

Article

Identification of High-Priority Tributaries for Water Quality Management in Nakdong River Using Neural Networks and Grade Classification

Kang-Young Jung ¹, Sohyun Cho ¹, Seong-Yun Hwang ¹ , Yeongjae Lee ¹, Kyunghyun Kim ² 
and Eun Hye Na ^{1,*} 

¹ Yeongsan River Environment Research Center, National Institute of Environmental Research, Gwangju 61011, Korea; happy3313@korea.kr (K.-Y.J.); shcho35@korea.kr (S.C.); hsyliark@korea.kr (S.-Y.H.); nier@korea.kr (Y.L.)

² Watershed Pollution Load Management Research Division, National Institute of Environmental Research, Incheon 22689, Korea; Matthias@korea.kr

* Correspondence: eunye@korea.kr; Tel.: +82-62-970-2181

Received: 1 October 2020; Accepted: 28 October 2020; Published: 3 November 2020



Abstract: To determine the high-priority tributaries that require water quality improvement in the Nakdong River, which is an important drinking water resource for southeastern Korea, data collected at 28 tributaries between 2013 and 2017 were analyzed. To analyze the water quality characteristics of the tributary streams, principal component analysis and factor analysis were performed. COD (chemical oxygen demand), TOC (total organic carbon), TP (total phosphorus), SS (suspended solids), and BOD (biochemical oxygen demand) were classified as the primary factors. In the self-organizing maps analysis using the unsupervised learning neural network model, the first factor showed a highly relevant pattern. To perform the grade classification, 11 parameters were selected. Six parameters are concentrations of the main parameters for the water quality standard assessment in South Korea. We added the pollution load densities for the selected five primary factors. Joochungang showed the highest pollution load density despite its small watershed area. According to the results of the grade classification method, Joochungang, Topyeongcheon, Hwapocheon, Chacheon, Gwangyeocheon, and Geumhogang were selected as tributaries requiring high-priority water quality management measures. From this study, it was concluded that neural network models and grade classification methods could be utilized to identify the high-priority tributaries for more directed and effective water quality management.

Keywords: self-organizing maps; neural network model; grade classification; Nakdong river; tributary; management priority

1. Introduction

River water quality is highly affected by climatic conditions, atmospheric inputs, and anthropogenic inputs such as domestic sewage, industrial wastewater, and runoff from urban and agricultural areas [1]. A river system comprises both the mainstream and the tributaries, and the latter carry pollutants discharged from natural and anthropogenic sources to the mainstream of the river. Thus, for water quality improvement of the river, the water quality of its tributaries should be considered first [2,3]. Higher-quality water resources can be secured if the pollutants associated with human activities in the tributaries are managed before they flow into the mainstream [4].

The monitoring data of water quality and flow rate are important for identifying the behavior of pollutants in the watershed and for establishing management policies to improve the water quality in polluted rivers [5,6]. In previous studies, various classification techniques based on monitoring data

have been attempted to provide the quantitative and qualitative information for effective water quality management. Yan et al. [7] developed a model based on an adaptive neuro fuzzy inference system to assess and classify the river water quality status. Cho et al. [8], Na et al. [9], and Jung et al. [10] implemented the tributary grouping method based on the river water quality standard and flow regime. The tributaries with pollutant concentrations and flow rates greater than the reference values were classified as Group I, which require water quality management policies. Recently, multivariate statistical techniques such as cluster analysis (CA), principal component analysis (PCA), and discriminant analysis (DA) have been applied for evaluation and classification of river water [1,3,4,11–13]. Although the grouping of data (monitoring stations or parameters) is a very useful approach for obtaining better information about the water quality status of the river system, it is insufficient to rank management priorities in the same groups.

In this study, the grade classification method linked to the neural network model was proposed to identify the high-priority tributaries for water quality management of the Nakdong River in South Korea. This study focused on improving classification methods. To evaluate and classify the water quality and flow rate characteristics of the tributaries more effectively, we implemented multivariate statistical techniques such as principal component analysis, factor analysis and cluster analysis, and compared the results using the self-organizing map (SOM) based on artificial neural networks. In addition, we proposed measures to prioritize tributaries using a grade classification method considering the various factors that affect river water quality, such as flow rate, water quality parameters, delivery load, pollution load density, and land use.

2. Materials and Methods

2.1. Study Area

The Nakdong River, which is the second largest river system in South Korea, originates in the mountain area of the central-eastern Korean Peninsula and flows south to the South Sea through the estuary dam. The target watershed area is approximately 23,817 km², the channel length is 521.5 km, the watershed circumference is 1097.13 km, the watershed average width is 46.03 km, the watershed average elevation is 291.2 m, and the watershed slope is 32.26%. The major land use types are 67.5% forest and 23.5% agricultural area [14–19]. In this river watershed, there are major cities, including Daegu and Busan, and large-scale industrial complexes in the middle and downstream watersheds.

The average annual precipitation for 10 years from 2004 to 2013 recorded in the Nakdong River watershed is 1180.7 mm, and the precipitation from July to September accounts for 54.0% of total precipitation. The Nakdong River is an important water resource in the watershed. The total available water resources are approximately 6.2 billion tons, which are used for residential and industrial water (21.6%), agricultural water (51.0%), and river maintenance water (27.4%) [20]. Currently, the Nakdong River system supplies drinking water for more than 13 million people. The drinking water intake stations for the Pusan metropolitan area, which is the second largest city in South Korea, are located in the downstream area of this river. However, high anthropogenic pressure in the form of domestic sewage, industrial wastewater, and agricultural runoff has resulted in a serious deterioration of water quality in the downstream area of the Nakdong River for the last several decades. Thus, much effort is required to manage water quality in the watershed [21].

In this study, 28 major tributaries of the Nakdong River flowing into the mainstream were selected. The water quality monitoring sites for this study are located at the end of each tributary that join the Nakdong River mainstream (Figure 1).

In Figure 1, seven artificial dams are built upstream of the main stream and tributary stream. These dams affect the water quality, flow, and ecology of downstream rivers. In previous studies, it was also confirmed that artificial dams affect physical, biological and chemical changes in downstream rivers [22–26].

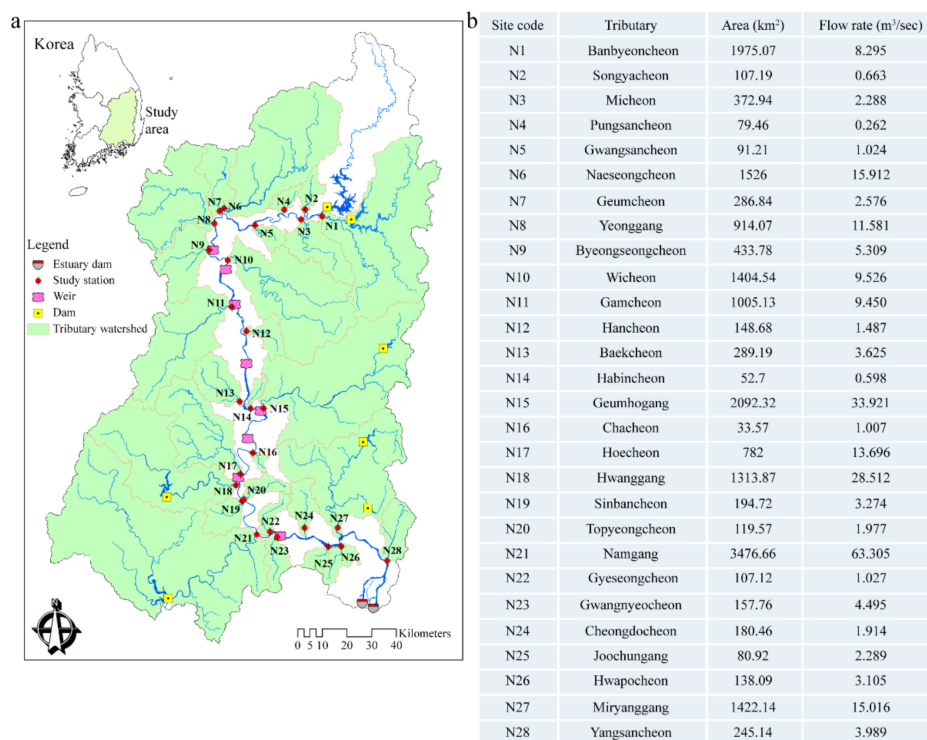


Figure 1. (a) Monitoring sites in Nakdong River tributaries and (b) physical characteristics.

2.2. Data Collection

The data used in this study were obtained from Water Environment Information System operated by the National Institute of Environmental Research, the Korean Ministry of Environment (<http://water.nier.go.kr>). The month and annual mean values collected at 8-day intervals for 5 years from January 2013 to December 2017 at 28 monitoring sites were used. Eleven parameters were monitored: hydrogen ion concentration (pH), dissolved oxygen concentration (DO), electrical conductivity (EC), water temperature (WT), biochemical oxygen demand (BOD), chemical oxygen demand (COD), total organic carbon (TOC), total nitrogen (TN), total phosphorous (TP), suspended solids (SS), and flow rate.

2.3. Data Analysis

The delivery load is the amount of pollutants that are generated in the watershed, discharged directly to the river with or without treatment, and reach the end of the drainage area undergoing decomposition by physical, chemical, and biological processes. In this study, the delivery loads at 28 tributaries were analyzed using Equation (1), and the pollution contribution rate (%) was calculated to determine the impact on the Nakdong River; it was also used for grade classification.

$$\text{Delivery load (kg/d)} = \text{Flow rate (m}^3\text{/s)} \times \text{Water quality concentration (mg/L)} \times 86,400 \text{ (s/d)} \times 10^{-3} \text{ (kg/mg)} \times 10^{-3} \text{ (L/m}^3\text{)} \quad (1)$$

Pollution load density is a factor that directly influences river water quality and is a method to verify the pollution potential of the water quality and the degree of pollution load according to the watershed area without considering the environmental characteristics of the watershed. Yoon et al. [27] reported that the water pollution level increases with the pollution load per watershed area. In the present study, the pollution potential of each tributary was examined based on the pollution load density calculated by dividing the delivery load (kg/d) by the watershed area (km²); it was also used to select the watersheds that needed to be prioritized.

To calculate the pollution load density, the watershed areas of 28 tributaries were required. To determine the watershed area in this study, the target watersheds were delineated using the ArcSWAT model based on the DEM (Digital Elevation Model) data of 30 m × 30 m grids created using a 1:25,000 digital map.

Land use was classified into nine categories (urban, industrial, field, paddy, forest, grass, bare land, wetland, and water; Supplementary Materials, Figure S1). Land use analysis was performed using a middle-category land cover map provided by the Korean Ministry of Environment.

2.4. Statistical Analysis

Water quality parameters were analyzed by performing principal component analysis, factor analysis, and cluster analysis using R software (Version 3.6.1) [28,29]. Principal component analysis calculates the principal components and condenses data so that the complex structure between correlated variables can be explained easily while minimizing the loss of information [30]. Factor analysis combines highly correlated variables into one homogeneous factor around the commonality of the variables [15,31]. Cluster analysis classifies the observational data into groups with high similarity. The cluster analysis method uses a hierarchical clustering method in which clustering occurs sequentially. It also examines the similarity of individuals belonging to the same cluster and differences between individuals belonging to different clusters [32,33]. Observed data are subjected to a Z-scale standardization process with a mean of 0 and a standard deviation of 1. This process reduces the influence of the differences in variance among water quality items and increases the explanatory power of factors by adjusting the size of variables [34,35].

In this study, to simplify the factor analysis method, the commonly used Varimax orthogonal rotation method was used for factor axis rotation. The Euclidean square distance method was applied to measure the distance between clusters, and the Ward method was used to minimize information loss between clusters [1,36,37]. Current research mostly defines the relationships between factors using analytical techniques based on artificial neural networks. In this study, the results of factor analysis and cluster analysis were compared using the self-organization map (SOM), which is a type of unsupervised machine learning neural network model. SOM is used in a variety of fields for analyzing nonlinear relationships, processing a wide range of data and irregular data, interpreting information on multiple variables, and visual image effects [38–40].

2.5. Grade Classification Method

In this study, eleven parameters were selected for grade classification. Six parameters (concentrations of BOD, COD, TOC, TN, TP, and SS) represent the main parameters for the water quality standard assessment in South Korea (www.me.go.kr). We added the pollution load densities of five water quality parameters (BOD, COD, TOC, TP, and SS), which we classified as the primary factors of water quality by conducting a principal component analysis and a factor analysis.

The eleven parameters for grade classification of the Nakdong River tributaries were divided into seven score ranges (Table 1). The same score ranges as those in the water quality standard of South Korea were applied to the concentrations of BOD, COD, TOC, TN, TP, and SS. For the pollution load densities of BOD, COD, TOC, TP, and SS, the score ranges of all seven grades were set by classifying them (5%, 10%, 25%, 50%, 75%, 90%) using box plots.

Among the 11 parameters, those classified as the first factor were selected as parameters to be assigned weights, the basic score was multiplied by 2, and the results were arithmetically summed up. Statistically, the mean of the data or the value at the center of data is used as the value representing the central trend of a data group. In this study, the value ($\bar{X}$) obtained by arithmetically averaging the data surveyed for the parameters in Table 1, $X_1, X_2, X_3, \dots, X_n$, was set as the representative value to calculate the grade classification score.

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_i + \dots + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n} \quad (2)$$

Table 1. Water quality parameters and scores.

No.	Parameter	Score Range						
		1	2	3	4	5	6	7
1	BOD ^{(1)*}	<1	<2	<3	<5	<8	<10	above 10
2	COD ^{(1)*}	<2	<4	<5	<7	<9	<11	above 11
3	TOC ^{(1)*}	<2	<3	<4	<5	<6	<8	above 8
4	TN ^{(2)*}	<1	<2	<3	<4	<5	<6	above 6
5	TP ^{(1)*}	<0.02	<0.04	<0.1	<0.2	<0.3	<0.5	above 0.5
6	SS ^{(1)*}	<5	<10	<15	<20	<25	<35	above 35
7	COD load ⁽³⁾ (kg/d/km ²)	<1	<3	<5	<7	<10	<15	above 15
8	TOC load ⁽³⁾ (kg/d/km ²)	<1	<2	<3	<4	<7	<10	above 10
9	TP load ⁽³⁾ (kg/d/km ²)	<0.01	<0.02	<0.05	<0.08	<0.16	<0.25	above 0.25
10	SS load ⁽³⁾ (kg/d/km ²)	<2	<5	<10	<15	<20	<30	above 30
11	BOD load ⁽³⁾ (kg/d/km ²)	<0.3	<0.6	<1.5	<3	<6	<9	above 9

(1) Water quality standard for rivers (concentration), (2) Water quality standard for lakes (concentration), (3) Pollutant load density, * BOD: biochemical oxygen demand, COD: chemical oxygen demand, TOC: total organic carbon, TN: total nitrogen, TP: total phosphorus, SS: suspended solids.

3. Results and Discussion

3.1. Results of the Statistical Analysis of Monitoring Data

3.1.1. Variations in Water Quality

Variations in water quality at each of the 28 sites in the study area were analyzed for 11 water quality parameters surveyed from 2013 to 2017. Seven parameters, which consisted of the organic pollution indices BOD, COD, and TOC; the eutrophication indices TN and TP; the suspended solid index SS; and flow rate, were selected. The data of these seven parameters surveyed in 8-day intervals are displayed as box plots in Figures 2 and 3, and the data summary is presented in Supplementary Materials (Tables S1 and S2).

The BOD had a range of 0.8–4.9 mg/L on average, which is similar to the median range of 0.7–4.7 mg/L; the first quartile had a range of 0.4–3.4 mg/L; and the third quartile had a range of 1.0–6.2 mg/L. Geumhogang (N15), Topyeongcheon (N20), and Joochungang (N25) showed relatively higher BOD values than other streams. The COD ranged from 1.7 to 231.5 mg/L, the mean value ranged from 3.5 to 11.9 mg/L, the first quartile ranged from 2.5 to 9.3 mg/L, and the third quartile ranged from 3.8 to 14.0 mg/L. Chacheon (N16) and Topyeongcheon (N20) showed high COD values. The TOC had a range of 0.7–75.1 mg/L, the mean value ranged from 1.9 to 8.4 mg/L, the first quartile ranged from 1.3 to 6.4 mg/L, and the third quartile ranged from 2.1 to 9.8 mg/L. Chacheon (N16) showed a high TOC value. In previous studies, these streams are streams with high organic matter concentrations [41–47]. The TN had a range of 0.306–18.978 mg/L, and the median had a range of 1.804–5.921 mg/L. Gwangsancheon (N5) showed a higher TN than the other sites, with a maximum value of 18.978 mg/L. The TN concentration in the Gwangsancheon stream was confirmed in previous studies due to the influence of the non-point source pollution of the surrounding arable land [48]. The average TP ranged from 0.020 to 0.167 mg/L, the first quartile ranged from 0.014 to 0.089 mg/L, and the third quartile ranged from 0.024 to 0.221 mg/L. The SS had a range of 0.2–3623.3 mg/L with large variations. Songyacheon (N2) and Hwapocheon (N26) showed high SS values. Especially, Songyacheon (N2) showed many outliers with high SS values. The flow rate ranged from 0.001 to 751.460 m³/s, and the median had a range of 0.092–36.669 m³/s. There were 11 streams with an average flow rate higher than 5 m³/s. Namgang (N21) had the highest flow rate at 63.305 m³/s, followed by Geumhogang (N15), Hwanggang (N18), Naeseongcheon (N6), Miryanggang (N27), and Hoecheon (N17). The streams with a large watershed area also had a large flow rate. The variations in the water quality of the tributaries of the Nakdong River water system indicate that the tributaries in the upstream region with less anthropogenic pollution have relatively good water quality, whereas the tributaries in the midstream and downstream of the

Nakdong River have high pollution levels. These streams pass through agricultural area (Joosungang (N25), Chacheon (N16), Topyeongcheon (N20), and Hwapocheon (N26)) and the downtown area (Geumhogang (N15), Namgang (N21), and Yangsancheon (N28)).

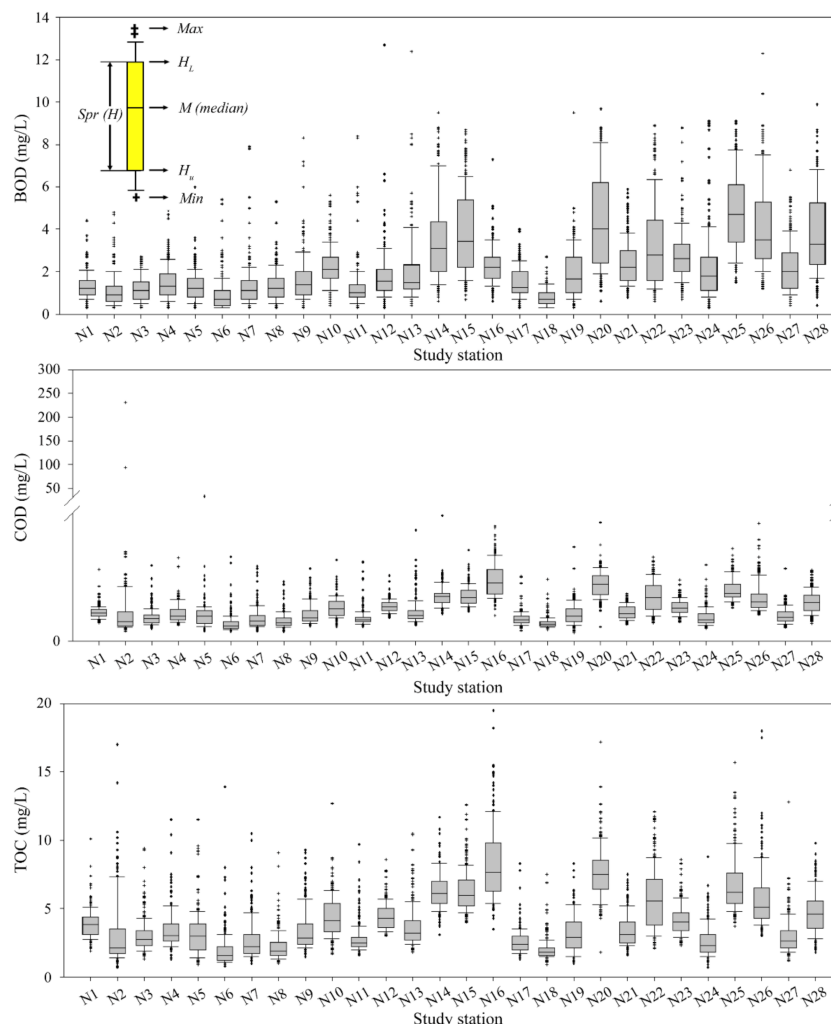


Figure 2. Box plot of the water quality and flow rate data collected at 8-day intervals in tributaries of the Nakdong River.

3.1.2. Pollution Contribution Rate

To calculate the pollution contribution rate (%) of each tributary to the Nakdong River mainstream, the ratio of the load of each tributary to the sum of the delivery load in each tributary was determined using the mean values from 28 sites surveyed for five years from 2013 to 2017. Geumhogang (N15) and Namgang (N21), which pass through cities and industrial complexes, are influenced by the surrounding urban point sources and thus show generally high load ratios. Namgang (N21) showed the highest load ratios of BOD, COD, and TOC among all tributaries at 28.1%, 25.1%, and 24.2%, respectively. Geumhogang (N15) showed the highest load ratios of TP and TN among all tributaries at 23.4% and 27.9%, respectively. Namgang (N21) had the highest flow rate among all tributaries, with an average flow rate of 63.305 m³/s. The streams with a high flow rate were also found to have a high load. The flow rate ratio of Namgang (N21) to the Nakdong River is 29.0%, and the pollution load contribution rates of BOD, TP, and TN are 44.2%, 23.0%, and 25.3%, respectively. Thus, Namgang (N21) showed the largest pollution contribution rate among the major tributary streams flowing to the Nakdong River [49]. In contrast, Songyacheon (N2), Gwangsancheon (N5), Chacheon (N16), Hancheon (N12),

Topyeongcheon (N20), and Joochungang (N25) generally showed low load ratios of BOD, COD, TOC, and TP, and low flow rates. The calculation results of the delivery load and pollution contribution rate are presented in detail in Supplementary Materials (Tables S3 and S4).

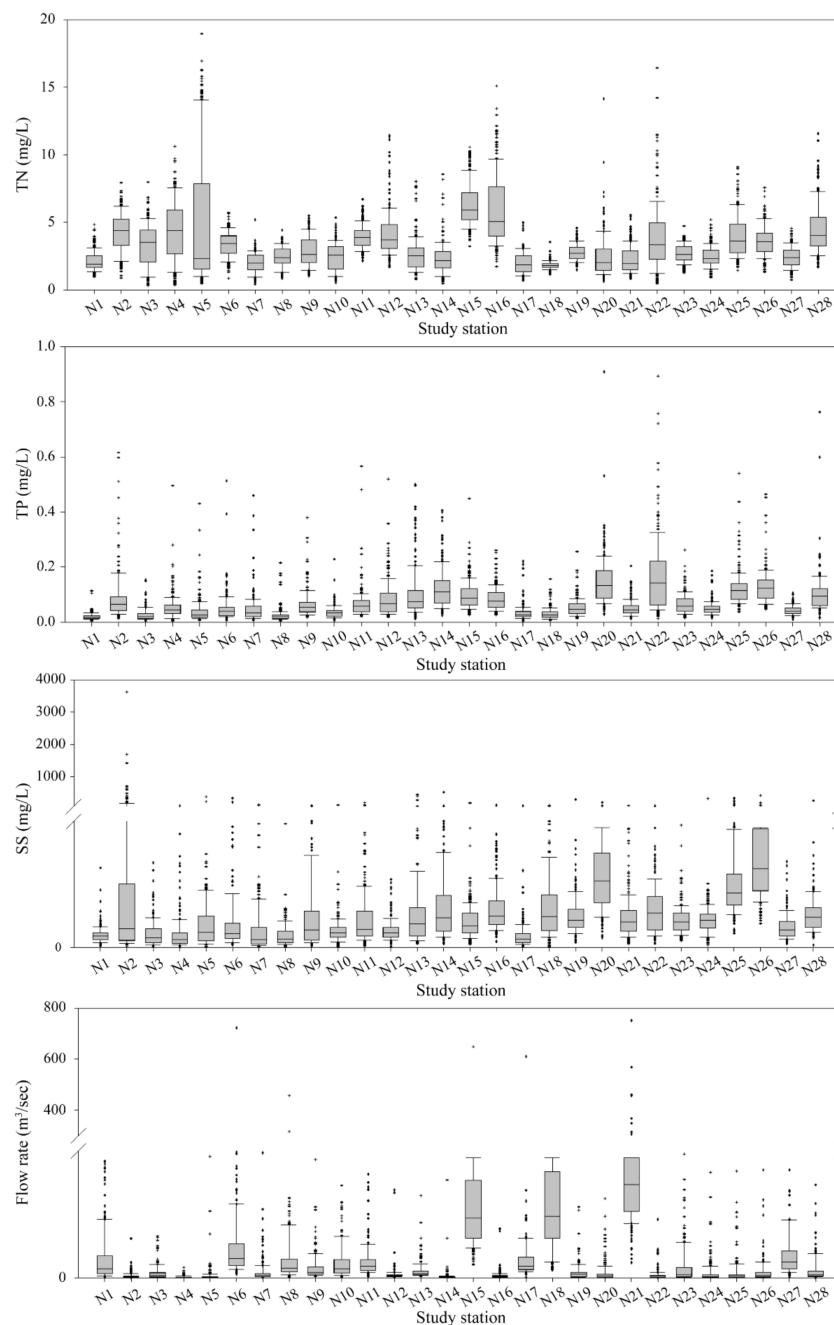


Figure 3. Box plot of the water quality and flow rate data collected at 8-day intervals in tributaries of the Nakdong River (continued).

3.1.3. Pollution Load Density

The Chacheon watershed showed the highest pollution load density, although it has a small watershed area. Chacheon (N16) has a watershed catchment area of 33.6 km², which is the smallest among all the 28 tributary watersheds. Joochungang (N25) and Habincheon (N14) also showed a high pollution potential, that is, pollution load density, despite their small watershed areas. The pollution load density indicates the delivery load per unit area and can be used to indirectly evaluate the pollution

intensity of the watershed. Lee et al. [50] selected streams that need to be prioritized by calculating and comparing the delivery load density of each tributary of the Namgang watershed. Jung et al. [36] calculated the discharge load density by dividing the total amount of pollutants discharged into public waters by the watershed area to classify the tributary streams that have high pollution levels in the Yeongsan River water system. The results of the pollution load density are presented in detail in Supplementary Materials (Tables S3 and S4).

3.1.4. Land Use

In the Nakdong River water system, Namgang (N21) has the largest watershed area of 3476.7 km², followed by Geumhogang (N15) at 2092.3 km². Chacheon (N16) has the smallest watershed area of 33.6 km². Geumhogang (N15) and Namgang (N21) have the largest urban areas of 163.9 km² and 114.4 km², respectively. Geumhogang (N15) also has the largest industrial area of 20.9 km² among 28 tributary streams. Most watersheds have a large forest area, but Chacheon (N16) has an agricultural area of 13.6 km² (40.6%), larger than the forest area of 10.8 km² (32.1%). Joochungang (N25) also has a larger paddy area of 25.2 km² (31.1%) and fields of 17.3 km² (21.4%) than the forest area of 19.8 km² (24.5%), and the urban area in the entire watershed is also somewhat large at 7.9 km² (9.8%). Excluding forests, most surveyed tributary streams have high ratios of paddy areas and fields. Geumhogang (N15) has an urban area corresponding to 21.6% of the total area, excluding forests. Streams that flow through urban or agricultural areas in the watershed can increase the concentrations of organic pollutants and nutrients in the Nakdong River. Kim et al. [46] reported that the concentration of organic matter is affected by the land use type of the watershed. Worrall and Burt [51] demonstrated that the concentration of organic matter reflects watershed characteristics. The results of the analysis of the land use types of the 28 tributary streams are presented in detail in Supplementary Materials (Tables S5 and S6).

3.2. Multivariate Statistical Analysis of Data

3.2.1. Principal Component Analysis and Factor Analysis

Principal component analysis and factor analysis were performed using monthly average values of the parameters at 28 monitoring sites. Four principal components with 1.0 or higher eigenvalue were extracted considering only the principal component axes. The first, second, third, and fourth factors contributed 40%, 23%, 20%, and 17%, respectively. The principal components from the first to the fourth factors accounted for almost all (100%) of the total variance. To verify whether the variables used for the descriptive statistics are suitable for factor analysis, the Kaiser–Meyer–Olkin (KMO) test was conducted. The KMO test measure was calculated as 0.63; it should be at least 0.5 to enable factor analysis. The spherical verification value of the Bartlett's test was 0.000 ($p < 0.05$). Bartlett's test verifies whether the variable used is a diagonal matrix. If the p value is less than 0.05, the diagonal matrix is rejected, meaning that factor analysis is possible. Therefore, both tests confirm that the correlations of the variables were statistically significant and the factor analysis was valid.

The analysis results are shown in Figure 4. COD, TOC, TP, SS, and BOD were classified as the first factor; WT and DO as the second factor; EC, TN, and flow rate as the third factor; and pH as the fourth factor. To investigate the correlations among the parameters classified as the first factor, pattern analysis was performed using SOM (Figure 5). For the number of output neurons in the two-dimensional grid for visualization, 4900 grids ($N = 70 \times 70$) were used for horizontal and vertical directions. The results showed that the variables of the first factor showed a color pattern in the same direction (pattern) without opposite patterns, which is different from the pattern of the second factor. Based on the analysis results, the tributaries of the Nakdong River have the largest pollution levels caused by organic pollutants and nutrients, followed by the seasonal factor and the biological metabolism factor (Figure 4).

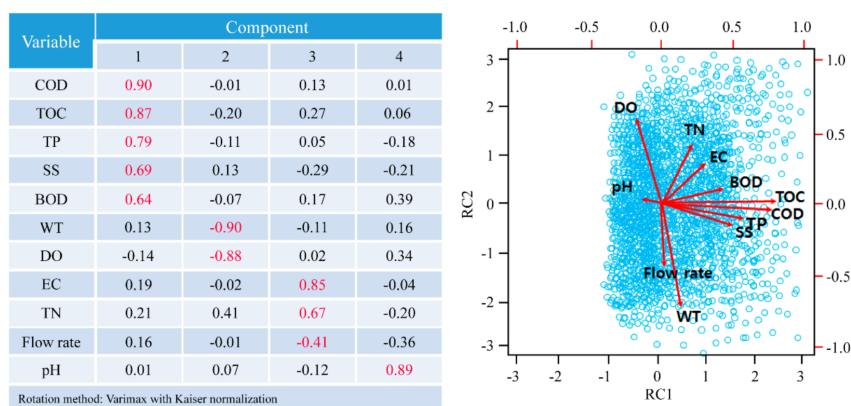


Figure 4. Rotated component matrix by factor analysis and the biplot.

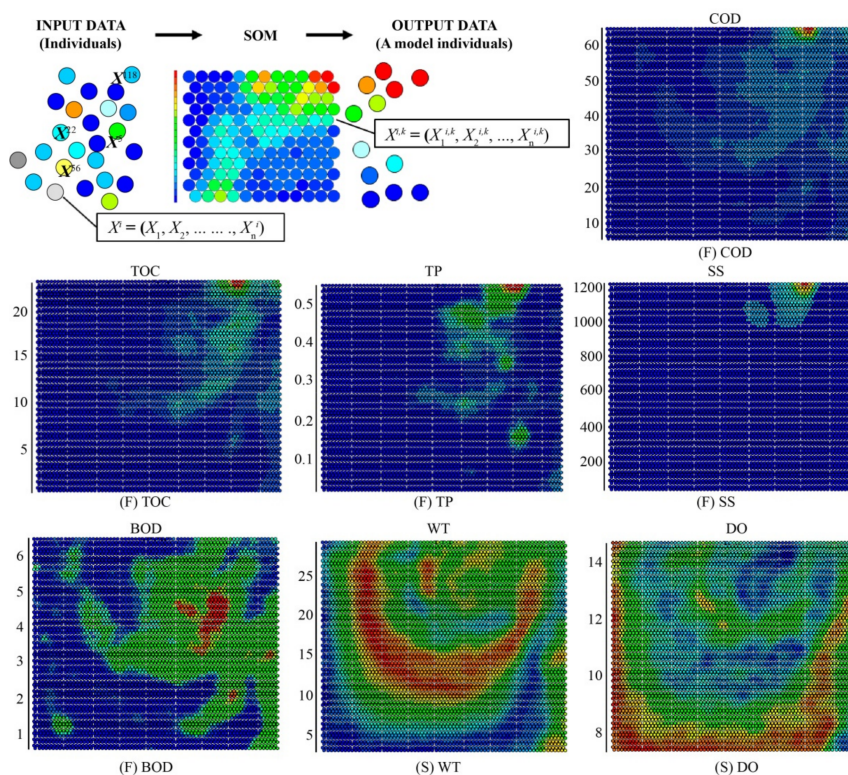


Figure 5. Results of the SOM (self-organizing map) analysis using first and second factors (Red: high value; blue: low value).

3.2.2. Cluster Analysis

The annual average values of 11 parameters at each of the 28 monitoring sites were used to perform the spatial cluster analysis. The clustered data were expressed as a dendrogram, and the results of the cluster analysis based on the factor analysis results are displayed in Figure 6. A dendrogram reveals the similarity of water quality changes between major tributaries and the mainstream, facilitates the classification of clusters, and provides visual information [52]. Furthermore, R uses the Nb Clust library to estimate the number of clusters. When the Nb Clust function was executed, two groups were recommended and classified considering all the pollutants in the watershed. Cluster 1 included tributaries upstream of the Nakdong River, with low pollution levels and similar water quality characteristics. The streams that flow through large and small cities have a high share of agricultural area and high water pollution potential due to the release of pollutants from surrounding factories or

livestock yards and were classified as Cluster 2. Geumhogang (N15), the representative tributary of the Nakdong River, which has a high level of pollution caused by the release of domestic and industrial wastewater, was also included in this group.

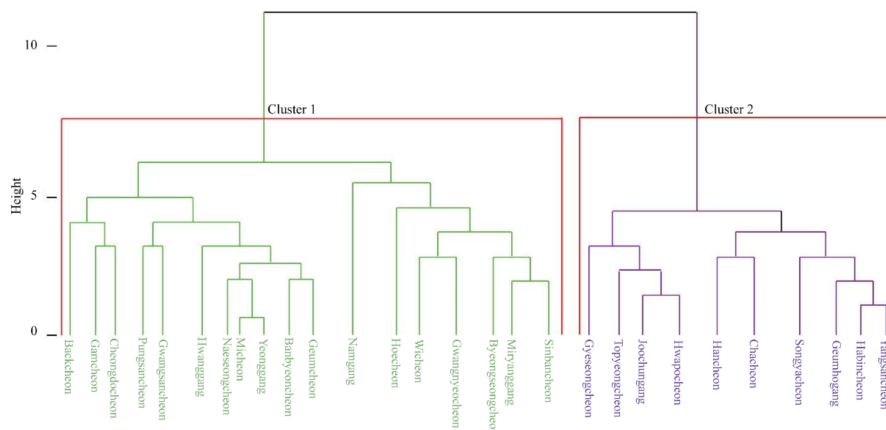


Figure 6. Dendrogram of 28 monitoring sites using hierarchical cluster analysis based on the water quality in the Nakdong River tributaries (Cluster 1: low pollution; Cluster 2: high pollution).

Geumhogang (N15) has a large impact on the Nakdong River because it experiences large environmental changes and flows through a large city. The tributaries classified in the same group, Chacheon (N16), Gyeseongcheon (N22), Habincheon (N14), Topyeongcheon (N20), Joochungang (N25), Yangsancheon (N28), and Hwapocheon (N26) also have high water pollution potential due to the watershed environmental characteristics. According to Singh et al. [38], cluster analysis results can be grouped by the water pollution level of the monitoring sites. Shrestha and Kazama [1] and Shen et al. [53] classified and reported the groups according to the characteristics of pollutants around the river and the impact of water pollution and urban sewage related to land use. In this study, the groups were classified by the characteristics of surrounding pollutants or the water pollution level of the stream. The results of the SOM analysis based on the recommended number of clusters are shown in Figure 7, which demonstrates the characteristics of the cells with the radar plots. Radars of different colors indicate different water quality factors. The classification of sites to cells (clusters) can be checked by the site name in each cell. The sites with water quality characteristics that have large effects of DO, pH, and flow rate were classified into Cluster 1, whereas the sites with water quality characteristics that have large effects of COD, TOC, BOD, TP, TN, SS, and WT were classified into Cluster 2. This cluster analysis method is useful for classifying streams with similar water quality characteristics and for distinguishing the effects of tributaries on the mainstream. Although cluster analysis is useful for classifying streams with similar characteristics based on the water quality management measures, it is difficult to prioritize the cluster analysis results within the same group.

3.3. Grade Classification Results

When the representative values of 28 sites were arithmetically summed according to the grade classification criteria in Table 1, the scores ranged from 30 to 100. The results were distinguished by scores, and the tributaries are listed in Table 2 by rank. The top 10 tributaries were those with high pollutant concentrations. Joochungang (N25), Topyeongcheon (N20), Hwapocheon (N26), Chacheon (N16), Gwangyeocheon (N23), Geumhogang (N15), and Yangsancheon (N28) have high population density due to urbanization, a large impervious area, and significant levels of point and nonpoint source pollution due to the high land use ratio of the fields. In Joochungang (N25) and Chacheon (N16), the urban dry areas and fields accounted for 62.3% and 58.8%, and 85.6% and 86.7%, respectively, excluding the forest areas. Therefore, these areas are highly affected by the high density of pollution sources.

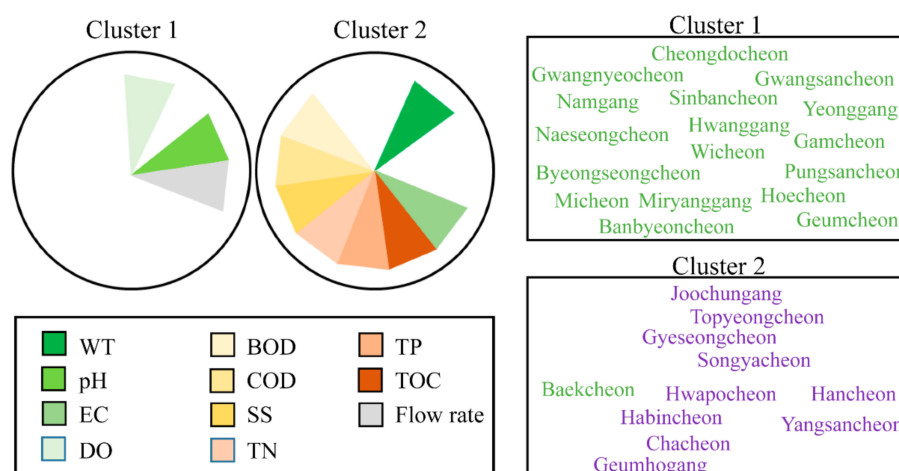


Figure 7. Results of the SOM analysis.

Table 2. Nakdong River tributaries ranked by water quality improvement priority score.

Ranking	Station	Score	Code No.
1	Joochungang	100	N 25
2	Topyeongcheon	96	N 20
3	Hwapocheon	95	N 26
4	Chacheon	94	N 16
5	Gwangnyeocheon	84	N 23
6	Geumhogan	82	N 15
7	Yangsancheon	79	N 28
8	Habincheon	76	N 14
9	Gyeseongcheon	72	N 22
10	Namgang	67	N 21
11	Baekcheon	66	N 13
12	Sinbancheon	66	N 19
13	Songyacheon	58	N 2
14	Byeongseongcheon	58	N 9
15	Hancheon	58	N 12
16	Hwanggang	56	N 18
17	Gwangsancheon	54	N 5
18	Gamcheon	53	N 11
19	Cheongdocheon	51	N 24
20	Hoecheon	49	N 17
21	Miryanggang	48	N 27
22	Wicheon	44	N 10
23	Geumcheon	44	N 7
24	Yeonggang	43	N 8
25	Naeseongcheon	38	N 6
26	Pungsancheon	35	N 4
27	Micheon	33	N 3
28	Banbyeoncheon	30	N 1

Furthermore, these tributaries have high pollution load density and bad water quality, resulting in high scores. Thus, water environment management measures for these streams should be considered to improve the water quality of the Nakdong River mainstream. Hancheon (N12), Baekcheon (N13), Sinbancheon (N19), Songyacheon (N2), and Byeongseongcheon (N9) also have high water pollution potential, and thus attention should also be paid to the water quality management of these streams. Figure 8 illustrates the scores and rankings of the tributaries along with pollution intensities according to the pollution load density in each watershed for easy identification.

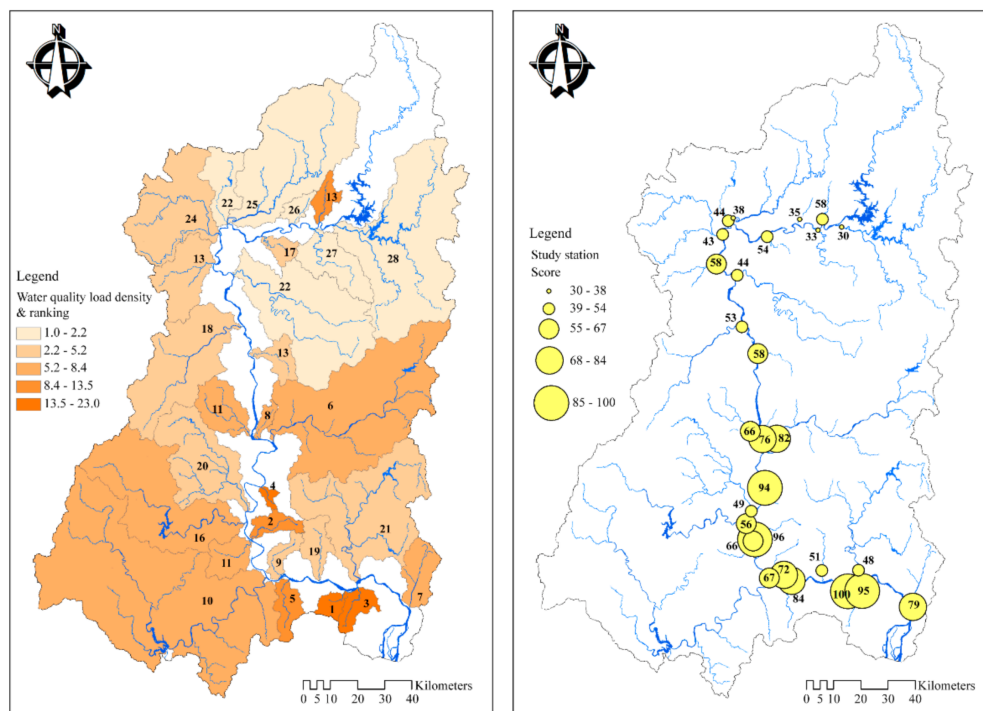


Figure 8. GIS maps of water quality load density and grade classification results.

4. Conclusions

In this study, we calculated the pollution contribution rates and conducted pollution load density analysis, land use analysis, and multivariate statistical analysis based on the water quality and flow rate monitoring data collected from 2013 to 2017 at 28 sites in the tributaries flowing into the Nakdong River. The results were scored and graded to further improve the existing methods. In addition, a neural network model (self-organizing map, SOM) was applied to the data. When the pollution contribution rates were calculated using the delivery load, among the 28 tributaries that flowed into the Nakdong River mainstream, Geumhogang (N15) and Namgang (N21) showed high load ratios. Chacheon (N16), Joochungang (N25), Topyeongcheon (N20), and Hwapocheon (N26) located in the midstream and downstream of the Nakdong River were found to have high pollution intensities per unit area. The major water quality parameters in Nakdong River tributaries were COD, TOC, SS, TP, and BOD. In general, tributaries with low pollution were classified as Cluster 1. Joochungang (N25), Topyeongcheon (N20), Hwapocheon (N26), Chacheon (N16), Gwangyeocheon (N23), Geumhogang (N15), Yangsancheon (N28), and Habincheon (N14), which have high pollution input from domestic and industrial wastewater and agricultural lands were classified as Cluster 2. The SOM analysis was found to be effective for visualizing and clustering a wide range of pollutant characteristics. According to the grade classification criteria, the streams that require high-priority water management measures were Joochungang (N25), Chacheon (N16), Topyeongcheon (N20), Hwapocheon (N26), and Geumhogang (N15), which have high pollutant concentrations, flow through a large city, or have a high share of agricultural area. The proposed prioritization method can be a reasonable way to select rivers with high pollution levels and progressively improve the water quality of the Nakdong River. However, future studies should analyze long-term monitoring data and focus on finding a grading method by using a greater number of environmental variables. In conclusion, the results of the grade classification method can be used to determine the high-priority streams that need water quality improvement. By targeting those streams, the proposed method can serve as a decision-making tool for more directed and effective water quality management and policy formulation.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2071-1050/12/21/9149/s1>, Figure S1: Distribution of land use types in the Nakdong River tributaries watershed (example), Table S1: Descriptive statistics of water quality and flow rate of tributaries at Nakdong River, Table S2: Descriptive statistics of water quality and flow rate of tributaries at Nakdong River (continued), Table S3: Delivery load and water quality load density of total phosphorus, organic matters and flow rate from tributaries of the Nakdong River, Table S4: Delivery load and water quality load density of total phosphorus, organic matters and flow rate from tributaries of the Nakdong River (continued), Table S5: Classifieds Nakdong River tributary land use area, Table S6: Classifieds Nakdong River tributary land use area (continued).

Author Contributions: Conceptualization, K.-Y.J. and E.H.N.; methodology, K.-Y.J.; software, K.-Y.J. and S.-Y.H.; validation, K.-Y.J., Y.L. and E.H.N.; formal analysis, K.-Y.J. and S.-Y.H.; investigation, K.-Y.J. and S.C.; resources, Y.L.; data curation, K.-Y.J. and S.C.; writing—original draft preparation, K.-Y.J.; writing—review and editing, K.-Y.J. and E.H.N.; visualization, K.-Y.J. and S.-Y.H.; supervision, E.H.N.; project administration, K.-Y.J., K.K. and E.H.N.; funding acquisition, Y.L. and K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Institute of Environment Research (NIER), grant number NIER-2019-03-01-013.

Acknowledgments: This paper was supported by National Institute of Environment Research (NIER) in 2019.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Shrestha, S.; Kazama, F. Assessment of Surface Water Quality Using Multivariate Statistical Techniques: A Case Study of the Fuji River Basin, Japan. *Environ. Modell. Softw.* **2007**, *22*, 464–475. [\[CrossRef\]](#)
- Anh, J.M.; Lyu, S. Selection of Priority Tributaries for Point and Non-Point Source Pollution Management. *KSCE J. Civ. Eng.* **2020**, *24*, 1060–1069.
- Kim, Y.; Gal, B.; Park, J.; Kim, S.; Im, T. Classification of Nakdong River Tributaries under Priority Management base on their Characteristics and Water Quality Index. *J. Korean Soc. Environ. Eng.* **2018**, *40*, 73–81. [\[CrossRef\]](#)
- Im, T.H.; Na, S.; Shin, S.; Son, Y. Water Quality Analysis in Nakdong River Tributaries for the Determination of Priority Management Areas. *J. Korean Soc. Environ. Eng.* **2016**, *38*, 558–565. [\[CrossRef\]](#)
- Park, S.H.; Moon, E.H.; Choi, J.H.; Cho, B.W.; Kim, H.S.; Jeong, W.H.; Yi, S.J.; Kim, Y.I. Analysis of Distribution Characteristics of Flowrate and Water Quality in Tributary at Chungcheongnam-do. *J. Korean Soc. Environ. Eng.* **2011**, *29*, 107–113.
- Barakat, A.; Baghdadi, M.E.; Rais, J.; Aghezzaf, B.; Slassi, M. Assessment of spatial and seasonal water quality variation of Oum Er Rbia River (Morocco) using multivariate statistical techniques. *Int. Soil Water Conserv. Res.* **2016**, *4*, 284–292. [\[CrossRef\]](#)
- Yan, H.; Zou, Z.; Wang, H. Adaptive neuro-fuzzy inference system for classification of water quality status. *J. Environ. Sci.* **2010**, *22*, 1891–1896. [\[CrossRef\]](#)
- Cho, B.; Choi, J.; Yi, S.; Kim, Y. Selection Priority of Tributary Catchments for Improving Water Quality using Stream Grouping Method. *J. Korean Soc. Water Environ.* **2012**, *28*, 18–25.
- Na, S.; Lin, T.H.; Lee, J.Y.; Kwon, H.; Cheon, S.U. Flow Rate-Water Quality Characteristics of Tributaries and a Grouping Method for Tributary Management in Nakdong River. *J. Wetl. Res.* **2015**, *17*, 380–390. [\[CrossRef\]](#)
- Jung, K.Y.; Kim, K.H.; Lee, J.W.; Lee, I.J.; Yoon, J.S.; Lee, K.L.; Im, T.H. Selection of Priority Management Target Tributary for Effective Watershed Management in Nam-River Mid-Watershed. *J. Korean Soc. Water Environ.* **2013**, *29*, 514–522.
- Kowalkowski, T.; Zbytniewski, R.; Szpejna, J.; Buszewski, B. Application of chemometrics in river water classification. *Water Res.* **2006**, *40*, 744–752. [\[CrossRef\]](#)
- Azhar, S.C.; Aris, A.Z.; Yusoff, M.K.; Ramli, M.F.; Juahir, H. Classification of river water quality using multivariate analysis. *Procedia Environ. Sci.* **2015**, *30*, 79–84. [\[CrossRef\]](#)
- Jung, K.Y.; Lee, K.L.; Im, T.H.; Lee, I.J.; Kim, S.; Han, K.Y.; Ahn, J.M. Evaluation of water quality for the Nakdong River watershed using multivariate analysis. *Environ. Technol. Inno.* **2016**, *5*, 67–82. [\[CrossRef\]](#)
- Choi, H.G.; Han, K.Y.; Choi, S.Y. Comparison of Organic Matter Distribution in Major Tributaries of the Nakdong River. *J. Environ. Impact Assess.* **2011**, *20*, 833–843.
- Park, H.S.; Park, C.K.; Song, M.K.; Baek, K.H.; Shin, S.K. Evaluation of water quality characteristics using factor analysis in the Nakdong river. *J. Korean Soc. Water Environ.* **2001**, *17*, 693–701.
- Park, J.H. Characteristics and Fate of Refractory Organic Matter in the Nakdong River. Master's Thesis, Department of Environmental Engineering, Pukyong National University, Busan, Korea, 2012; pp. 4–80.

17. Lee, S.H.; Kim, B.R.; Lee, H.W. A study on water quality after construction of the weirs in the middle area in nakdong river. *J. Korean Soc. Environ. Eng.* **2014**, *36*, 264–285. [[CrossRef](#)]
18. Hwang, H.S.; Park, B.K.; Kim, Y.S.; Park, K.J.; Cheon, S.U.; Lee, S.J. Research on the Applicability of the Load Duration Curve to Evaluate the Achievement of Target Water Quality in the Unit Watershed for a TMDL. *J. Korean. Soc. Water. Environ.* **2011**, *27*, 885–895.
19. Jung, K.Y.; Kim, H.T.; Kim, S.S.; Kim, S.; Shin, D.S.; Kim, K.H. Application of the Load Duration Curve (LDC) to Evaluate the Achievement Rate of Target Water Quality in the Nakdong River Unit Watersheds. *J. Environ. Sci.* **2017**, *26*, 433–445. [[CrossRef](#)]
20. Jung, K.Y. Advanced Control Strategies on Water Quality Management in Nakdong River. Master's Thesis, Kyungpook National University, Daegu, Korea, 2009; pp. 94–100.
21. National Institute of Environmental Research (NIER). Report of the National Pollution Source Survey, Based on the 2012, 11-1480523-000429-10. *Natl. Inst. Environ. Res.* **2014**, *14*, 122.
22. Park, B.J.; Kim, H.S.; Jung, K.S.; Ji, H.K. A Study on the Impact Range Calculation at the Downstream of Dam. *J. Korea Water Resour. Assoc.* **2008**, *41*, 1009–1021. [[CrossRef](#)]
23. Han, J.H.; An, K.G. Water Quality Variation Dynamics between Artificial Reservoir and Effected Downstream Watershed: The Case Study. *Korean. J. Limnol.* **2008**, *41*, 382–394.
24. Kim, G.W. A study on the Impact of Continuous Dams in the North Han River on the Downstreams Water Quality. Master's Thesis, Ewha Womans University, Seoul, Korea, 2020; pp. 1–91.
25. Barillier, A.J.; Coste, M. Experimental reservoir water release impact of the water quality on a river 60 km downstream (upper seine river, France). *Wat. Res. Limnol.* **1993**, *27*, 635–643. [[CrossRef](#)]
26. Chung, S.W.; Kim, Y.K. Impact of a Flushing Discharge from an Upstream Dam on the NH₃-N Concentrations during Winter Season in Geum River. *J. Korean. Soc. Water. Environ.* **2005**, *21*, 609–616.
27. Yoon, Y.S.; Yu, J.J.; Kim, M.S.; Lee, H.J. Computation and Assessment of Delivery Pollutant Loads for the Streams in the Nakdong River Basin. *J. Korean Soc. Water Environ.* **2006**, *22*, 277–287.
28. Kim, S.S.; Lee, D.Y.; Moon, B.H. Study on the Temporal Characterization of Water Quality Data Using Multivariate Statistical Methods at Urban Streams. *KSWST J. Wat. Treat.* **2015**, *23*, 75–86. [[CrossRef](#)]
29. Kim, J.H. A Study on Data Analysis Tool Software Performance Comparison Using Multivariate Analysis of Variance. Master's Thesis, Soongsil University, Seoul, Korea, 2017; pp. 11–16.
30. Kim, M.A.; Lee, J.K.; Zoh, K.D. Evaluation of the Geum River by Multivariate Analysis: Principal Component Analysis and Factor Analysis. *J. Korean Soc. Water Environ.* **2007**, *23*, 161–168.
31. Sakamoto, H.; Hamaguchi, R.; Iwase, M. Evaluation of river water pollution in the chanai district of eastern hokkaido by the multivariate analysis method. *Water Waste* **1988**, *30*, 239–245.
32. Vega, M.; Pardo, R.; Barrado, E.; Deban, L. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res.* **1998**, *32*, 3581–3592. [[CrossRef](#)]
33. Usman, A.; An, K.G. Reservoir Water Quality Assessment Based on Chemical Parameters and the Chlorophyll Dynamics in Relation to Nutrient Regime. *Pol. J. Environ. Stud.* **2019**, *28*, 1043–1061.
34. Zhang, X.; Wang, Q.; Liu, Y.; Wu, J.; Yu, M. Application of multivariate statistical techniques in the assessment of water quality in the Southwest New Territories and Kowloon, Hong Kong. *Environ. Monit Assess.* **2011**, *173*, 17–27. [[CrossRef](#)]
35. Pinto, C.; Calazans, G.M.; Oliveira, S.C. Assessment of spatial variations in the surface water quality of the Velhas River Basin Brazil using multivariate statistical analysis and nonparametric statistics. *Environ. Monit Assess.* **2019**, *191*, 164. [[CrossRef](#)]
36. Jung, S.J.; Kim, K.S.; Seo, D.J.; Kim, J.H.; Lim, B.J. Evaluation of Water quality characteristics and grade classification of Yeongsan river tributaries. *J. Korean Soc. Water Environ.* **2013**, *29*, 504–513.
37. Singh, K.P.; Malik, A.; Mohan, D.; Sinha, S. Multivariate Statistical Techniques for the Evaluation of Spatial and Temporal Variations in Water Quality of Gomti River (India)-a Case Study. *Water Res.* **2004**, *38*, 3980–3992. [[CrossRef](#)] [[PubMed](#)]
38. Ahn, J.M.; Lee, I.J.; Jung, K.Y.; Kim, J.E.; Lee, K.C.; Cheon, S.U.; Lyu, S.W. Characteristics of Trend and Pattern for Water Quality Monitoring Networks Data using Seasonal-Kendall SOM and RDA on the Mulgeum in the Nakdong River. *J. Environ. Sci.* **2016**, *25*, 361–371. [[CrossRef](#)]
39. Yang, D.S.; Im, T.H.; Lee, I.J.; Jung, K.Y.; Kim, G.H.; Kwon, H.G.; Yoo, J.C.; Ahn, J.M. Patterns and Trends of Water Level and Water Quality at the Namgang Junction in the Nakdong River Based on Hourly Measurement Time Series Data. *J. Environ. Sci.* **2018**, *27*, 63–74. [[CrossRef](#)]

40. Paik, B.C.; Kim, C.K.; Kim, T.R. A Study on the Pattern Analysis of Correlation between the River Flow and Water Quality using a SOM (Self Organizing Maps) Technique. *J. Korean Soc. Urban Environ.* **2011**, *11*, 153–160.
41. Jung, K.Y. Statistical Analysis of Water Quality Grade Classification and Water Environmental Pollution Assessment. Ph.D. Thesis, Kyungpook National University, Daegu, Korea, 2017; pp. 60–113.
42. Jung, K.Y.; Cho, S.H.; Ha, D.W.; Kang, T.W.; Lee, Y.J.; Han, K.Y.; Kim, K.H. An Analysis of Long-term Changes in Water Quality of Geumho River using Statistical Techniques. *J. Environ. Sci.* **2018**, *27*, 883–899. [[CrossRef](#)]
43. Jung, K.Y.; An, J.M.; Lee, K.L.; Lee, I.J.; Yu, J.J.; Cheon, S.U.; Kim, K.S.; Han, K.Y. Temporal and Spatial Analysis of Non-biodegradable Organic Pollutants in the Geumho River System. *J. Environ. Sci.* **2015**, *24*, 1343–1362.
44. Kim, J.C.; Lee, J.H.; Yoo, C.S.; Kim, S.D. Impact of urbanization on hydrology of Geumho river watershed: A model study. *J. Korean. Soc. Water Environ.* **2007**, *23*, 535–542.
45. Kim, Y.H.; Han, J.H.; An, K.G. Physico-chemical Water Quality Gradients Along the Main Axis of the Headwater-to-Downstream of Geumho River and Their Influences on Fish Guilds. *J. Korean. Soc. Water Environ.* **2012**, *28*, 561–573.
46. Kim, H.S.; Hong, J.J.; Seong, J.U.; Choi, K.S.; Park, J.C. Comparison of Organic Matter Distribution in Major Tributaries of the Nakdong River. *J. Korean. Soc. Water Environ.* **2013**, *29*, 618–624.
47. Park, K.H. The impact of environmental characteristics in the Geumho river watershed on stream water quality. *J. Korea Assoc. Geogr. Inf. Stud.* **2003**, *6*, 85–98.
48. An, J.M.; Jung, K.Y.; Kim, S.; Lee, H.J.; Shin, S.H.; Yang, D.S.; Shin, D.S.; Na, S.M. Development and Application of the Grid-Distributed Model for Contribution Rate Analysis on Non-point Source Pollution According to Landuse. *J. Korean Soc. Water Environ.* **2017**, *33*, 78–89.
49. NRERC. *08 Management Report of Target Water Quality Monitoring System in Nakdong River Basin, 11-1480523-000444-10*; National Institute of Environmental Research: Gyeongsangbuk-do, Korea, 2009.
50. Lee, J.W.; Kwon, H.G.; Kwak, I.S.; Youn, J.S.; Cheon, S.U. The estimation of contribution ratio for sub stream in Nam river basin. *J. Environ. Impact Assess.* **2012**, *21*, 745–755.
51. Worrall, F.; Burt, T.P. Trend in DOC Concentration in Great Britain. *J. Hydrol.* **2007**, *346*, 81–92. [[CrossRef](#)]
52. Kim, G.A.; Kim, Y.J.; Song, M.J.; Ji, K.W.; Yu, P.J.; Kim, C.W. Evaluation of Water Quality Characteristics in the Nakdong River using Multivariate Analysis. *J. Korean Soc. Water Environ.* **2007**, *23*, 814–821.
53. Shen, Y.N.; Lu, J.; Chen, D.J.; Shi, Y.M. Response of stream pollution characteristics to catchment land cover in Cao-E River Basin China. *Pedosphere* **2011**, *21*, 115–123. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).