It was found that the basin areas of Sites 1 and 10 were the largest, covering 23.6 and 27.7 km^2, respectively. In particular, Site 10 consisted of the largest agricultural (4.4 km^2) and industrial (0.115 km^2) areas, which have the highest risk of inflow from sewage and wastewater (Fig. 1b).

2.2 Water quality analysis methods

This study area employed monthly average data of water quality analysis for three years (from January 2017 to December 2019), targeting the 17 sites in the mainstream of the Nakdong River. Precipitation (PCP) data were collected from the Water Resources Management Information System (http://www.wamis.go.kr) provided by the Han River Flood Control Office of South Korea. Among the water quality parameters, a total of four components were measured at the site using a multiparameter water quality meter (YSI-650MDS, USA): water temperature (WT), dissolved oxygen (DO), electrical conductivity (EC), and pH. Ten further water quality parameters were analyzed as indirect indicators of organic matter pollution: biochemical oxygen demand (BOD), chemical oxygen demand (COD), total organic carbon (TOC); the nutrient salt nitrogen-based compounds total nitrogen (T-N), ammoniacal nitrogen (NH_3-N), and nitrate nitrogen (NO_3-N); the phosphorus-based compounds total phosphorus (T-P) and inorganic dissolved phosphorus (PO_4-P); and suspended solids (SS) and chlorophyll-a (Chl-a) (Table 1). The field survey and study analysis methods were based on the official test methods for water quality according to the Ministry of Environment of Korea (MOE 2012).

2.3 Statistical analysis methods

In this study, multivariate statistical analysis techniques were applied using SPSS 20.0 software. The widely used pearson’s correlation analysis was employed to determine correlations between the water quality parameters; as the p-value was set to less than 0.05, coefficients with higher values were assumed to have non-significant correlations. CA was applied to identify groups of samples with similar water quality parameters (Kumari *et al*. 2013). This study used Ward’s methods, a hierarchical CA designed to minimize the loss of data between clusters, and each cluster was set by converting similarities into distances using

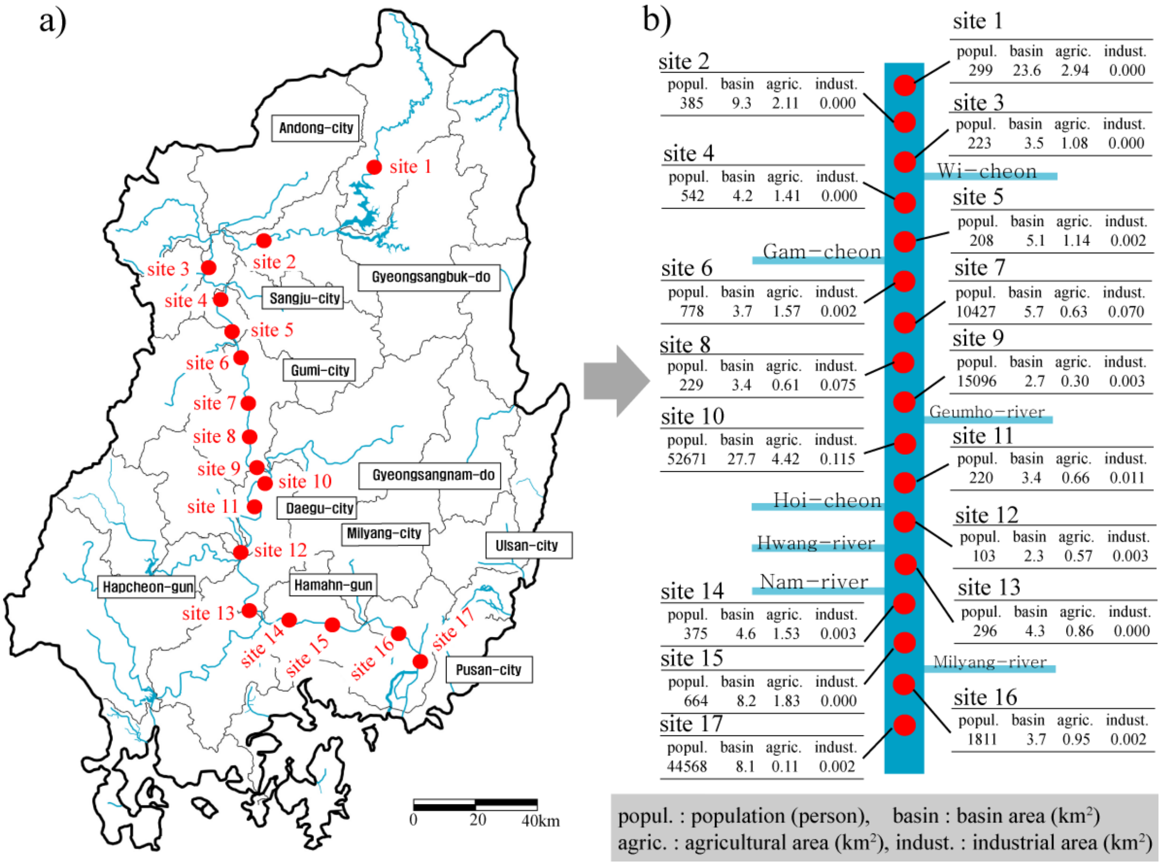


Fig. 1 Monitoring sites in the Nakdong river (a): site map, (b): summary map, population, basin, agriculture, and industrial areas

Table 1 Water quality parameter, abbreviations, analytical methods, and equipment used in this study

Parameter	Abbreviation	Analytical Methods (equipment model, country)
Biochemical oxygen demand	BOD	Dilution and seeding 5-days (YSI 5100-115v, USA)
Chemical oxygen demand	COD	Potassium permanganate (Water bath)
Total organic carbon	TOC	High-temperature combustion (AnalytikJena multi N/C 3100, Germany)
Total nitrogen	T-N	Continuous flow analysis (Skalar SAN++, Netherlands)
Ammonium nitrogen	NH_3-N	Ion chromatography (J.L.Science Metrohm850, Swiss)
Nitrate nitrogen	NO_3-N	UV-visible spectrometry (Aliance Smartchem200, France)
Total phosphorus	T-P	Continuous flow analysis (Skalar SAN++, Netherlands)
Inorganic dissolve phosphorus	PO_4-P	UV-visible spectrometry (Aliance Smartchem200, France)
Suspended solids	SS	Filtration methods (GF-C, and Dry oven)

Euclidean distance (Otto 1998). Based on these results, clusters were illustrated as a dendrogram (Mckenna 2003). Prior to performing FA/PCA, the Kasier-Meyer-Olkin (KMO) and Barlett's tests were conducted to determine whether clusters were applicable to water quality parameter data. Based on the CA results, FA/PCA was performed on all clusters together as well as each configured cluster. In addition, FA/PCA was performed on the correlation matrix between the different parameters followed by Varimax rotation, with the same being used to examine the association between water parameters (Kuppusamy and Giridhar 2006).

2.4 Pollution index method

In order to estimate the pollution level in the study area, the pollution index for each water quality parameter was calculated as follows (Su *et al*. 2011).

$$PI = C_i/C_0 \quad (\mathrm{i} = 1, 2, \ldots \mathrm{n}) \tag{1}$$

where *Ci* is the concentration of the analyzed water quality parameter and *Co* is the standard concentration of each parameter. In this study, standard concentrations were based on the environmental standards for river water quality of Korea; except for T-N which does not currently have a

Table 2 Environmental standards for river water quality of Korea

	class	BOD (mg/L)	COD (mg/L)	TOC (mg/L)	SS (mg/L)	DO (mg/L)	T-P (mg/L)	T-N* (mg/L)
a	Excellent	≤ 1	≤ 2	≤ 2	≤ 25	≥ 7.5	≤ 0.02	≤ 0.2
b	Good	≤ 2	≤ 4	≤ 3	≤ 25	≥ 5.0	≤ 0.04	≤ 0.3
	Above average	≤ 3	≤ 5	≤ 4	≤ 25	≥ 5.0	≤ 0.10	≤ 0.4
	Fair	≤ 5	≤ 7	≤ 5	≤ 25	≥ 5.0	≤ 0.20	≤ 0.6
	Poor	≤ 8	≤ 9	≤ 6	≤ 100	≥ 2.0	≤ 0.30	≤ 1.0
	Very poor	≤ 10	≤ 11	≤ 8		≥ 2.0	≤ 0.50	≤ 1.5

*T-N: Environmental standard for lake water quality of Korea

standard concentration for rivers, and so was set based on the environmental standard for lake water quality of Korea. The environmental standard water quality in Korea is divided into six classes; class I is subdivided into classes Ia and Ib, with class Ib corresponding to “good” and thus indicating a good water quality condition (Table 2). Therefore, the concentration of Ib was used as the standard concentration; if a PI value was less than 1, this would indicate a non-contaminated environment, while higher PI values indicated a higher level of contamination.

3. Results and discussion

3.1 Results of water quality

Fig. 2 shows the analysis of the water quality components for each site during the study period. Overall, the study area showed an increasing trend in pH (7.0–9.1) in the middle and lower reaches compared to upstream. Water temperature and DO showed similar patterns at all sites; DO was between 5.8–16.8 mg/L without forming oxygen-deficient water masses (< 2.0 mg/L) at all sites. It is known that pH affects chemical and biological processes, and that temperature affects the availability of oxygen concentration (Kowalkowski *et al*. 2006). In particular, pH affects the growth of algal populations (Na *et al*. 2016); as algae such as phytoplankton grow, carbonates and bicarbonates in the water are absorbed and increased by photosynthesis (Kim *et al*. 2002).

The concentration of Chl-a, a component indicating algal density, tended to increase in the middle and lower reaches (> 19 mg/m^3) compared to upstream (< 19 mg/m^3). As a result, the study area showed more active growth of algae in the middle and lower reaches than in the upstream part. It has been estimated that for the middle and lower reaches, there is a high risk of eutrophication in which algae growth expands and DO is decreased, especially in the summer when water temperature increases. EC exceeded 300 µs/cm at sites in the middle and lower reaches (Sites 10, 11, and 12), and SS was also increased in that it had higher concentrations in the middle and lower reaches compared to upstream, which was a similar pattern to EC. It is assumed that this spatial pattern of increase in SS was due not only to seasonal factors such as precipitation, but also to a large amount of suspended matter inflowed from tributaries which are distributed more heavily among the middle and lower reaches than in the upstream part (Jung and Kim 2017). BOD showed “excellent” water quality corresponding to Ia at Sites 1 and 2 located in the farthest upstream part, increasing to 2.2 mg/L at Site 10 in the middle-stream, and then showing a relatively lower water quality of class II at Site 14. COD was found to be less than 5 mg/L in the upstream part (Sites 1 and 2) indicating no contamination; however, all other sites showed a level of pollution corresponding to class III with a concentration exceeding 5 mg/L; in particular, the maximum COD concentration (6.7 mg/L) was observed at both Sites 11 and 12. As TOC was less than 4 mg/L at Sites 1, 2, and 3, which belonged to the upstream part, those sites were classified as Ib and class II; all other sites showed values exceeding 4 mg/L corresponding to class III, with midstream sites (Sites 10–13) showing the maximum TOC concentrations. Organic matter indicators such as BOD, COD, and TOC showed distinctive increasing trends from Site 10, where the maximum concentration appeared; it can be seen that there was a high concentration of organic matter downstream. Nitrogen-based nutrient salts such as T-N, NH_3-N, and NO_3-N had maximum concentrations in the midstream part; T-N also showed a high concentration exceeding 3 mg/L from Sites 10 to 13. NH_3-N and NO_3-N also showed an increasing trend in the midstream part; Sites 10 and 11 had the highest concentrations, with NH_3-N exceeding 0.1 mg/L and NO_3-N exceeding 2.5 mg/L. Phosphorus-based nutrient salts such as T-P and PO_4-P also showed a similar pattern to organic matter and nitrogen-based nutrient salts, showing the highest concentration in the midstream part; T-P was found to belong to classes Ia and Ib at all sites. Overall for the study area, results showed high concentrations of organic matter and nutrient salts at Site 10, affecting the middle and lower reaches.

3.2 Correlation analysis

Pearson’s correlation analysis was performed to identify the correlation between all water quality components in this study (Table 3). PCP showed high positive correlation coefficients with T-P (0.650) and PO_4-P (0.617) along with water temperature (0.659); PCP also had positive correlations with TOC (0.396), COD (0.461), and SS (0.386). SS showed positive correlations with TOC (0.523), COD (0.510), T-P (0.500), and PO_4-P (0.323). It is assumed

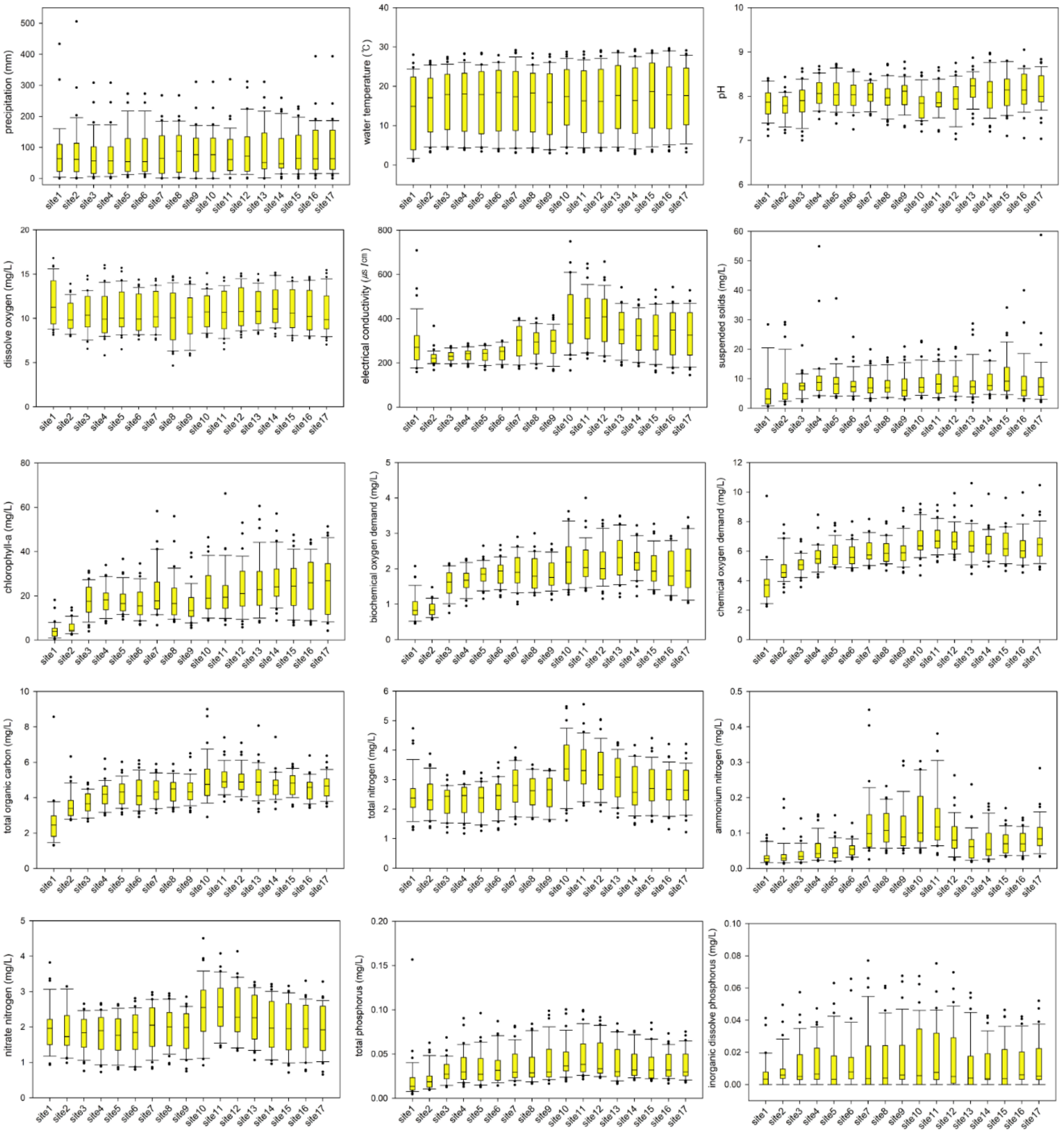


Fig. 2 Box plot of water quality data in each sites of the study area (from 2017 to 2019)

that these results were due to the increased concentration of SS, organic matter, and phosphorus-based nutrient salts inflowing into the water system because of non-point source pollution, such as soil runoff during summer coupled with high precipitation, as well as to the simultaneous discharge of phosphorus adsorbed onto SS (Lee and Kim 2017). DO showed a high negative correlation with water temperature (-0.916), likely caused by a typical seasonal factor in which the oxygen depletion rate increases due to the decomposition of organic matter during summer when the water temperature also increases (Na *et al*. 2015). As Chl-a had no significant correlation with PCP or water temperature, it was interpreted to have little or no seasonal effect. This result differed from other research outcomes, which have showed that Chl-a was generally highly correlated with water temperature. This can be explained in that during summer, when water temperature increases, blue-green algae *Microcystis* species tend to proliferate, leading to an increase in Chl-a, a phenomenon in which the low-temperature diatoms *Stephanodiscus sp*. show optimal growth despite the low water temperature (Joung *et al*. 2013, Seo *et al*. 2010). In addition, as there were positive correlations of Chl-a with organic matter components such as BOD (0.721), COD (0.515), and TOC (0.502), it can be assumed that (marine) autogenous organic matter increased in relation to the algal growth (Kim *et al*. 2013). Although

Table 3 Pearson’s correlation coefficients among water quality parameter

	PCP	**WT**	**pH**	**DO**	**EC**	**SS**	**Chl-a**	**BOD**	**COD**	**TOC**	**T-N**	**NH$_3$-N**	**NO$_3$-N**	**T-P**	**PO$_4$-P**
PCP	1														
WT	0.659**	1													
pH	-0.139**	0.107**	1												
DO	-0.609**	-0.916**	0.053	1											
EC	-0.431**	-0.464**	0.137**	0.496**	1										
SS	0.386**	0.216**	-0.118**	-0.198**	-0.215**	1									
Chl-a	0.037	0.044	0.422**	0.157**	0.231**	0.087*	1								
BOD	-0.003	-0.016	0.392**	0.186**	0.447**	0.068	0.721**	1							
COD	0.461**	0.409**	0.124**	-0.301**	0.008	0.510**	0.515**	0.551**	1						
TOC	0.396**	0.338**	0.107**	-0.224**	0.045	0.523**	0.502**	0.585**	0.895**	1					
T-N	-0.303**	-0.603**	-0.112**	0.597**	0.644**	-0.083*	0.141**	0.280**	0.032	0.074	1				
NH$_3$-N	0.010	-0.079	-0.112**	0.003	0.366**	0.001	0.033	0.295**	0.216**	0.250**	0.323**	1			
NO$_3$-N	-0.346**	-0.665**	-0.143**	0.657**	0.526**	-0.101*	0.094*	0.200**	-0.080*	-0.005	0.949**	0.176**	1		
T-P	0.650**	0.546**	-0.263**	-0.506**	-0.312**	0.500**	0.148**	0.158**	0.672**	0.581**	-0.126**	0.187**	-0.188**	1	
PO$_4$-P	0.617**	0.537**	-0.363**	-0.545**	-0.473**	0.323**	-0.112**	-0.095*	0.473**	0.380**	-0.210**	0.093*	-0.228**	0.871**	1

*: correlation is significant at the 0.05 level, **: correlation is significant at the 0.01 level

Chl-a indicated low yet significant positive correlations with T-N (0.141) and T-P (0.148), although it had negative correlation with PO$_4$-P (-0.112). It is believed that, unlike other nutrient salts, PO$_4$-P can be easily transformed within the body of phytoplankton in water (Kappers 1980). T-N showed high positive correlation with NO$_3$-N (0.949), and these two components also showed high correlation with EC (T-N: 0.597 and NO$_3$-N: 0.657), indicating that nitrogen-based nutrient salts fully functioned as electrolytic factors in the water. Regarding the correlations between organic matter components, compared to the correlation coefficients of BOD with COD (0.551) and with TOC (0.585), the correlation of COD with TOC showed a relatively higher coefficient of 0.895. This indicates that the variation in COD and TOC, and those in COD/TOC and BOD, did not correspond to each other, inferring that non-biodegradable organic matters were more distributed than organic (Gwak and Kim 2015). Overall, PCP and water temperature showed high correlations with SS and phosphorus-based nutrient salts, while COD and TOC showed a higher correlation coefficient than those of COD/TOC with BOD. Based on this, it is assumed that non-biodegradable organic matter was distributed in the study area.

3.3 Cluster analysis

Cluster analysis involves arranging objectives into clusters based on the similarities inside the clusters and dissimilarities of different clusters. The clusters were divided by their unique characteristics, and often informed the interpretation of the water quality data (Vega *et al.* 1998). CA was performed based on the analyzed water quality components and is represented by a dendrogram (Fig. 3). The study area was largely divided into three clusters (I, II, and III), while clusters I and II could be subdivided into four sub-clusters, Ia, Ib, IIa, and IIb, comprising five clusters in total including cluster III. Cluster III belonged to the upstream part of the study area (Sites 1 and 2), showing the lowest concentrations of most water quality components compared to other clusters, thus it was interpreted that this cluster had the lowest level of water quality pollution. Cluster II consisted of the upper (Sites 3–6, sub-cluster IIb) and middle (Sites 7–9, sub-cluster IIa) stream sites. Cluster II had a higher level of water quality pollution than cluster III, but had a lower level than cluster I. Cluster I consisted of sites in the middle (Sites 10–12, sub-cluster Ib) and lower (Sites 13–17, sub-cluster Ia) reaches. The average number of populations in the classified clusters is in the order of Ib > Ia > IIa > IIb > III, and land use is in the order of Ib > IIa > IIb > Ia > III for both agricultural and industrial area. As such, sub-cluster Ib represented sites with the largest population and the widest areas of agriculture and industry in the study area. The CA was performed and classified based on the water quality pollution levels and characteristics of the study area (Singh *et al.* 2005). According to the pollution levels, the relationship from highest to lowest was: cluster I (sub-clusters Ib > Ia) > cluster II (sub-clusters IIa > IIb) > cluster III. Further, a large amount of organic matter and nutrient salts was assumed to be present in sub-cluster Ib, likely affecting sub-cluster Ia in the lower reaches.

The CA in this study indicated land use impact on the Nakdong River and the location of sewage treatment facilities, which was reflective of previous classification of geographic groups based on the pollutants introduced by domestic and industrial wastewater and agricultural activities (Shrestha and Kazama 2007). According to previous studies, areas were classified into the upstream part, which had good quality of water, and areas near Deagu

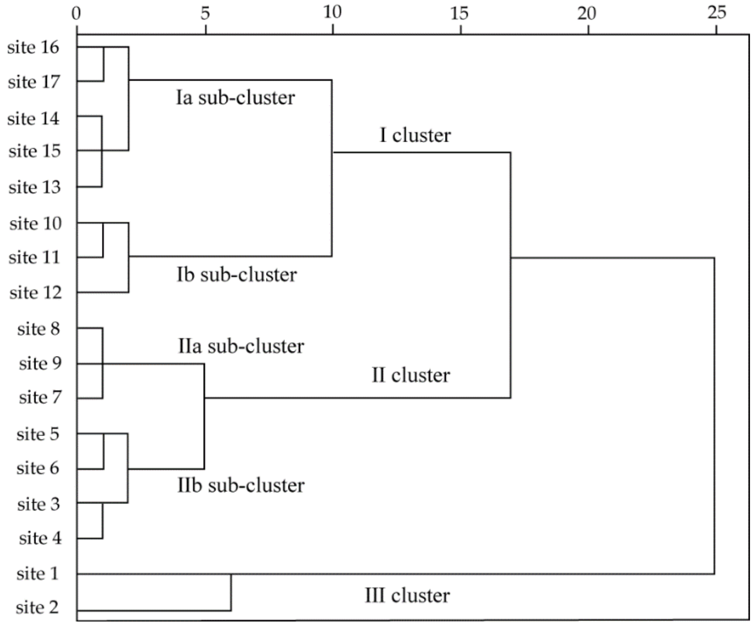


Fig. 3 Dendrogram of the monitoring sites using hierarchical cluster analysis in the study area

city, which were heavily affected by anthropogenic pollutants from cities and industrial complexes (Han *et al*. 2009). In addition, based on water quality pollution levels, the river had been divided into two groups: upstream, and the middle and lower reaches. The middle and lower reaches have been shown to have high pollution levels and to be strongly affected by domestic and industrial wastewater and effluent flowing into the Geumho-river (Jung *et al*. 2016a). In this study, clusters were classified on the basis of Site 10 located in Deagu city, according to various pollutants in the wastewater effluent discharged from urban, agricultural, and industrial areas concentrated near the site, which flow into the mainstream through the Geumho-river and show varying water quality characteristics.

3.4 Factor analysis/Principal component analysis

Factor analysis and principal component analysis is a powerful pattern recognition tool that combines many variables based on similar common dimensions and reduces them to a small number of factors in order to identify principal components (PCs) of environmental change. For this study, the method derived a unique pattern that was mutually independent of changes in water quality parameters in order to analyze the trend of the entire dataset (Simeonev *et al*. 2003). In addition, FA extracted varifactors (VFs) from loading factors, allowing them to be classified based on the degree of influence of each water quality parameter on the water quality characteristics. VF could also include unobservable, hypothetical, or latent variables, while each PC comprised a linear combination of observed water quality parameters (Wunderlin *et al*. 2001). In this study, FA/PCA was conducted on a total of six clusters (total cluster, sub-clusters Ia, Ib, IIa, and IIb, and cluster III), identifying PCs which had the greatest influence on factor impacts and water quality changes in each cluster. It is known that FA/PCA analysis is possible only when the KMO-test has a minimum value of 0.5, and Barlett's test has a p-value of less than 0.05 (Park *et al*. 2013, Kim and Kim 2017). In this study, the KMO-test had values of more than 0.7 in the total cluster and sub-clusters Ia, Ib, IIa, and IIb, and 0.633 in cluster III, while Barlett's test had values of 0.00 in all clusters, demonstrating the applicability of FA/PCA. Eigenvalues were also extracted, which provide a measure of the significance of the factor; the highest eigenvalue indicates the highest significance, with 1.0 or greater considered as significant (Kim and Mueller 1978, Pekey *et al*. 2004). Accordingly, in this study, VFs with an eigenvalue higher than 1.0 were extracted.

The FA results including VFs, factor loading, and total and cumulative variance are shown in Table 4. If the factor loading value of the water quality parameter exceeded 0.75, it was classified as "strong," values between 0.75 and 0.50 were "moderate," and values less than 0.5 were "weak" (Liu *et al*. 2003). Total of four VFs were extracted from the total cluster; VF1 had an eigenvalue of 4.445 and total variance of 29.630%, with T-P of 0.887 and PO_4-P of 0.791, which were classified as "strong", and had COD (0.735), TOC (0.699), SS (0.709), and PCP (0.665) which corresponded to "moderate". VF2 indicated an eigenvalue of 4.126, a total variance of 27.509%, with NO_3-N (0.910), T-N (0.875), WT (-0.839), DO (0.830), and EC (0.556) corresponding to strong. In addition, in cluster I, sub-clusters IIa and IIb, and cluster III, the factors extracted in VF1 and VF2 showed similar results to the total cluster. These results were similar to those of a previous study (Han *et al*. 2009), showing that anthropogenic pollutants such as organic matter and T-P corresponded to "strong" and had great influence on the water quality of the Nakdong River. In sub-cluster Ib, T-N (0.893), NO_3-N (0.944), DO (0.904), and WT (-0.901) of VF1 corresponded to strong, and in VF2, SS (0.892), T-P (0.863), PO_4-P (0.797), COD (0.721), PCP (0.639), and TOC (0.614) corresponded to both strong and moderate, showing different results from the other five

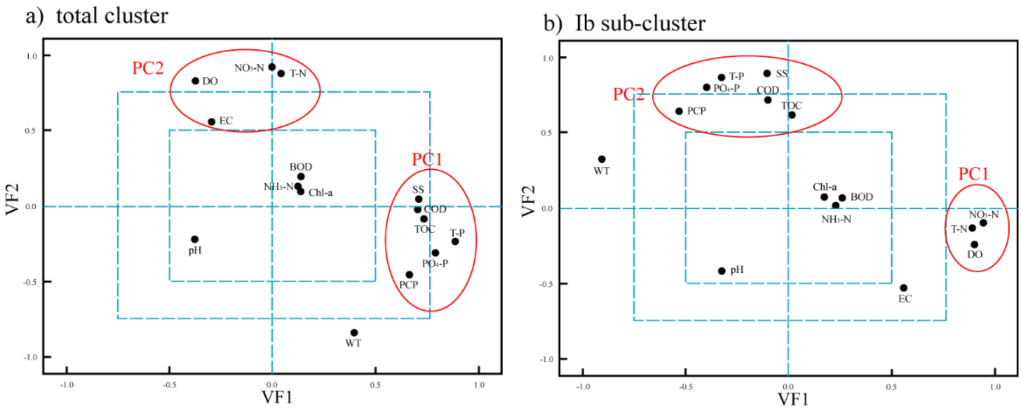
Table 4 Result of factor analysis (FA) for each cluster in the study area

a) total cluster (all sites)					**b) Ia sub-cluster**					**c) Ib sub-cluster**				
parameter	VF 1	VF 2	VF 3	VF 4	parameter	VF 1	VF 2	VF 3	VF 4	parameter	VF 1	VF 2	VF 3	VF 4
T-P	**0.887**	-0.235	0.014	0.160	T-P	**0.899**	-0.262	0.065	0.016	NO_3-N	**0.944**	-0.090	0.064	0.114
PO_4-P	**0.791**	-0.311	-0.255	0.129	PO_4-P	**0.868**	-0.294	-0.217	0.034	DO	**0.904**	-0.229	0.049	-0.005
COD	**0.735**	-0.084	0.567	0.135	SS	**0.740**	-0.141	0.311	-0.001	WT	**-0.901**	0.325	0.115	-0.084
SS	**0.709**	0.045	0.107	-0.322	PCP	**0.696**	-0.336	0.162	0.105	T-N	**0.893**	-0.130	0.144	0.256
TOC	**0.699**	-0.005	0.581	0.133	EC	-0.643	**0.530**	0.131	0.272	EC	**0.554**	-0.525	0.363	0.294
PCP	**0.665**	-0.457	-0.030	0.022	NO_3-N	-0.177	**0.926**	-0.033	0.027	SS	-0.103	**0.892**	0.093	-0.177
NO_3-N	0.005	**0.910**	-0.008	0.153	T-N	-0.208	**0.905**	-0.011	0.176	T-P	-0.326	**0.863**	-0.120	0.131
T-N	0.040	**0.875**	0.068	0.317	WT	0.357	**-0.872**	0.157	0.046	PO_4-P	-0.397	**0.797**	-0.305	0.108
WT	0.395	**-0.839**	0.100	0.068	DO	-0.376	**0.857**	0.139	-0.085	COD	-0.100	**0.721**	0.546	0.122
DO	-0.368	**0.830**	0.114	-0.143	Chl-a	0.146	0.122	**0.852**	-0.134	PCP	-0.532	**0.639**	0.036	-0.111
EC	-0.298	**0.556**	0.341	0.479	BOD	-0.002	0.228	**0.843**	0.110	TOC	0.015	**0.614**	**0.536**	0.365
Chl-a	0.134	0.109	**0.849**	-0.044	TOC	0.514	-0.209	**0.752**	0.009	Chl-a	0.168	0.080	**0.869**	-0.215
BOD	0.140	0.190	**0.834**	0.282	pH	-0.358	-0.231	**0.737**	-0.060	BOD	0.259	0.072	**0.766**	0.448
pH	-0.378	-0.221	**0.716**	-0.126	COD	0.597	-0.301	**0.662**	0.034	pH	-0.320	-0.415	**0.716**	-0.136
NH_3-N	0.126	0.130	0.052	**0.867**	NH_3-N	0.046	0.060	-0.046	**0.968**	NH_3-N	0.233	0.021	-0.053	**0.897**
Eigen value	4.445	4.126	2.732	1.463	Eigen value	4.121	3.988	3.220	1.100	Eigen value	4.445	4.126	2.732	1.463
%total Var.	29.630	27.509	18.214	9.756	%total Var.	27.476	26.651	21.469	7.334	% total Var.	29.630	27.509	18.214	9.756
Cul. %	29.630	57.139	75.353	85.109	Cul. %	27.476	54.127	75.596	82.930	Cul. %	29.630	57.139	75.353	85.109

d) IIa sub-cluster					**e) IIb sub-cluster**					**f) III cluster**				
parameter	VF1	VF2	VF3		parameter	VF1	VF2	VF3	VF4	parameter	VF1	VF2	VF3	VF4
COD	**0.874**	-0.099	0.228		T-P	**0.890**	-0.126	-0.297	-0.018	T-P	**0.875**	0.221	-0.078	0.043
SS	**0.858**	0.059	0.126		COD	**0.880**	-0.064	0.179	0.158	PO_4-P	**0.871**	0.098	-0.052	-0.214
T-P	**0.835**	-0.278	-0.246		PO_4-P	**0.818**	-0.048	-0.418	-0.057	COD	**0.867**	0.247	-0.089	0.086
PO_4-P	**0.810**	-0.316	-0.264		SS	**0.805**	0.124	0.053	0.089	SS	**0.826**	0.381	0.038	-0.228
TOC	**0.809**	-0.017	0.332		PCP	**0.792**	-0.334	-0.244	0.111	TOC	**0.758**	-0.026	-0.086	0.429
EC	-0.760	0.421	-0.049		TOC	**0.741**	0.007	0.361	0.185	PCP	**0.655**	0.524	0.120	-0.037
PCP	0.641	-0.524	-0.101		EC	**-0.729**	0.326	0.422	0.216	NO_3-N	0.386	-0.342	0.017	0.361
pH	-0.593	-0.199	0.526		NO_3-N	0.070	**0.942**	-0.089	-0.168	T-N	0.222	**0.867**	-0.249	0.205
NO_3-N	0.035	**0.949**	0.073		T-N	0.125	**0.933**	-0.082	-0.098	WT	-0.233	**-0.860**	0.245	-0.138
T-N	-0.024	**0.942**	-0.062		WT	0.512	**-0.809**	-0.089	-0.048	DO	-0.053	**-0.661**	0.378	0.151
WT	0.402	**-0.861**	-0.019		DO	-0.512	**0.762**	0.246	-0.030	EC	0.435	**0.651**	-0.029	-0.058
DO	-0.315	**0.854**	0.216		BOD	0.049	0.076	**0.868**	-0.080	Chl-a	-0.004	-0.211	**0.964**	-0.022
Chl-a	0.001	0.179	**0.804**		pH	-0.241	-0.155	**0.792**	-0.097	BOD	-0.073	-0.223	**0.952**	-0.019
BOD	-0.016	0.398	**0.651**		NH_3-N	0.504	-0.146	0.028	**0.705**	pH	-0.248	0.031	-0.012	**0.851**
NH_3-N	-0.163	0.163	**-0.609**		Chl-a	0.121	0.166	0.485	**-0.659**	NH_3-N	0.540	0.387	-0.013	**0.609**
Eigen value	5.137	4.156	2.095		Eigen value	5.466	3.326	2.365	1.119	Eigen value	4.765	3.253	2.143	1.606
%total Var.	34.247	27.705	13.970		%total Var.	36.440	22.172	15.769	7.457	%total Var.	31.766	21.686	14.288	10.708
Cul. %	34.247	61.952	75.922		Cul. %	36.440	58.612	74.381	81.838	Cul. %	31.766	53.452	67.740	78.447

clusters. In addition, Chl-a, BOD, and pH, which are related to eutrophication, belonged to VF3 in all clusters.

The factor loading values of the total cluster and sub-cluster Ib are presented in a scatter plot in order to identify PC1 and PC2, which had highest influences on the study area, and PC1 and PC2 in sub-cluster Ib with different VF1 and VF2 (Fig. 4). There is a difference in the principal components affecting cluster sub-cluster Ib compared to other clusters, and it can be estimated that there is a difference in the pollutant sources flowing into the sites


a) total cluster

b) Ib sub-cluster

Fig. 4 Scatter plot of loading for the total cluster (a), and sub-cluster Ib (b)

corresponding to sub-cluster Ib. PC1, which had the greatest influence on water quality characteristics in the study area, included phosphorus-based nutrient (T-P and PO_4-P), non-degradable organic matter (COD and TOC), PCP, and SS, while PC2 had T-N, NO_3-N, and DO. PC1 in sub-cluster Ib consisted of NO_3-N and T-N; compared to other clusters, nitrogen-based nutrient salts had the greatest effect on water quality characteristics, while phosphorus-based nutrient salts, organic matter, and SS, identified as PC1 in the total cluster and other clusters, were found to be PC2. It is known that nutrient salts, non-degradable organic matter, and SS, which have great influence on the water quality of the Nakdong River, are mostly introduced from urban and industrial areas near the sites, while phosphorus-based nutrient salts and SS have generally increased because of point source pollution and non-point source pollution, such as agricultural land and soil erosion (Venkatraman *et al.* 2014). In particular, domestic wastewater, especially that containing detergents, industrial effluents, and fertilizer run-off, contributes to elevated levels of phosphates in water (Iscen *et al.* 2008). Nitrogen-based nutrient salts, which corresponded to PC2 in the total cluster and PC1 in sub-cluster Ib in this study, are known to be an anthropogenic pollutant sources of domestic wastewater and agricultural activities (Singh *et al.* 2005). Fertilizers containing nitrogen compounds are used for agricultural activities, distributing a high concentration of nitrogen pollutants within the Nakdong River basin (Zhou *et al.* 2006). In addition, nitrogen-based nutrient salts, among pollutants inflowed through tributaries, generally function as PCs in areas where urban and industrial areas are concentrated (Yoon *et al.* 2019). Based on this, for this study it can be inferred that anthropogenic sources of pollution were discharged from urban, industrial, and agricultural areas near sub-cluster Ib. In particular, a high concentration of nitrogen pollutants was introduced through the Geumho-river, a large tributary joining Site 10. According to the results of FA/PCA, in the study area, phosphorus-based nutrient salts (T-P and PO_4-P) and non-degradable organic matter (COD and TOC) had the greatest effect on water quality characteristics. Compared to other clusters, sub-cluster Ib (Sites 10, 11, and 12) was considered to be relatively heavily affected by nitrogen-based nutrient salts (T-N, and NO_3-N).

3.5 Pollution index

Based on the FA/PCA results, the concentrations of nutrient salts and organic matter, which acted as major factors (PC1 and PC2) for variations of water quality characteristics in the study area, were calculated as PIs (Fig. 5). Specifically, the PIs were based on the mean concentrations of T-P and T-N (parameters representing phosphorus and nitrogen-based nutrient salts), COD and TOC (parameters indicating non-degradable organic matter), and SS. COD and TOC showed high pollution levels with averages of 1.45 and 1.44, respectively. In clusters II and III in the upstream part, the values generally recorded less than 1.5, while values exceeded 1.5 from Site 10 to the downstream sites, which belonged to sub-cluster Ib. COD and TOC, indicating non-biodegradable organic matter, were likely heavily influenced by the wastewater effluent discharged from urban and industrial areas. In particular, biodegradable organic matter can be processed and removed in sewage and wastewater treatment plants; however, non-biodegradable substances are known to remain in the effluent (Imai *et al.* 2002). At Geumho-river joining Site 10, there is a concentration of sewage and wastewater treatment plants, introducing a large amount of non-biodegradable organic compounds in the effluent into the mainstream of the Nakdong River (Jung *et al.* 2015). In addition, there is a study that the Geumho-river's water quality steeply decreased along with the inflow of tributaries' flow as well as effluents from the wastewater treatment plants in city of Gyeongsan and Daegu (Bae 2020). Therefore, in order to reduce the pollution caused by non-biodegradable organic compounds at the sites in the middle and lower reaches, the effluent from the treatment facilities near Geumho-river should be properly managed. SS recorded an average PI of 0.28, indicating no pollution, and showing a relatively higher value in the lower reaches than in the upstream part. T-P had an average PI of 0.73, indicating an overall pollution-free environment; however, cluster Ib showed a higher degree of contamination with a value close to 1 (> 0.9). The changes in T-P concentration are influence by changes in the discharge loads from pollution sources in watershed (Jung *et al.* 2019). Therefore, the PI increased according to the increased in the discharge load from the pollution sources of the Geumho-river joining

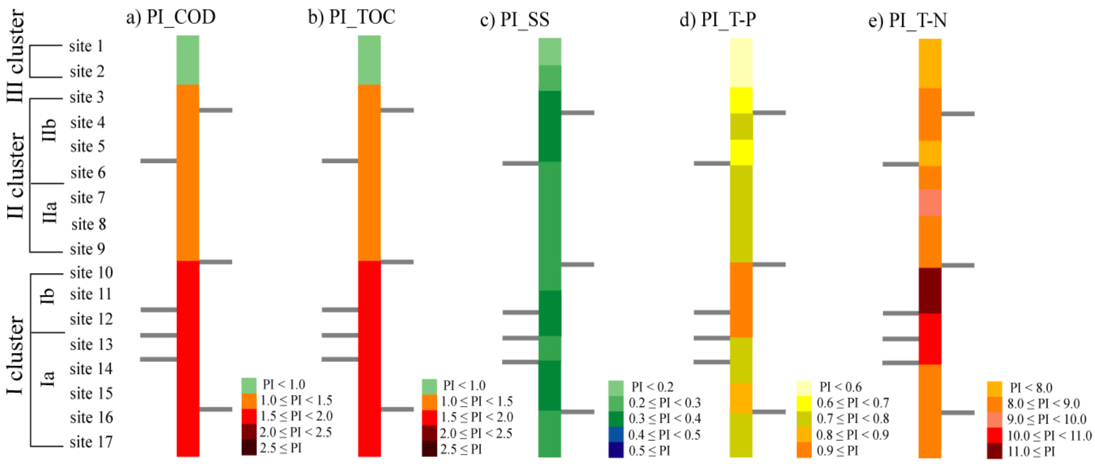


Fig. 5 Pollution indexes of (a) COD; (b) TOC; (c) SS; (d) T-P; and (e) T-N

the cluster Ib, and there is a high possibility of water pollution by nitrogen-based nutrient salts for those sites. T-N showed the highest PI with an average of 13.54, which may be the result of applying the lake standard to the T-N standard concentration. Therefore, it is necessary to set the T-N river standard concentration in order to determine the pollution level of T-N in rivers with higher accuracy and reliability. T-N, similar to other water quality parameters, tended to increase in the middle and lower reaches, especially in sub-cluster Ib, showing a high value exceeding 15. These results, along with the outcomes from FA/PCA, indicate that the variations in water quality characteristics by nitrogen-based nutrient salts were higher than those in other clusters, indicating that worsening water quality from nitrogen contaminants was serious in sub-cluster Ib. The PI of five water quality parameters was generally higher in the middle and lower reaches than in the upstream part, and in particular, sub-cluster Ib showed the highest pollution levels. It was found that spatial variations in the integrated PI were consistent with previously determined spatial variations of the level of urbanization (Wang *et al*. 2008). For Site 10 in sub-cluster Ib, which had the highest PI value and was consistent with the highest degree of urbanization due to population density, it is interpreted that the influx of anthropogenic pollutants due to human activities significantly influenced the water quality. Within the study area, a higher number of larger tributaries such as Geumho-river, Hwang-river, Nam-river, and Milyang-river join the middle and lower reaches as compared to the upstream part of the mainstem Nakdong River. Among them, Geumho-river is located within a larger city and has a higher population as compared to other tributaries; further, coupled with its higher concentration of agricultural and industrial areas, it is known that pollutants such as organic matters and nutrient salts, which are discharged from sewage treatment plants and domestic wastewater, flow into the mainstream of the Nakdong River from this location, affecting its downstream reaches (Jung *et al*. 2016b, Lee and Kim 2017). In this study, the concentrations of non-degradable organic compounds and nutrient salts increased sharply at Site 10 compared to Site 9, leading to a tendency to worsen water pollution. Thus, it is inferred that non-degradable organic matter, and nitrogen- and phosphorus-based nutrient salts, which were distributed in sewage and wastewater discharged from urban, agricultural, and industrial areas concentrated near Site 10, flowed into the mainstream through the Geumho-river, worsening water quality pollution in the middle and lower reaches.

4. Conclusions

In order to evaluate the characteristics and spatial variation of water quality (PCP, WT, DO, EC, BOD, COD, TOC, T-N, NH_3-N, NO_3-N, T-P, PO_4-P, SS, and Chl-a) in the Nakdong River, a total of 17 sites were surveyed for three years from 2017–2019. Four statistical analysis techniques (pearson's correlation analysis, CA, and FA/PCA) were applied based on the analyzed water quality parameters to calculate the PI of main contributing factors. Generally, the study area showed that concentrations of phosphorus- and nitrogen-based nutrient salts and organic matter increased at Site 10. The study area was divided into a total of five clusters based on the characteristics of water quality pollution, which show descending order of pollution as: cluster I (sub-clusters Ib > Ia) > cluster II (sub-clusters IIa > IIb) > cluster III. Phosphorus-based nutrient salts (T-P and PO_4-P), SS, and non-degradable organic matter (COD and TOC) corresponded to VF1 and VF2, which were found to be PCs with a strong impact on water quality. In addition, it was found that sub-cluster Ib was relatively heavily affected by nitrogen-based nutrient salts of nitrogen (NO_3-N and TN) compared to other clusters. Water pollution (PI) shows deterioration by non-degradable organic matter, nitrogen- and phosphorus-based nutrient salts in the middle and lower reaches as compared to upstream. Based on these results, anthropogenic pollution sources such as sewage and wastewater discharged from urban, agricultural, and industrial areas concentrated near Site 10 likely flowed into the mainstream through the Geumho-river and deteriorated the water quality in the middle and lower reaches. Therefore, to improve overall water quality in the Nakdong River, pollution reduction and purification measures for nutrient salts and non-degradable organic matters should be preferentially implemented for the site of Geumho-river joining Site 10 in sub-cluster Ib. In other words, in order to improve water pollution in the main stream of river, which is used as a source of drinking water, it is necessary to remove the source of pollution by effectively treating wastewater from cities, agriculture, and industrial complexes flowing in from tributaries.

Acknowledgments

This research was supported by a grant from the Nakdong River Environment Research Center, National Institute of Environmental Research of Korea (NIER-2021-01-01-042).

References

Ahn, K. and Yang W.M. (2007), “Water quality characteristics in Geum river watershed”, *Kor. J. Limnol.*, **40**(1), 110-120.

Alves, D.D., Riegel, R.P., de Quevedo, D.M., Osorio, D.M.M, da Costa, G.M., do Nascimento, C.A. and Teloken, F. (2018), “Seasonal assessment apportionment of surface water pollution using multivariate statistical methods: Sinos river, south Brazil”, *Environ. Monit. Assess.*, **190**(7), 1-12. https://doi.org/10.10067/s10661-018-6759-3.

Bae, H.K. (2020), “The effect of combined sewer overflow on river’s water quality”, *Membr. Water Treat.*, **11**(1), 49-57. http://doi.org/10.12898/mwt.2020.11.1.049.

Chang, N.B. (2005), “Sustainable water resources management under uncertainty”, *Stoch. Environ. Res. Risk Assess.*, **19**(2), 97-98. https://doi.org/10.1007/s00477-004-0217-1.

Choi, H.G., Han, G.Y. and Choi, S.Y. (2011), “A stochastic analysis of the quality with discharge variation in upper Nakdong river basin”, *J. Environ. Imp. Assess.*, **20**(6), 833-843. https://doi.org/10.14249/eia.2011.20.6.833.

Choi, K.Y., Im, T.H., Lee, J.W. and Choi, S.Y. (2012), “Evaluation of water quality characteristics in the Nakdong river using statistical analysis”, *J. Kor. Soc. Water Environ.*, **45**(11), 1157-1168. https://doi.org/10.3741/JKWRA.2012.45.11.1157.

Dutta, S., Dwivedi, A. and Kumar, M.S. (2018), “Use of water quality index and multivariate statistical techniques for the assessment of spatial variations in water quality of a small river”, *Environ. Monit. Assess.*, **190**(12), 1-17. https://doi.org/10.1007/s10661-018-7100-x.

Gwak, B.R. and Kim, I.K. (2015), “The evaluation of water quality using a multivariate analysis in Changyeong-Haman weir section”, *J. Kor. Soc. Water Environ.*, **29**(6), 625-632. https://doi.org/10.11001/jksww.2015.29.6.625.

Fan, X., Cui, B., Zhao, Z. and Zhang, H. (2010), “Assessment of river water quality in Pearl river delta using multivariate statistical techniques”, *Proc. Environ. Sci.*, **2**, 1220-1234. https://doi.org/10.1016/j.proenv.2010.10.133.

Han, S.H., Kim, E.S. and Kim, S.D. (2009), “The water quality management in the Nakdong river watershed using multivariate statistical techniques”, *Kor. J. Civil Eng.*, **13**(2), 97-105. https:// doi.org/10.1007/s12205-009-0097-5.

Hussain, M., Ahmed, S.M. and Abderrahman, W. (2008), “Cluster analysis and quality assessment of logged water at an irrigation project, eastern Saudi Arabia”, *J. Environ. Manage.*, **86**, 279-307. https://doi.org/10.1016/j.jenvman.2006.12.007.

Imai, A., Fukushima, T., Matsushige, K, Kim, Y.H. and Choi, K. (2002), “Characterization of dissolved organic matter in effluents from wastewater treatment plants”, *Water Res.*, **34**(6), 859-870. https://doi.org/10.1016/S0043-1354(01)00283-4.

Iscen, C.F., Emiroglu, O., Ilhan, S., Arslan, N., Yilmaz, V. and Ahiska, S. (2008), “Application of multivariate statistical techniques in the assessment of surface water quality in Uluabat lake, Turkey”, *Environ. Monit. Assess.*, **144**, 269-276. https://doi.org/10.1007/s10661-007-9989-3.

Joung, S.H., Park, H.K., Lee, H.J. and Lee, S.H. (2013), “Effect of climate change for diatom bloom at winter and spring season in Mulgeum station of the Nakdong river”, *J. Kor. Soc. Water Environ.*, **29**(2), 155-164.

Jung, K.Y., Ahn, J.M., Lee, K.L., Lee, I.J., Yu, J.J., Cheon, S.U., Kim, K.S. and Han, G.Y. (2015), “Temporal and spatial analysis of non-biodegradable organic pollutants in the Geumho river system”, *J. Environ. Sci. Int.*, **24**(11), 1343-1362. http://doi.org/10.5322.JESI.2015.24.11.1343.

Jung, K.Y., Lee, K.L., Im, T.H., Lee, I.J., Kim, S., Han, K.Y. and Ahn, J.M. (2016a), “Evaluation of water quality for the Nakdong river watershed using multivariate analysis”, *Environ. Tech. Inn.*, **5**, 67-82. https://dx.doi.org/10.1016/j.eti.2015.12.001.

Jung, K.Y., Ahn, J.M., Kim, S., Lee, I.J. and Yang, D.S. (2016b), “Evaluation of water quality characteristics and water quality improvement grade classification of Geumho river tributaries”, *J. Environ. Sci. Int.*, **25**, 767-787. http://doi.org/10.5322/JESI.2016.25.6.767.

Jung, S.Y. and Kim, I.K. (2017), “Analysis of water quality factor and correlation between water quality and Chl-a in middle and downstream weir section of Nakdong river”, *J. Kor. Soc. Environ. Eng.*, **39**, 89-96. http://doi.org/10.4491/KSEE.2017.39.2.89.

Jung, K.Y., Ahn, J.M., Cho, S.H., Lee, Y.J., Han, K.Y., Shin, D.S. and Kim, K.H. (2019), “Evaluation of long-term water quality management policy effect using nonparametric statistical methods”, *Membr. Water Treat.*, **10**(5), 339-352. https://doi.org/10.12989/mwt.2019.10.5.339.

Kappers, F.I. (1980), “The cynobaterium Microcystics Aeruginosa kg. and the nitrogen cycle of the hypertrophic lake Brielle (the Netherlands)”, *Hyper. Eco.*, **2**, 37-43. https://doi.org/10.1007/978-94-009-9203-05.

Kim, G.H., Jung, K.Y., Yoon, J.S. and Cheon, S.U. (2013), “Temporal and spatial analysis of water quality data observed in lower watershed of Nam river Dam”, *J. Kor. Soc. Hazard Mitig.*, **13**(6), 429-437. http://dx.doi.org/10.9798/KOSHAM.2013.13.6.429.

Kim, J.I., Choi, J.W. and An, K.G. (2014), “Spatial and temporal variations of water quality in an urban Miho stream and some influences of the tributaries on the water quality”, *J. Environ. Sci. Int.*, **23**(3), 433-455. https://doi.org/10.5322/JESI.2014.23.3.433.

Kim, J.O. and Mueller, C.W. (1978), *Introduction to Factor Analysis: What It Is and How Do It, Quantitative Applications In The Social Sciences,* Sage University Press, Newbury Park, U.S.A.

Kim, M.S., Chung, Y.R., Suh, E.H. and Song, W.S. (2002), “Eutrophication of Nakdong river and statistical analysis of environmental factors”, *Algae,* **17**, 105-115. http://doi.org/10.4490/Algae.2002.17.2.105.

Kim, S.M., Kim, S.M. and Park, T.Y. (2010), “Spatial analysis of BOD data in Namgang-Dam watershed for TMDLs”, *J. Kor. Soc. Agric. Eng.*, **52**(3), 81-88.

Kim, S.R. and Kim, S.M. (2017), “Spatial water quality analysis of main stream of Nakdong river considering the inflow of tributaries”, *J. Kor. Soc. Water. Environ.*, **33**(6), 640-649. https://doi.org/10.15681/KSWIE2017.33.6.640.

Kolovos, A., Christakos, G., Serre, M.L. and Miller, C.T. (2002), “Computational Bayesian maximum entropy solution of a stochastic advection reaction equation in the light of site-specific information", *Water Resour. Res.*, **38**(12), 1318-1334. https://doi.org/10.1029/2001WR000743.

Kowalkowski, T., Zbytniewski, R., Szpejna, J. and Buszewski, B. (2006), "Application of chemometrics in river water classification", *Water Res.*, **40**, 744-752. http://doi.org/10.1016/j.waters.2005.11.042.

Kuppusamy, M.R. and Giridhar, V.V. (2006), "Factor analysis of water quality characteristics including trace metal speciation in the coastal environmental system of Chennai Ennore", *Environ. Int.*, **32**, 174-179. https://doi.org/10.1016/j.envint.2005.08.008.

Kumari, M., Tripathi, S., Pathak, V. and Tripathi, B.D. (2013), "Chemometric characterization of river water quality", *Environ. Monit. Assess.*, **185**, 3081-3092. https://doi.org/10.1007/s10661-012-2774-y.

Lee, H.J., Park, H.K., Lee, J.H., Park, A.R. and Cheon, S.U. (2016), "Coliform pollution status of Nakdong river and tributaries", *J. Kor. Soc. Water Waste*, **32**(3), 271-280. https://doi.org/10.15681/KSWE.2016.32.3.271.

Lee, S.M. and Kim, I.K. (2017), "Analysis of correlation between cyanobacterial population and water quality factors in the middle and downstream region of Nakdong river", *J. Kor. Soc. Water Waste,* **31**, 93-101. http://doi.org/10.11001/jksww.2017.31.1.093.

Liu, C.W., Lin. K.H., and Kuo, Y.M. (2003), "Application of factor analysis in the assessment of ground water quality in a blackfoot disease area in Taiwan", *Sci. Total Environ.*, **313**(1), 77-89. https://doi.org/10.1016/s0048-9697(02)00683-6.

Mckenna, J.E. (2003), "An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis", *Environ. Model Soft.*, **18**(3), 205-220. https://doi.org/10.1016/S1364-8152(02)00094-4.

Ministry of Environment (MOE) (2012), *Official Test Methods of Water Quality*, MOE, Sejong, Korea.

Ma, X., Wang, L., Yang, H., Li, N. and Gong, C. (2020), "Spatiotemporal analysis of water quality using multivariate statistical techniques and the water quality identification index for the Qinhuai river basin, China", *Warter*, **12**(10), 2764-2783. https://doi.org/10.3390/w12102764.

Na, E.H., Park, S.Y., Kim, J.H., Im, S.S. and Kim, K.H. (2015), "A study on spatial and temporal patterns of water quality the middle area of the Nakdong river", *J. Kor. Soc. Water Environ.*, **31**(6), 723-731. http://doi.org/10.15681/KSWE.2015.31.6.723.

Na, S.M., Kwon, H.G., Shin, S.M., Son, Y.G., Shin, D.S. and Im, T.H. (2016), "A study on seasonal pollutant distribution characteristics of contaminated tributaries in Nakdong river basin", *J. Wetland Res.*, **18**, 301-312. http://doi.org/10.17663/JWR.2016.18.3.301.

Otto, M. (1998), *Multivariate Methods In: Analytical Chemistry*. Wiley/VCH, Weinheim, Getmany.

Ouyang, Y., Nkedi-kizza, P., Wu, Q.T., Shinde, D. and Huang, C.H. (2006), "Assessment of seasonal variations in surface water quality", *Water Res.*, **39**, 2621-2635. https://doi.org/10.1016/j.water.2006.08.030.

Park, J.H., Jung, J.W., Kim, D.Y., Kim, K.S., Han, S.W., Kim, H.O. and Lim, B.J. (2013), "Water quality characteristics of the major tributaries in Yeongsan and Sumjin river basin using statistical analysis", *J. Environ. Imp. Assess.*, **22**(2), 171-181. https://doi.org/10.14249/eja.2013.22.2.171.

Pekey, H., Karakas, D. and Bakoglu, M. (2004), "Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses", *Mar. Pollut. Bull.*, **49**(9-10), 809-818. https://doi.org/10.1016/j.marpolbul.2004.06.029.

Razmkhah, H., Abrishamchi, A., Torkian, A. (2010), "Evaluation of spatial and temporal variation in water quality by pattern recognition techniques: A case study on Jajrood river (Tehran, Iran)", *J. Environ. Manage.*, **91**, 852-860. https://doi.org/10.1016/j.jenvman.2009.11.001.

Ren, W., Zhong, Y., Meligrana, L., Anderson, B., Watt, W.E., Chen, J. and Leung, H.L. (2003), "Urbanization, land use, and water quality in Shanghai 1947-1996", *Environ. Int.*, **29**(5), 649-659. https://doi.org/S0160-4120(03)00051-5.

Seo, J.K., Lee, H.K. and Chung, I.K. (2010), "The community dynamics of microbial food web during algal blooms by Stephanodiscus spp. in downstream of Nakdong river", *Kor. J. Environ. Bio.*, **28**(3), 172-178.

Seo, M.J., Lee, H.J., and Kim, Y.S. (2019), "Relationship between coliform bacteria and water quality factors at weir stations in the Nakdong river, South Korea", *Warter*, **11**(6), 1171-1187. https://doi.org/10.3390/w11061171.

Shrestha, S. and Kazama, F. (2007), "Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan", *Environ. Model Soft.*, **22**, 464-475. https://doi.org/10.1016/j.envsoft.2006.02.001.

Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M. and Kouimatzis, T. (2003), "Assessment of the surface water quality in Northern Greece", *Water Res.*, **37**, 4119-4124. https://doi.org/10.1016/S0043-1354(03)00398-1.

Singh, K.P., Malik, A., Mohan, D. and Sinha, S. (2005), "Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques – a case study", *Analy. Chim. Acta.*, **538**, 355-374. https://doi.org/10.1016/j.aca.2005.02.006.

Suh, M.J., Cho, C.D., Im, T.H., Kim, S.H., Yoon, H.J., Kim, Y.S. and Kim, G.H. (2019), "Statistical analysis of the spatio-temporal water quality characteristics of the Nakdong river", *J. Environ. Sci. Int.*, **28**(3), 303-320. https://doi.org/10.5322/JESI.2019.258.3.303.

Su, S., Li, D., Zhang, Q., Xia, R., Huang, F. and Wu, J. (2011), "Temporal trend and source apportionment of water pollution in different functional zones of Qiantang river, China", *Water Res.*, **45**(4), 1781-1795. https://doi.org/10.1016/j.waters.2010.11.030.

Vega, M., Pardo, R., Barrado, E., Deban, 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/10.10161/S0043-1354(98)00138-9.

Venkatraman, S., Chung, S.Y., Lee, S.Y. and Park, N. (2014), "Assessment of river water quality via environmentric multivariate statistical tools and water quality index: A case study of Nakdong river basin, Korea", *Carpa. J. Ear. Environ. Sci.*, **9**(2), 125-132.

Wang, H.Y. (2002), "Assessment and prediction of overall environmental quality of Zhuzhou city, human province, China", *J. Environ. Manag.*, **66**(3), 329-340. https://doi.org/10.1006/jema.2002.0590.

Wang, J., Da, L., Song, K. and Li, B. (2008), "Temporal variations of surface water quality in urban, suburban, and rural areas during rapid urbanization in Shanghai, China", *Environ. Poll.*, **152**(2), 387-393. https://doi.org/10.1016/j.envpol.2007.06.050.

Wunderlin, D.A., Daiz, M.P., Ame, M.V., Pesce, S.F., Hued, A.C. and Bistoni, M.A. (2001), "Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality, a case study: Suquira river basin (Cordoba-Argentina)", *Water Res.*, **35**, 1894-2881. https://doi.org/10.1016/s0043-1354(00)00592-3.

Xian, G., Crane, M. and Junshan, S. (2007), "An analysis of urban development and its environmental impact on the Tampa bay watershed", *J. Environ. Manag.*, **85**(4), 965-976. https://doi.org/10.1016/j.jenvman.2006.11.012.

Yang, L., Mei, K. Liu, X., Wu, L., Zhang, M., Xu, J. and Wang, F. (2013). "Spatial distribution and source apportionment of water pollution in different administrative zones of Wen-Rui-Tang (WRT) river watershed, China”, *Environ. Sci. Pollut. Res.*, **20**, 5341-5352. https://doi.org/10.1007/s11356-013-1536-x.

Yang, Y.H., Zhou, F. Guo, H.C., Sheng, H., Liu, H., Dao, X. and He, C.J. (2010). “Analysis of spatial and temporal water pollution patterns in lake Dianchi using multivariate statistical methods”, *J. Environ. Manag.*, **85**(4), 965-976. https://doi.org/10.1016/j.jenvman.2006.11.012.

Yoon, H.Y., Kim, J.H., Chae, M.H., Cho, Y.H. and Cheon, S.U. (2019), “Assessment of water quality in the Miho stream using multivariate statistics”, *J. Environ. Impact. Assess.*, **28**(4), 373-386 https://doi.org/10.14249/eia.2019.28.4.373.

Zeng, X. and Rasmussen, T.C. (2005), “Multivariate statistical characterization of water quality in lake Lanier, Georgia, USA”, *J. Environ. Qual.*, **34**, 1980-1991. https://doi.org/10.2134/jeq2004.0337.

Zhang, Y., Guo, F., Meng, W. and Wang, X.Q. (2009), “Water quality assessment and source identification of Daliao river basin using multivariate statistical methods”, *Environ. Monit. Assess.*, **170**, 407-416. https://doi.org/10.1007/s10661-009-1242-9.

Zhou, F., Liu, Y., Guo, H.C. (2006), “Application of multivariate statistical methods to water quality assessment of the watercourses in Northwestern New Territories, Hong Kong”, *Environ. Monit. Assess.*, **132**, 1-13. https://doi.org/10.1007/s10661-006-9497-x.

JK