

Article

Surface Water Quality Assessment and Contamination Source Identification Using Multivariate Statistical Techniques: A Case Study of the Nanxi River in the Taihu Watershed, China

Zhi-Min Zhang ^{1,*} , Fei Zhang ² , Jing-Long Du ¹ and De-Chao Chen ¹

¹ College of Geography Science and Geomatics Engineering, Suzhou University of Science and Technology, Suzhou 215009, China; jldu@mail.usts.edu.cn (J.-L.D.); dcchen@mail.usts.edu.cn (D.-C.C.)

² College of Resources and Environment Science, Xinjiang University, Urumqi 830046, China; zhangfei3s@xju.edu.cn

* Correspondence: zmzhang@mail.usts.edu.cn; Tel.: +86-137-7179-7831

Abstract: Understanding the spatiotemporal patterns of water quality is crucial because it provides essential information for water pollution control. The spatiotemporal variations in water quality for the Nanxi River in the Taihu watershed of China were evaluated by a water quality index (WQI) and multivariate statistical techniques; additionally, the potential sources of contamination were identified. The data set included 22 water quality parameters collected during the monitoring period from 2015 to 2020 for 14 monitoring stations. WQI assessment revealed that approximately 85% of monitoring stations were classified as “medium-low” water quality, and most showed continuous improvement in water quality. Cluster analysis divided the 14 monitoring stations into three clusters (low contamination, medium contamination and high contamination). Discriminant analysis identified pH, petroleum, volatile phenol, chemical oxygen demand, total phosphorus, F, S, fecal coliform, SO₄, Cl, NO₃-N, total hardness, NO₂-N and NH₃ as important parameters affecting spatial variations. Factor analysis identified four potential contamination source types: nutrient, organics, feces and oil. This study demonstrated the usefulness of multivariate statistical techniques in assessing large data sets, identifying contamination source types, and better understanding spatiotemporal variations in water quality to restore and protect water resources.

Keywords: Nanxi River; multivariate statistical techniques; water quality index; water quality assessment



Citation: Zhang, Z.-M.; Zhang, F.; Du, J.-L.; Chen, D.-C. Surface Water Quality Assessment and Contamination Source Identification Using Multivariate Statistical Techniques: A Case Study of the Nanxi River in the Taihu Watershed, China. *Water* **2022**, *14*, 778. <https://doi.org/10.3390/w14050778>

Academic Editor: Roko Andricevic

Received: 19 January 2022

Accepted: 27 February 2022

Published: 1 March 2022

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



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Deterioration of the water environment is a prominent problem in worldwide watershed management and seriously threatens the security of the water ecological environment [1]. Natural factors (such as climate, topography, geology and soil) and human activities (such as urbanization, industrial production and agricultural practice) affect the surface water quality of an area [2–5]. The seasonal changes in precipitation, hydrological conditions and stream runoff have marked effects on stream flow and the consequent pollutant concentrations in surface water [1,6–8]. Dynamic spatiotemporal assessment of water quality can be used to analyze water contamination problems, identify potential contamination source types, and provide information support and reference to effectively manage water resources [3].

To effectively prevent and control surface water contamination, reliable water quality data for in-depth research is necessary. Considering the spatiotemporal variation in the physicochemical and biological characteristics of surface water, a long-term monitoring plan to accurately assess water quality should be developed [9]. Environmental protection departments in China have established sound water quality monitoring networks and continuous water quality monitoring procedures that monitor the physical properties

(e.g., temperature, pH and electrical conductivity, etc.), total organic components, nutrients and inorganic components, as well as the biological and microbial conditions. In water quality assessments, multiple water quality parameters are typically collected at multiple monitoring stations in different monitoring periods, and this process generates a complex data matrix [10]. Due to the potential multivariable correlations among monitoring stations, monitoring periods and water quality parameters, this complex data set is often challenging to analyze and explain [11–13]. In a comprehensive assessment of water quality, the challenge is to determine whether the changes in water quality should be attributed to the contamination of rivers by human activities or biogeochemical changes in natural processes [14]. Furthermore, the water quality parameters that can best describe the spatiotemporal changes and identify contamination source types should be determined.

The water quality index (WQI) is a useful method for evaluating the change and trend of water environment quality by synthesizing multiple original parameters to a single index [10]. As a water quality assessment model, WQI determines the relative weight of each parameter based on its importance in water environment protection and integrates multiple variables into a dimensionless variable to represent the comprehensive water quality status and grade [15–18]. WQI has played an increasingly crucial role in the water quality assessment of rivers, lakes and groundwater [19–24].

With the increased number and dimension of measurement parameters in samples, the problem of allocating unknown samples and mining valuable information becomes increasingly complex. Therefore, using multivariate statistical techniques and data reduction simultaneously to obtain satisfactory results is necessary [10]. Multivariate statistical techniques, such as cluster analysis (CA), discriminant analysis (DA), principal component analysis (PCA) and factor analysis (FA) have been widely applied to evaluate water quality, can simplify data dimensions from complex water quality data matrices and remove redundant information without losing valuable information [8,25–29]. Multivariate statistical techniques can identify spatiotemporal patterns of water quality and analyze the possible factors causing spatiotemporal variations in water quality and affecting the health of water ecosystems [1,30].

The Nanxi River in the Taihu watershed, as a rapidly urbanized area in China, is experiencing high disturbance from human activities and serious water contamination problems [31]. The spatiotemporal variations in surface water quality and the identification of contamination source types are critical for sustainable watershed water quality management. However, studies focusing on the identification of contamination source types in the Nanxi River are limited. The primary aims of this study are as follows: (1) to evaluate the contamination levels of different monitoring stations and periods to examine the spatiotemporal distributions of water quality using WQI; (2) to extract the clustering information of monitoring stations, and determine the most important classification variables for the spatial variations in water quality; and (3) to analyze the potential impact factors of water quality in three regions with different contamination levels and explore possible contamination source types (natural processes or human activities).

2. Materials and Methods

2.1. Study Area

The Nanxi River (119°08′–119°36′ E, 31°1′–31°41′ N) is the main river in the western part of the Taihu watershed in China (Figure 1). The total extension of the study area is 1535.87 km² and includes 39% farmland, 23% water area, 22% forestland and 16% built-up land. It belongs to the subtropical monsoon climate zone, with an average annual temperature of 16 °C and an average annual precipitation of 1147 mm, 70% of which occurs in the rainy season from May to October. The area comprises low mountains, hills, plain polders and other landform types, with elevations of 1–702 m. The main types of soil are paddy soils, yellow-brown soils, and yellow cinnamon soils [32]. The regional zonal vegetation is an evergreen and deciduous broad-leaved mixed forest. The main crops are rice, rape, tea and sericulture, etc. [31].

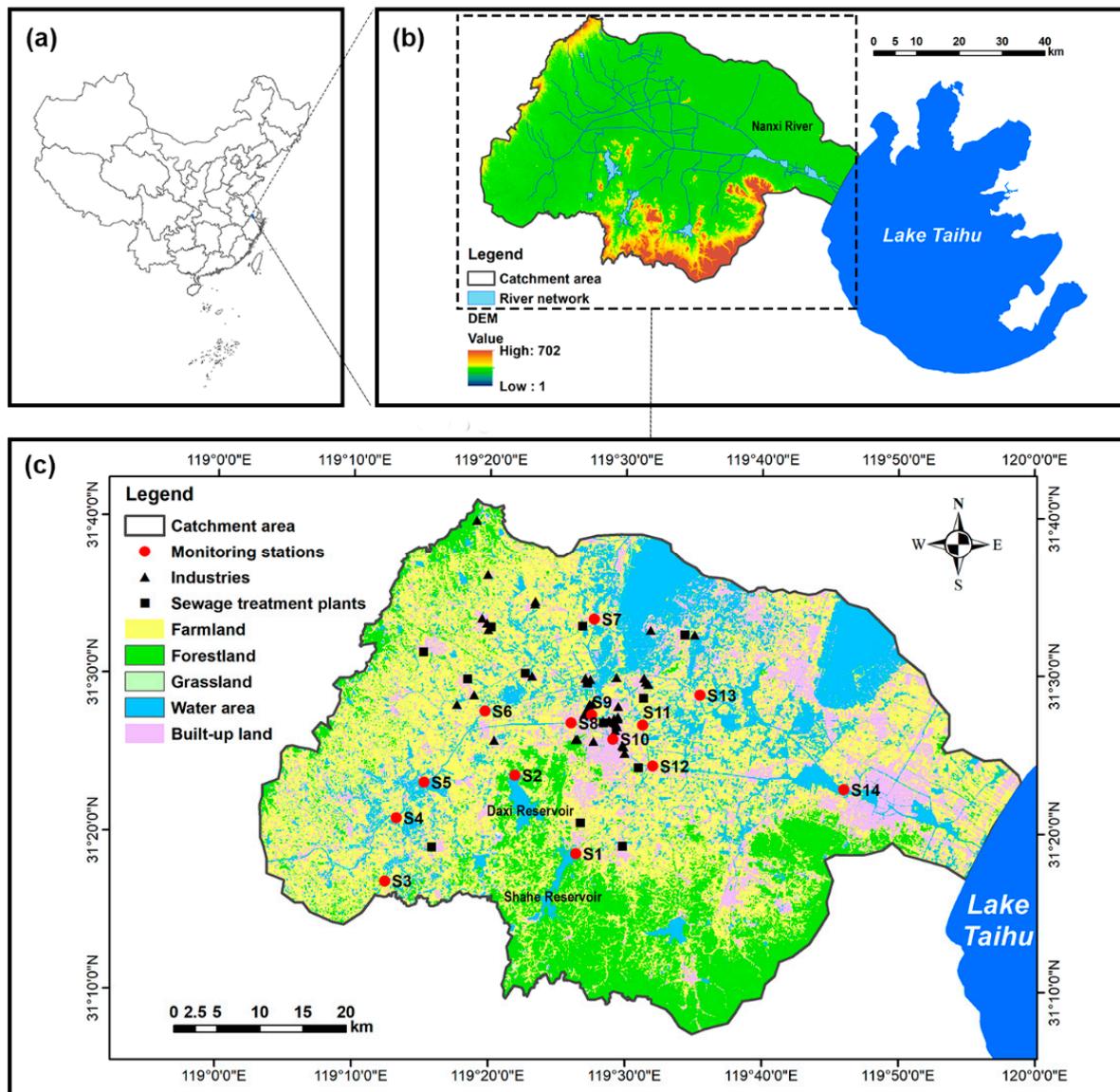


Figure 1. (a) Location of the study area in China; (b) Topography and river network of the study area; (c) Study area with 14 monitoring stations in the Nanxi River.

This area is relatively developed in the Taihu watershed, with a population of approximately 763,000. The area has many chemicals, synthetic materials, and mechanical, electronic and cement factories. The basin has fertile paddy soil, which is very suitable for various agricultural activities. The main contamination source types in this area include municipal and industrial wastewater, livestock and poultry breeding, planting and aquaculture. In 2019, the total amount of industrial and domestic sewage discharge was 2248.8×10^4 and 3244.9×10^4 tons, respectively, including 3043.5 tons of chemical oxygen demand (COD), 538.5 tons of ammonia nitrogen ($\text{NH}_3\text{-N}$), and 14.8% of the industrial sewage treatment rate; the COD, total nitrogen (TN) and total phosphorus (TP) of agricultural contamination sources were 3567.3, 338.2 and 202 tons, respectively. Thirteen sewage treatment plants (STPs) exist to treat domestic and industrial wastewater in the area. The maximum daily treatment capacity of these STPs is $15 \times 10^4 \text{ m}^3$ [33].

2.2. Monitoring Stations and Water Quality Data

Water samples were collected from 14 monitoring stations (Figure 1) in the Nanxi River every month from January 2015 to November 2020. All the water samples were

collected, stored, transported, and analyzed according to the Technical Specifications Requirements for Monitoring of Surface Water and Waste Water (HJ/T 91-2002) [34] and the Environmental Quality Standards for Surface Water (GB3838-2002) [35] to ensure the quality of the data. The final data were from the Liyang Environmental Protection Bureau.

These water quality data belonged to monthly routine sampling, which reflected the daily water quality status of each monitoring station, but cannot capture the dynamics of pollutants generated by episodic events (e.g., storms and pollution leakage accidents, etc.). Therefore, if possible, additional sampling of water quality before and after episodic events will be required in the future.

We selected 22 water quality parameters for this study. The abbreviations, units and descriptive statistics of them are summarized in Table 1.

Table 1. Water quality parameters, abbreviations, units and descriptive statistics.

Parameter	Abbreviation	Unit	Minimum	Maximum	Mean	S.D.	C.V.
Water temperature	Temp	°C	3.4	34.0	18.1	8.4	0.463
pH	pH		5.74	8.91	7.50	0.38	0.050
Electrical conductivity	EC	ms/m	9.4	181.0	38.0	16.6	0.437
Dissolved oxygen	DO	mg/L	1.8	13.0	6.5	1.9	0.291
Permanganate index	COD _{Mn}	mg/L	1.9	15.7	5.5	1.4	0.261
Biochemical oxygen demand	BOD ₅	mg/L	1.8	9.0	3.6	1.0	0.287
Ammonia nitrogen	NH ₃ -N	mg/L	0.046	31.600	0.916	1.355	1.479
Petroleum	Petrol	mg/L	0.00	0.27	0.06	0.04	0.721
Volatile phenol	VP	mg/L	0.000	0.010	0.002	0.001	0.581
Chemical oxygen demand	COD	mg/L	5.0	87.9	19.1	6.8	0.354
Total nitrogen	TN	mg/L	0.24	37.40	2.87	2.01	0.699
Total phosphorus	TP	mg/L	0.010	0.444	0.131	0.071	0.546
Fluoride	F	mg/L	0.07	0.83	0.34	0.12	0.345
Sulfide	S	mg/L	0.000	0.173	0.046	0.033	0.715
Fecal coliform	F. Coli	CFU/L	360	9130	4226	2429	0.575
Sulphate	SO ₄	mg/L	15.9	53.1	32.1	7.4	0.231
Chloride	Cl	mg/L	0	195	54	30	0.557
Nitrate nitrogen	NO ₃ -N	mg/L	0.020	1.200	0.294	0.127	0.434
Total suspended solids	TSS	mg/L	7	239	33	19	0.574
Total hardness	T-Hard	mg/L	41	255	147	37	0.251
Nitrite nitrogen	NO ₂ -N	mg/L	0.005	0.294	0.063	0.041	0.660
Nonionic ammonia	NH ₃	mg/L	0.001	0.172	0.014	0.019	1.364

Mean represents the mean value; S.D. represents the standard deviation; C.V. represents the coefficient of variation.

2.3. Water Quality Index

WQI is an effective method for water quality assessments [36]. According to the impact of each water quality parameter on human water health and its relative importance in aquatic organisms [15], different weights were assigned in Table 2 [23]. Temp, pH, EC, DO, COD_{Mn}, COD, BOD₅, NH₃-N, TP, TN, Petrol, F. Coli, SO₄, Cl, NO₂-N, NO₃-N, TSS and T-Hard were used, and the measured values were normalized.

The formula used to calculate WQI is as follows:

$$WQI = \frac{\sum_{i=1}^n C_i P_i}{\sum_{i=1}^n P_i} \quad (1)$$

where n is the total number of parameters involved in the calculation, C_i is the normalization factor of parameter i , and P_i is the relative weight of parameter i . The minimum value of P_i is 1, and the maximum weight specified is 4. These values were determined based on previous studies [23,37–39].

The calculated WQI is a dimensionless value from 0 to 100. Based on the WQI scores, surface water quality was divided into five categories [23]: excellent (90–100), good (70–89), moderate (60–69), low (40–59), and bad (0–39).

Table 2. Weights and normalization factors of the parameters used to calculate the water quality index.

Parameters	Units	Relative Weight (P_i)	Normalization Factor (C_i)										
			100	90	80	70	60	50	40	30	20	10	0
Temp	°C	1	21/16	22/15	24/14	26/12	28/10	30/5	32/0	36/−2	40/−4	45/−6	>45/−6
pH		1	7	7–8	7–8.5	7–9	6.5–7	6–9.5	5–10	4–11	3–12	2–13	1–14
EC	ms/m	1	<75	<100	<125	<150	<200	<250	<300	<500	<800	<1200	>1200
DO	mg/L	4	≥7.5	>7	>6.5	>6	>5	>4	>3.5	>3	>2	≥1	<1
COD _{Min}	mg/L	3	<1	<2	<3	<4	<6	<8	<10	<12	<14	≤15	>15
COD	mg/L	3	<10	<15	<16	<18	<20	<25	<30	<34	<37	≤40	>40
BOD ₅	mg/L	3	<2	<3	<3.4	<3.7	<4	<5	<6	<7	<9	≤10	>10
NH ₃ -N	mg/L	3	<0.1	<0.15	<0.3	<0.5	<1	<1.3	<1.5	<1.7	<1.9	≤2	>2
TP	mg/L	1	<0.01	<0.02	<0.05	<0.1	<0.15	<0.2	<0.25	<0.3	<0.35	≤0.4	>0.4
TN	mg/L	2	<0.1	<0.2	<0.35	<0.5	<0.75	<1	<1.25	<1.5	<1.75	≤2	>2
Petrol	mg/L	2	<0.01	<0.02	<0.03	<0.04	<0.05	<0.3	<0.5	<0.7	<0.9	≤1	>1
F. Coli	CFU/L	3	<100	<200	<1000	<2000	<10,000	<15,000	<20,000	<30,000	<35,000	≤40,000	>40,000
SO ₄	mg/L	2	<25	<50	<75	<100	<150	<250	<400	<600	<1000	≤1500	>1500
Cl	mg/L	1	<25	<50	<100	<150	<200	<300	<500	<700	<1000	≤1500	>1500
NO ₂ -N	mg/L	2	<0.005	<0.01	<0.03	<0.05	<0.1	<0.15	<0.2	<0.25	<0.5	≤1	>1
NO ₃ -N	mg/L	2	<0.5	<2	<4	<6	<8	<10	<15	<20	<50	≤100	>100
TSS	mg/L	4	<20	<40	<60	<80	<100	<120	<160	<240	<320	≤400	>400
T-Hard	mg/L	1	<25	<100	<200	<300	<400	<500	<600	<800	<1000	≤1500	>1500

Normalization factors are according to GB3838-2002 and weight, as proposed by Wu et al. (2018) [23].

2.4. Statistical Analysis

The water quality data sets over six years (2015–2020) were checked to eliminate possible missing and abnormal values [40]. The parameter F. Coli had missing values, which were replaced by sequence mean values. Data conforming to normal distribution are needed for most multivariate statistical techniques. Therefore, kurtosis and skewness statistics were analyzed to test whether each water quality parameter conformed to a normal distribution [5,13]. The original data showed that the kurtosis value was between −1.888 and 357.738, and the skewness value was between −1.215 and 16.671, indicating that the raw data were far from the normal distribution. Because most kurtosis and skewness values were greater than 0, the raw data were logarithmically converted ($x = \log_{10}(x)$) [41]. After logarithmic transformation, kurtosis and skewness were in the ranges of −0.640 to 4.577 and −2.279 to 0.096 respectively. To minimize the influence of different units and variances on the parameters, Z-scale standardization (mean value is 0, variance is 1) was performed on the data.

Cluster analysis (CA) was performed on the standardized data to explore the spatial similarity and clustering information on water quality. Principal component analysis/factor analysis (PCA/FA) was performed on the standardized data to explore possible contamination source types [9,28]. Discriminant analysis (DA) was performed on the raw data to extract the important variables reflecting the variations between groups [5]. STATISTICA 10 was used for statistical analysis.

Please refer to Supplementary Materials for the detail about the above multivariate statistical methods.

3. Results and Discussion

The descriptive statistics of 22 water quality parameters are summarized in Table 1. The pH values ranged from 5.74 to 8.91, which were basically within the standard limit of 6–9 allowed by GB3838-2002. The mean values of F, S, F. Coli, and VP in most water samples were far lower than the class III standard (GB3838-2002), while that of Petrol (0.058 mg/L) was slightly higher than the class III standard (0.05 mg/L). Among nutrients, the mean value of TN was 2.87 and far higher than the class III standard (1.0 mg/L); the mean values of NO₃-N and NO₂-N were 0.294 and 0.063 respectively, which were far lower than the class III standard (10 mg/L); the one of NH₃-N was 0.916 and lower than the class III standard (1.0 mg/L). TN is the sum of NO₃-N, NO₂-N, NH₃-N and organic nitrogen, which is the main indicator of water eutrophication. Thus, the main nutrient in the study area was organic nitrogen. The concentration levels of COD_{Min}, BOD₅ and COD deserve

attention because these parameters represent the levels of biological, chemical and organic contamination in surface water, respectively. The maximum values of these parameters were 15.7, 9.0 and 87.9 mg/L, respectively, all exceeding the class III standard (6, 4 and 20 mg/L, respectively). Therefore, the study area had a relatively high contamination level. The coefficients of variation for $\text{NH}_3\text{-N}$, NH_3 , Petrol, S, TN and $\text{NO}_2\text{-N}$ were relatively high, indicating significant temporal and spatial differences in the distributions of these water quality parameters.

3.1. Water Quality Assessment Using WQI

The water quality of most monitoring stations was classified as “medium-low”, accounting for approximately 84.52% (of which “medium” accounted for 64.29% and “low” accounted for 20.24%). Additionally, 13.10% of the water quality was “good”, and only 2.38% was “excellent” (Figure 2). The water quality of S1 and S2 was always above “good”, especially the water quality of S2, which was “excellent” in 2015 and 2018. Because these two monitoring stations are in the Daxi Reservoir and Shahe Reservoir within the urban centralized drinking water protection area, their water quality has been maintained in good condition due to the good natural ecological environment and strict contamination control measures. Other monitoring stations are in urbanized or agricultural areas.

From the interannual change trend of WQI (Figure 3), about half of the monitoring stations showed an increasing trend, most of which were generally stable. The water quality of monitoring stations S1 and S2 decreased slightly. Due to rapid urbanization and population growth, water environment security is facing increased pressure, and the protection of water sources should be further strengthened. The water quality of other monitoring stations showed continuous improvement, especially from 2016 to 2019. Since 2017, Changzhou city has adopted special actions: “two reductions” (reducing total coal consumption and backward chemical production), “six governance” (governing the Taihu Lake water environment, domestic garbage, black and smelly water bodies, livestock and poultry breeding contamination, volatile organic compound contamination, and hidden environmental dangers), and “three improvements” (improving the level of ecological protection, environmental-economic policy regulation, and environmental law enforcement and supervision) [42]. The environmental quality has been significantly improved, the total discharge of major pollutants has been markedly reduced, and the environmental risks have been effectively controlled.

3.2. Spatial Similarities and Clustering

Spatial CA generated a dendrogram, dividing the 14 monitoring stations into 3 clusters at $(D_{link}/D_{max}) \times 100 < 40$ (Figure 4). According to the physical, chemical and microbiological characteristics of water quality, each cluster was classified into its own contamination category. Cluster A included stations S1 and S2 and corresponded to low contamination. Cluster B contained six monitoring stations (S5, S6, S8, S9, S11 and S13) and was classified as medium contamination. Cluster C comprised six monitoring stations (S3, S4, S7, S10, S12 and S14) and was classified as high contamination.

In cluster A, S1 and S2 are in the Daxi Reservoir and Shahe Reservoir. The contamination of the six monitoring stations in cluster B mainly derives from nonpoint source contamination, such as agricultural runoff, livestock and poultry breeding, and fishpond drainage. The monitoring stations of cluster C are mainly located in urban areas and downstream reaches, and the possibility of water contamination is higher because of the comprehensive impacts of domestic sewage, industrial wastewater and upstream inflow water [5,7,10].

The above spatial CA results coincided with the average WQI of the monitoring stations. The WQI values of S1 and S2 were the highest; those of S3, S4, S7, S10, S12 and S14 were relatively low; those of S5, S6, S8, S9, S11 and S13 were at a medium level (Figure 5). Thus, CA can be used to provide reliable water quality classification throughout

monitoring stations; however, designing optimal spatial sampling strategies is warranted in the future [10,28,41,43].

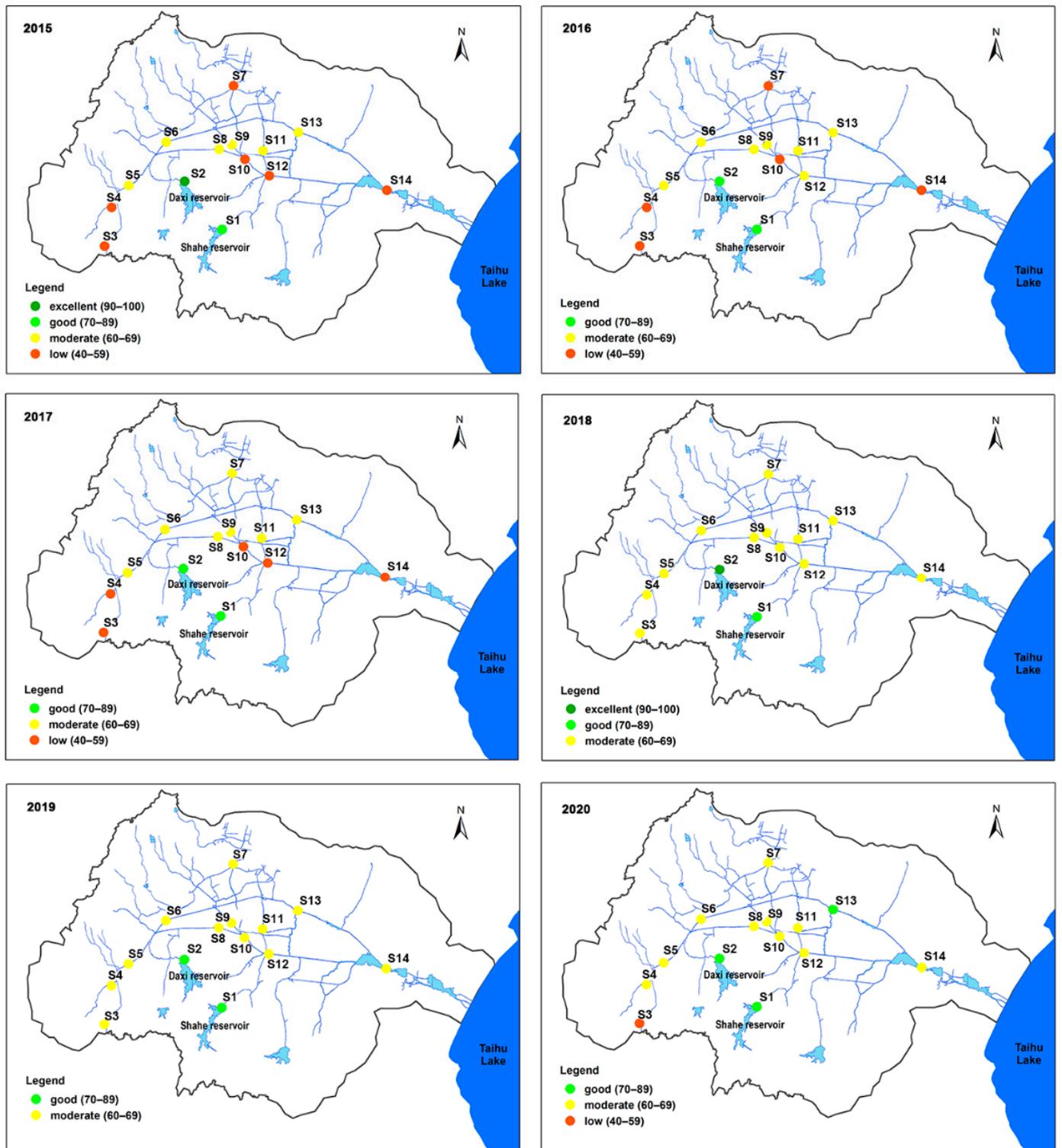


Figure 2. Spatial distribution pattern of the average annual WQI categories in the Nanxi River from 2015 to 2020.

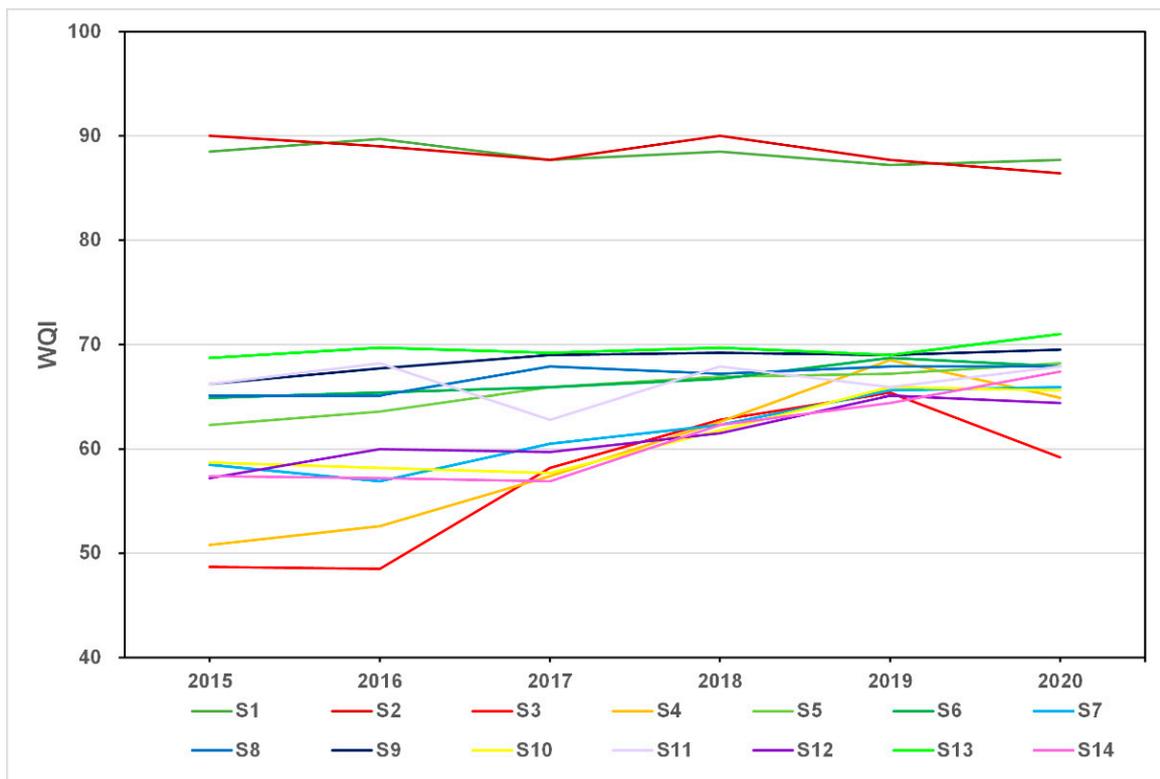


Figure 3. The average annual WQI scores of 14 monitoring stations in the Nanxi River from 2015 to 2020.

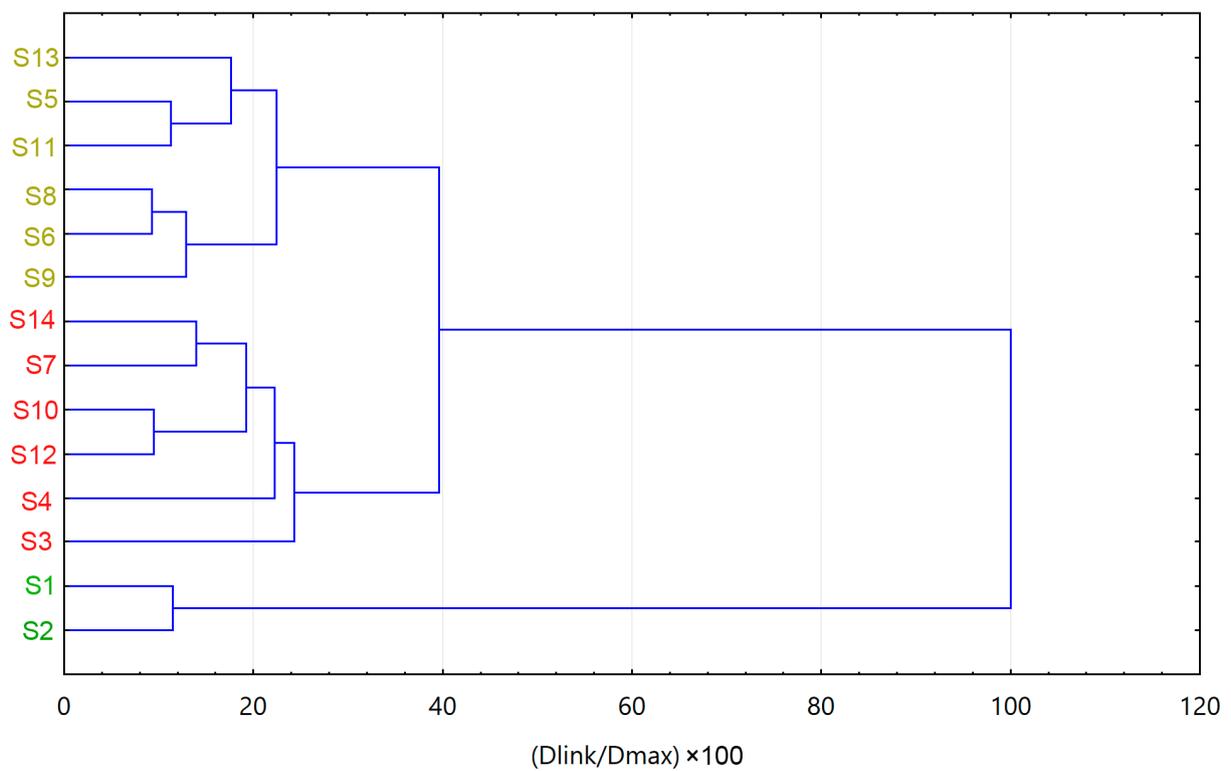


Figure 4. Dendrogram of spatial similarities and clustering of monitoring stations (S1–S14) in the Nanxi River.

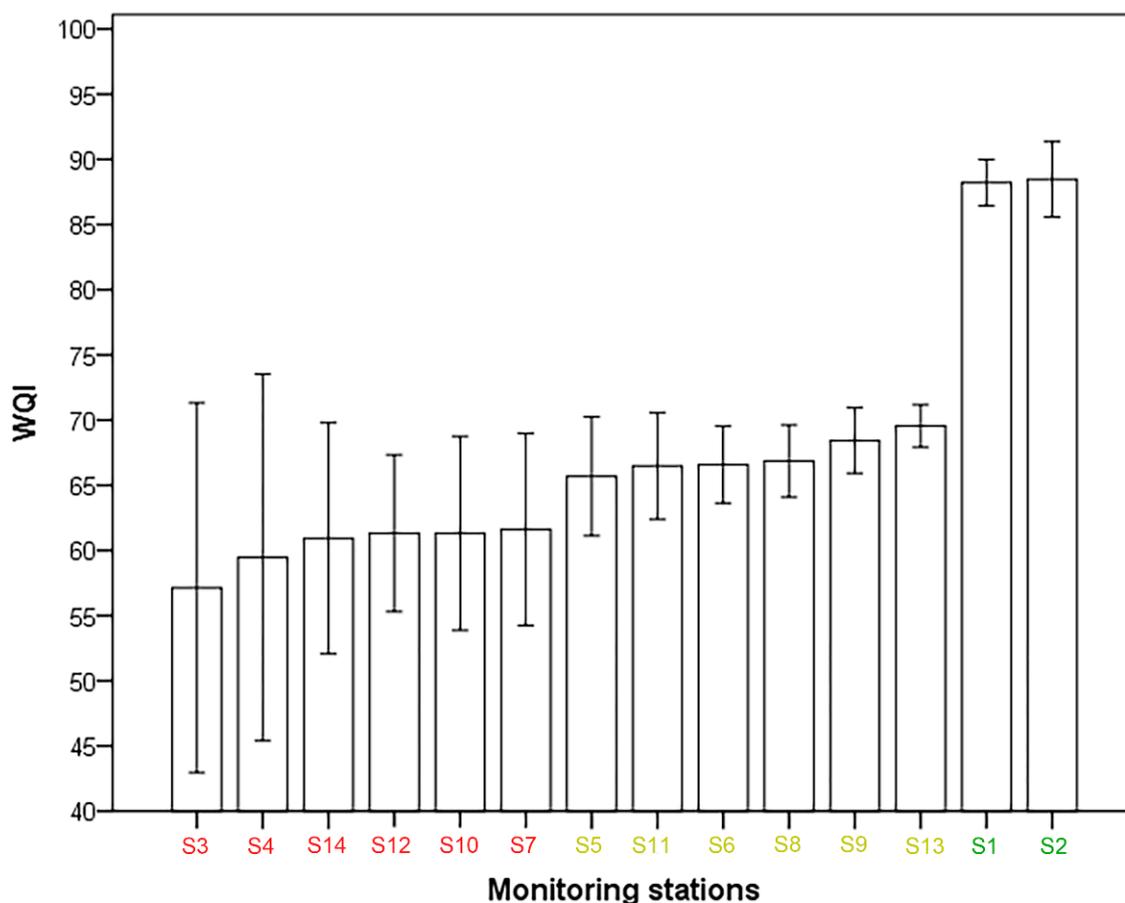


Figure 5. Average WQI of 14 monitoring stations in the Nanxi River.

3.3. Spatial Variations in Water Quality

Based on the CA data, discriminant analysis was used to detect the significance of the discriminant function and to identify the important variables reflecting the variation between clusters. The Wilks’ lambda and chi-square values in all discriminant functions were in the range of 0.036–0.509 and 504.269–2479.317, respectively, and the *p* values were all less than 0.01 (Table 3), indicating that the spatial DA was valid [13].

Table 3. Spatial discriminant analysis results for spatial variations in water quality.

Modes	Discriminant Function	R	Wilks’ Lambda	Chi-Square	<i>p</i> Value
Standard	1	0.963697	0.035502	2478.586	0.00
	2	0.708513	0.498009	517.624	0.00
Forward	1	0.963686	0.035547	2479.317	0.00
	2	0.708173	0.498491	517.255	0.00
Backward	1	0.960275	0.039628	2409.856	0.00
	2	0.700788	0.508897	504.269	0.00

Tables 4 and 5 show the discriminant function and classification matrix generated from the standard, forward stepwise and backward stepwise modes of DA. The standard and forward stepwise models of the discriminant function used 22 and 21 discriminant variables, respectively, and obtained the corresponding classification matrix, which correctly assigned approximately 88% of cases. However, in the backward stepwise mode, DA generated nearly 87% of the correct allocation to the classification matrix using only 14 discrimination parameters. Spatial DA showed that pH, Petrol, VP, COD, TP, F, S, F. Coli, SO₄, Cl, NO₃-N, T-Hard, NO₂-N, and NH₃ were the critical variables to distinguish the water quality of the three spatial clusters and explained most of the spatial variations in expected water quality.

Table 4. Discriminant function coefficients of discriminant analysis for spatial variations in water quality.

Parameters	Standard Mode			Forward Stepwise Mode			Backward Stepwise Mode		
	Cluster A	Cluster B	Cluster C	Cluster A	Cluster B	Cluster C	Cluster A	Cluster B	Cluster C
Temp	2.19	2.15	2.16						
pH	115.00	107.97	107.52	102.81	95.983	95.470	95.99	89.346	88.763
EC	−0.09	−0.10	−0.08	−0.02	−0.035	−0.016			
DO	5.19	4.17	4.10	1.01	0.066	−0.021			
COD _{Mn}	−1.36	1.34	1.28	−2.35	0.373	0.314			
BOD ₅	12.67	11.16	11.64	12.40	10.899	11.372			
NH ₃ -N	9.04	4.81	5.01	8.72	4.494	4.688			
Petrol	−87.09	−63.79	3.49	−107.72	−84.054	−16.893	−29.04	−15.614	53.431
VP	−5928.45	−1499.07	41.86	−3233.54	1148.149	2704.845	−3985.81	552.192	2072.728
COD	1.09	0.60	0.75	0.56	0.081	0.232	1.09	0.769	0.921
TN	−2.78	−0.32	−0.53	−3.21	−0.752	−0.964			
TP	38.74	93.48	99.55	21.98	77.019	82.983	73.23	130.491	137.176
F	7.53	40.57	37.79	8.13	41.154	38.375	8.79	42.410	39.641
S	−50.45	90.85	97.30	−101.99	40.229	46.377	−89.33	50.326	53.812
F. Coli	0.00	0.00	0.00	0.00	0.002	0.002	0.00	0.001	0.001
SO ₄	−0.27	0.47	0.46	0.35	1.070	1.065	0.38	1.196	1.188
Cl	0.10	0.17	0.16	−0.07	0.006	−0.005	−0.08	0.005	−0.002
NO ₃ -N	−3.63	19.19	22.51	−17.28	5.789	9.028	−9.40	13.946	17.819
TSS	0.22	0.26	0.25	0.23	0.268	0.259			
T-Hard	0.44	0.76	0.76	0.30	0.624	0.627	0.33	0.627	0.630
NO ₂ -N	−46.64	50.99	33.76	−11.68	85.329	68.309	22.86	110.897	95.396
NH ₃	−1329.83	−1190.28	−1191.98	−936.82	−804.223	−803.626	−838.04	−706.073	−697.089
Constant	−522.73	−554.90	−562.82	−433.50	−468.801	−475.693	−397.88	−434.420	−440.556

Table 5. Classification matrix of discriminant analysis for spatial variations in water quality.

Monitoring Stations	Percent Correct	Stations Assigned by Discriminant Analysis		
		Cluster A	Cluster B	Cluster C
Standard mode				
Cluster A	100.0000	108	0	0
Cluster B	95.6790	0	310	14
Cluster C	77.4691	0	73	251
Total	88.4921	108	383	265
Forward stepwise mode				
Cluster A	100.0000	108	0	0
Cluster B	95.3704	0	309	15
Cluster C	76.8518	0	75	249
Total	88.0952	108	384	264
Backward stepwise mode				
Cluster A	100.0000	108	0	0
Cluster B	94.7531	0	307	17
Cluster C	75.3086	0	80	244
Total	87.1693	108	387	261

Based on the discriminant parameters analyzed by DA, box and whisker plots of three clusters (cluster A, cluster B, cluster C) were constructed to evaluate the spatial variations in water quality (Figure 6). Most of the parameters showed significant differences between clusters. Overall, the average concentration of cluster A was much lower than that of clusters B and C, and the average concentration of cluster C was slightly higher than that of cluster B. Higher Petrol, COD and TP values were found in cluster C, indicating that organic contamination and eutrophication were the most serious water environment problems in cluster C. Additionally, lower pH values were found at the monitoring stations of cluster C, likely because of the hydrolysis of acidic substances (ammonia and organic acids) [5]. In

conclusion, the water contamination of cluster C was more serious than that of the other two clusters. Thus, the prevention and control of contamination sources and treatment capacity of point source contamination must be strengthened, such as strengthening the construction and treatment capacity of STPs.

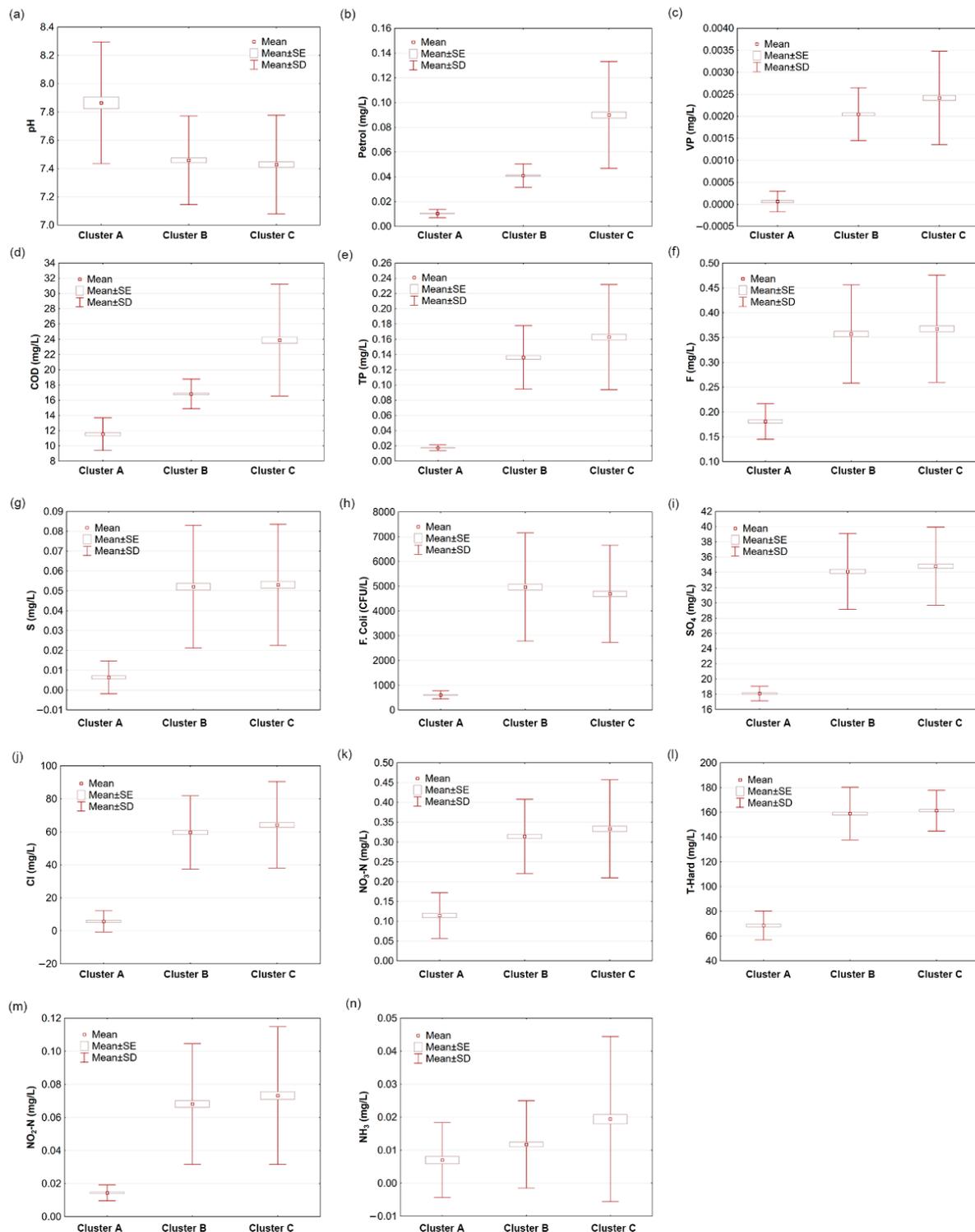


Figure 6. Spatial variations in water quality in three spatial clusters: (a) pH, (b) Petrol, (c) VP, (d) COD, (e) TP, (f) F, (g) S, (h) F. Coli, (i) SO₄, (j) Cl, (k) NO₃-N, (l) T-Hard, (m) NO₂-N, and (n) NH₃ in the Nanxi River.

3.4. Principal Component Determination and Contamination Source Identification

Because the contamination levels of the three spatial clusters (clusters A, B, and C) were significantly different, PCA/FA was used to identify the water contamination source types for the normalized data sets of the three spatial clusters.

PCA/FA of the three data matrices obtained six, eight and seven variance factors (VFs) with eigenvalues ≥ 1 , explaining 71.5%, 66.8% and 67.9% of the total variance in the corresponding data sets, respectively (Tables 6–8). Additionally, the loadings of parameters on VFs were categorized as “high”, “medium” and “low” based on absolute loading values of > 0.75 , $0.75\text{--}0.50$ and $0.50\text{--}0.30$ [44].

Table 6. Loadings of 22 water quality parameters on VFs for cluster A in the Nanxi River.

Parameters	VF1	VF2	VF3	VF4	VF5	VF6
Temp	0.145	−0.110	0.168	0.351	0.665	0.214
pH	0.043	0.182	0.026	0.075	0.764	0.386
EC	0.073	0.190	0.168	0.113	0.744	−0.335
DO	−0.368	0.047	−0.319	−0.286	−0.223	0.287
COD _{Mn}	−0.198	−0.084	−0.024	0.808	0.243	−0.017
BOD ₅	−0.068	0.103	0.130	0.852	0.101	−0.015
NH ₃ -N	0.091	−0.313	−0.489	0.093	−0.049	− 0.568
Petrol	0.829	−0.066	0.060	−0.128	0.134	0.011
VP	0.891	0.022	0.119	−0.160	0.130	0.060
COD	−0.180	−0.376	0.546	−0.032	−0.189	−0.091
TN	0.724	0.014	0.199	−0.144	0.038	−0.273
TP	0.074	−0.107	−0.618	0.347	0.141	−0.024
F	0.767	−0.141	0.006	0.092	−0.081	0.014
S	0.003	0.091	0.735	0.248	−0.138	−0.153
F. Coli	0.566	0.529	0.321	0.048	−0.004	−0.306
SO ₄	−0.011	− 0.750	−0.115	0.239	0.144	−0.084
Cl	0.238	0.046	0.889	0.061	0.041	0.084
NO ₃ -N	0.133	0.092	0.910	0.007	0.074	0.022
TSS	0.589	−0.061	−0.162	0.403	−0.188	0.395
T-Hard	0.205	− 0.826	0.096	−0.128	−0.243	0.022
NO ₂ -N	0.038	0.748	0.255	0.045	0.273	0.126
NH ₃	0.019	0.097	−0.220	0.062	0.884	−0.113
Eigenvalue	4.554	3.373	3.067	2.266	1.380	1.089
% Total variance	20.699	15.332	13.939	10.302	6.271	4.952
Cumulative % variance	20.699	36.031	49.969	60.271	66.542	71.494

VFs represent the variance factors after varimax raw rotation for principal components; bold values represent medium-high loadings.

Among the six VFs of cluster A, VF1 explained 20.7% of the total variance and had high positive loadings on Petrol, VP, TN and F. This factor indicated toxic organic contamination from farmland drainage, oily sewage discharge from ship operation, domestic sewage, industrial wastewater, atmospheric deposition and precipitation leaching. VF2 (15.3% of the total variance) had high negative loadings on SO₄ and TSS, and high positive loadings on NO₂-N. The presence of nitrite in water indicated that the decomposition process of organic matter continued, and the risk of organic matter contamination persisted. VF3 (13.9%) had high positive loadings on NO₃-N, Cl and S, indicating nutrients from agricultural runoff and atmospheric deposition and the natural source of soil erosion and salt ions (Cl, S) in the watershed [45]. VF4 (10.3%) had high positive loadings on BOD₅ and COD_{Mn}, representing organic contamination in sewage [6]. VF5 (6.3%) had high positive loadings on NH₃, pH and EC. Generally, EC indicates natural contamination, which may be due to soil erosion or an increase in the number of salt ions in water [44]. Additionally, VF6 (only 5.0%) had a medium negative loading on NH₃-N.

Table 7. Loadings of 22 water quality parameters on VFs for cluster B in the Nanxi River.

Parameters	VF1	VF2	VF3	VF4	VF5	VF6	VF7	VF8
Temp	−0.063	0.109	− 0.865	0.041	−0.052	−0.100	0.099	0.153
pH	0.042	0.021	0.220	−0.234	−0.117	0.125	0.077	0.787
EC	−0.214	0.036	−0.180	−0.032	−0.062	0.823	−0.024	0.046
DO	0.024	0.073	0.750	0.086	−0.126	−0.378	0.066	0.055
COD _{Mn}	0.191	− 0.847	0.041	−0.076	0.028	−0.006	−0.074	0.015
BOD ₅	−0.032	− 0.863	0.013	0.174	0.108	−0.020	0.041	−0.075
NH ₃ -N	0.211	−0.222	−0.088	−0.038	0.703	0.085	−0.019	−0.038
Petrol	−0.216	−0.357	0.073	0.067	0.629	−0.159	0.131	0.083
VP	0.074	0.141	−0.134	− 0.647	−0.073	−0.134	0.358	0.181
COD	−0.161	0.132	0.157	0.225	0.520	0.004	0.082	0.104
TN	−0.021	0.009	0.193	0.014	0.142	0.090	0.719	0.061
TP	0.049	−0.135	−0.160	0.621	0.164	0.046	−0.064	−0.073
F	− 0.810	0.169	−0.006	−0.117	−0.011	0.086	0.061	−0.010
S	0.706	0.171	0.121	−0.335	0.093	−0.062	−0.090	0.199
F. Coli	0.707	−0.178	0.010	0.400	−0.226	−0.052	0.095	0.061
SO ₄	−0.006	0.028	−0.166	−0.111	−0.103	0.007	0.786	0.066
Cl	0.392	0.026	0.354	0.086	0.161	0.568	0.291	0.061
NO ₃ -N	−0.185	0.013	0.064	− 0.664	0.038	0.314	−0.173	−0.140
TSS	−0.141	0.218	−0.169	−0.047	0.544	0.052	−0.206	−0.103
T-Hard	−0.093	−0.130	0.050	−0.158	0.035	0.202	0.131	−0.663
NO ₂ -N	0.165	0.085	0.189	0.572	−0.046	0.089	−0.490	0.143
NH ₃	0.081	−0.024	−0.355	0.119	0.135	0.045	0.092	0.812
Eigenvalue	2.837	2.440	2.038	1.980	1.548	1.419	1.297	1.146
% Total variance	12.897	11.089	9.262	8.998	7.038	6.452	5.897	5.209
Cumulative % variance	12.897	23.986	33.248	42.246	49.284	55.735	61.632	66.841

VFs represent the variance factors after varimax raw rotation for principal components; bold values represent medium-high loadings.

Regarding the data set of cluster B, among the eight VFs, VF1, which accounted for 12.9% of the total variance, represented a high negative loading on F but medium positive loadings on F. Coli and S, indicating microbial contamination from municipal sewage, livestock and poultry breeding. VF2 (11.1% of the total variance) represented high negative loadings on BOD₅ and COD_{Mn}, indicating organic contamination in urban sewage and industrial wastewater. VF3 (9.3%) represented a high positive loading on DO but a high negative loading on Temp. VF4 (9.0%) represented only a moderate positive loading on TP, revealing nutrient contamination (e.g., P), especially from sewage containing detergents, industrial wastewater and fertilizer. Point source contamination (such as wastewater from the phosphorus chemical industry) and nonpoint source contamination (such as animal breeding and agricultural fertilizer) from P, constitute common eutrophication-causing contamination in this area [46]. VF5 (7.0%) applied only a moderate positive loading on NH₃-N, representing the contamination of animal feces and agricultural fertilizers. VF6 (only 6.5%) presented a high positive loading only on EC, likely because of the mineral composition in river water [6]. VF7 (only 5.9%) presented a high positive loading on SO₄ and a medium positive loading on TN, representing industrial wastewater using sulfate or sulfuric acid. Finally, VF8 (only 5.2%) had a high positive loading on NH₃ and pH, likely because of industrial wastewater containing alkaline substances, such as NH₃.

Regarding the seven VFs of cluster C, VF1 (20.5% of the total variance) showed high positive loadings on NH₃-N and TN, representing nutrient contamination from agricultural runoff, municipal sewage and fertilizer plant wastewater (e.g., N). VF2 (12.2%) showed a high positive loading on F, representing industrial wastewater containing fluoride. VF3 (10.2%) showed a high positive loading on pH. VF4 (7.5%) showed a high positive loading on Temp and a moderate negative loading on DO, contrasting the results for VF3 of cluster B. VF5 (6.7%) showed a high positive loading on Petrol, representing contamination from oily sewage discharge from ship operations and wastewater from the petrochemical industry. VF6 (5.7) showed moderate positive loadings on TSS and EC. Agricultural runoff,

wastewater discharge, solid waste disposal and irrigation return increased the suspended solids loading in streams [45]. VF7 (5.0%) showed a high positive loading on SO₄, similar to VF7 of cluster B.

Table 8. Loadings of 22 water quality parameters on VFs for cluster C in the Nanxi River.

Parameters	VF1	VF2	VF3	VF4	VF5	VF6	VF7
Temp	−0.189	0.073	0.090	0.862	−0.116	−0.034	0.125
pH	−0.158	0.007	0.781	−0.105	−0.132	−0.052	0.135
EC	0.250	0.211	−0.077	0.206	−0.068	0.704	−0.077
DO	−0.216	0.030	−0.051	−0.743	−0.118	−0.155	0.090
COD _{Mn}	0.593	−0.065	0.051	0.032	0.414	−0.285	0.186
BOD ₅	0.534	−0.030	0.056	0.142	0.655	−0.255	0.098
NH ₃ -N	0.829	−0.001	−0.085	0.025	0.175	0.135	−0.189
Petrol	0.114	−0.105	0.086	−0.108	0.819	0.055	0.064
VP	0.001	−0.100	0.155	0.004	−0.632	−0.283	0.139
COD	0.546	−0.013	0.294	−0.245	0.472	0.075	0.071
TN	0.827	0.127	−0.065	−0.055	0.075	0.099	−0.001
TP	0.302	0.110	−0.201	0.061	0.648	0.167	−0.116
F	0.115	0.759	0.076	−0.063	0.044	0.214	0.200
S	−0.082	−0.673	0.355	−0.245	−0.133	0.154	−0.048
F. Coli	0.045	−0.721	−0.077	0.028	0.255	−0.249	0.071
SO ₄	0.053	0.197	0.074	0.080	−0.178	0.022	0.818
Cl	0.182	−0.435	0.133	−0.133	0.157	0.489	0.315
NO ₃ -N	0.262	0.268	−0.213	−0.041	−0.291	−0.013	−0.632
TSS	−0.106	0.174	−0.067	−0.033	0.153	0.706	0.091
T-Hard	0.148	0.105	−0.659	−0.187	−0.140	0.086	0.031
NO ₂ -N	0.191	−0.147	0.046	0.013	0.631	−0.043	−0.438
NH ₃	0.312	−0.001	0.689	0.404	0.104	0.007	0.135
Eigenvalue	4.513	2.685	2.250	1.661	1.469	1.252	1.098
% Total variance	20.515	12.202	10.227	7.549	6.678	5.693	4.989
Cumulative % variance	20.515	32.717	42.944	50.493	57.171	62.863	67.852

VFs represent the variance factors after varimax raw rotation for principal components; bold values represent medium-high loadings.

We have identified four contamination source types—nutrient, organics, feces and oil. Specifically, nutrient represented point source contamination, such as urban domestic wastewater and industrial wastewater from chemical fertilizer plants, and nonpoint source contamination, such as that related to agricultural activities and aquaculture. Second, organics were mainly derived from oxygen consumption and toxic organic matter from municipal sewage and industrial sewage. Third, feces were mainly derived from animal fecal drainage in the fishery and livestock breeding industries. Finally, oil represented the contamination characters from the petroleum chemical industry and oily sewage discharge from ship operation.

4. Conclusions

In the Nanxi River of the Taihu watershed in China, WQI and multivariate statistical techniques were used to assess the spatiotemporal variations in water quality and to identify contamination source types.

(1) The WQI findings indicated that the water quality of most monitoring stations was classified as “medium-low” and presented a continuous improvement trend. The water quality of S1 and S2 was always above “good”, especially the water quality of S2, which was “excellent” in 2015 and 2018.

(2) Cluster analysis divided the 14 monitoring stations into 3 clusters of low contamination, medium contamination and high contamination.

(3) Discriminant analysis used 14 parameters (pH, Petrol, VP, COD, TP, F, S, F. coli, SO₄, Cl, NO₃-N, T-Hard, NO₂-N, and NH₃) for important data reduction and provided an 87% correct allocation in the spatial variation analysis for the 3 clusters.

(4) PCA/FA was used to analyze the data sets of three spatial clusters and obtained six, eight and seven potential factors. The study showed that the sources of water contamination were mainly related to nutrients (livestock and poultry breeding, agricultural activities), salt ions (natural) and toxic organic contamination (urban sewage, industrial wastewater and ship operation) in cluster A; fecal coliform (livestock and poultry breeding), organic contamination (industrial and domestic sewage), temperature (natural), nutrients (point source: industrial wastewater and domestic sewage; nonpoint sources: livestock and poultry breeding, agricultural fertilizer) in cluster B; and fluoride (industrial wastewater), pH and temperature (natural), and petroleum (ship operation and industrial wastewater) in cluster C.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/w14050778/s1>, Supplementary Material.docx (the explanation about the statistical methods) and Supplementary Material.xlsx (the raw water quality data).

Author Contributions: Conceptualization, Z.-M.Z. and F.Z.; methodology, Z.-M.Z.; software, Z.-M.Z.; validation, F.Z.; formal analysis, Z.-M.Z.; investigation, D.-C.C.; resources, D.-C.C.; data curation, D.-C.C.; writing—Original draft preparation, Z.-M.Z.; writing—Review and editing, F.Z.; visualization, Z.-M.Z.; supervision, J.-L.D.; project administration, J.-L.D.; funding acquisition, F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number U1603241, U2003205.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in Supplementary Material.xlsx.

Acknowledgments: The authors would like to thank the reviewers for providing helpful suggestions to improve this manuscript.

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

References

1. Singh, K.; Malik, A.; Sinha, S. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—A case study. *Anal. Chim. Acta* **2005**, *538*, 355–374. [[CrossRef](#)]
2. Carpenter, S.R.; Caraco, N.F.; Correll, D.L.; Howarth, R.W.; Sharpley, A.N.; Smith, V.H. Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* **1998**, *8*, 559–568. [[CrossRef](#)]
3. Zhai, X.; Xia, J.; Zhang, Y. Water quality variation in the highly disturbed Huai River Basin, China from 1994 to 2005 by multi-statistical analyses. *Sci. Total Environ.* **2014**, *496C*, 594–606. [[CrossRef](#)] [[PubMed](#)]
4. Sundaray, K.S.; Panda, U.C.; Nayak, B.B.; Bhatta, D. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of the Mahanadi river–estuarine system (India)—A case study. *Environ. Geochem. Health* **2006**, *28*, 317–330. [[CrossRef](#)]
5. Xu, Y.; Xie, R.; Wang, Y.; Jian, S. Spatio-temporal variations of water quality in Yuqiao Reservoir Basin, North China. *Front. Env. Sci. Eng.* **2015**, *9*, 649–664. [[CrossRef](#)]
6. 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](#)]
7. 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](#)]
8. Vega, M.; Pardo, R.; Barrado, E.; Debn, L. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res.* **1998**, *32*, 3581–3592. [[CrossRef](#)]
9. Muangthong, S.; Shrestha, S. Assessment of surface water quality using multivariate statistical techniques: Case study of the Nampong river and Songkhram river, Thailand. *Environ. Monit. Assess.* **2015**, *187*, 548. [[CrossRef](#)]
10. Alberto, W.D.; Pilar, D.D.; Valeria, A.M.; Fabiana, P.S.; Cecilia, H.A.; De, B.M. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquía River Basin (Córdoba-Argentina). *Water Res.* **2001**, *35*, 2881–2894. [[CrossRef](#)]
11. Dixon, W.; Chiswell, B. Review of aquatic monitoring program design. *Water Res.* **1996**, *30*, 1935–1948. [[CrossRef](#)]

12. Zhou, F.; Guo, H.; Liu, Y.; Jiang, Y. Chemometrics data analysis of marine water quality and source identification in Southern Hong Kong. *Mar. Pollut. Bull.* **2007**, *54*, 745–756. [CrossRef] [PubMed]
13. Zhang, Y.; Guo, F.; Meng, W.; Wang, X. Water quality assessment and source identification of Daliao river basin using multivariate statistical methods. *Environ. Monit. Assess.* **2009**, *152*, 105–121. [CrossRef]
14. He, W.; Li, F.; Yu, J.; Chen, M.; Deng, Y.; Li, J.; Tang, X.; Chen, Z.; Yan, Z. Risk assessment and source apportionment of trace elements in multiple compartments in the lower reach of the Jinsha River, China. *Sci. Rep.* **2021**, *11*, 20041. [CrossRef] [PubMed]
15. Pesce, S.F.; Wunderlin, D.A. Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquia River. *Water Res.* **2000**, *34*, 2915–2926. [CrossRef]
16. Adimalla, N.; Li, P.; Venkatayogi, S. Hydrogeochemical evaluation of groundwater quality for drinking and irrigation purposes and integrated interpretation with water quality index studies. *Environ. Process.* **2018**, *5*, 363–383. [CrossRef]
17. Alver, A. Evaluation of conventional drinking water treatment plant efficiency according to water quality index and health risk assessment. *Environ. Sci. Pollut. R.* **2019**, *26*, 27225–27238. [CrossRef]
18. Hurley, T.; Sadiq, R.; Mazumder, A. Adaptation and evaluation of the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) for use as an effective tool to characterize drinking source water quality. *Water Res.* **2012**, *46*, 3544–3552. [CrossRef]
19. Şener, Ş.; Şener, E.; Davraz, A. Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey). *Sci. Total Environ.* **2017**, *584–585*, 131–144. [CrossRef]
20. Hoseinzadeh, E.; Khorsandi, H.; Wei, C.; Alipour, M. Evaluation of Aydughmush River water quality using the National Sanitation Foundation Water Quality Index (NSFWQI), River Pollution Index (RPI), and Forestry Water Quality Index (FWQI). *Desalin. Water Treat.* **2015**, *54*, 2994–3002. [CrossRef]
21. Suratman, S.; Sailan, M.I.M.; Hee, Y.Y.; Bedurus, E.A.; Latif, M.T. A preliminary study of water quality index in Terengganu River Basin, Malaysia. *Sains Malays.* **2015**, *44*, 67–73. [CrossRef]
22. Sun, W.; Xia, C.; Xu, M.; Guo, J.; Sun, G. Application of modified water quality indices as indicators to assess the spatial and temporal trends of water quality in the Dongjiang River. *Ecol. Indic.* **2016**, *66*, 306–312. [CrossRef]
23. Wu, Z.; Wang, X.; Chen, Y.; Cai, Y.; Deng, J. Assessing river water quality using water quality index in Lake Taihu Basin, China. *Sci. Total Environ.* **2018**, *612*, 914–922. [CrossRef] [PubMed]
24. Wu, H.; Yang, W.; Yao, R.; Zhao, Y.; Zhao, Y.; Zhang, Y.; Yuan, Q.; Lin, A. Evaluating surface water quality using water quality index in Beiyun River, China. *Environ. Sci. Pollut. R.* **2020**, *27*, 35449–35458. [CrossRef]
25. Helena, B.; Pardo, R.; Vega, M.; Barrado, E.; Fernandez, J.M.; Fernandez, L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res.* **2000**, *34*, 807–816. [CrossRef]
26. Liu, C.; Lin, K.; Kuo, Y. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Sci. Total Environ.* **2003**, *313*, 77–89. [CrossRef]
27. Reghunath, R.; Murthy, T.R.S.; Raghavan, B.R. The utility of multivariate statistical techniques in hydrogeochemical studies: An example from Karnataka, India. *Water Res.* **2002**, *36*, 2437–2442. [CrossRef]
28. Simeonov, V.; Stratis, J.A.; Samara, C.; Zachariadis, G.; Voutsas, B.; Anthemidis, A.; Sofoniou, M.; Kouimtzi, T. Assessment of the surface water quality in Northern Greece. *Water Res.* **2003**, *37*, 4119–4124. [CrossRef]
29. Bu, H.; Tan, X.; Li, S.; Zhang, Q. Temporal and spatial variations of water quality in the Jinshui River of the South Qinling Mts., China. *Ecotox. Environ. Safe.* **2010**, *783*, 907–913. [CrossRef]
30. Alexakis, D. Assessment of water quality in the Messolonghi–Etoliko and Neochorio region (West Greece) using hydrochemical and statistical analysis methods. *Environ. Monit. Assess.* **2011**, *182*, 397–413. [CrossRef]
31. Zhang, Z.; Zhang, F.; Du, J.; Chen, D.; Zhang, W. Impacts of land use at multiple buffer scales on seasonal water quality in a reticular river network area. *PLoS ONE* **2021**, *16*, e0244606. [CrossRef] [PubMed]
32. Standardization Administration of China. *Classification and Codes for Chinese Soil (GB/T 17296-2009)*, 3rd ed.; Standards Press of China: Beijing, China, 2009; pp. 2–113.
33. Liyang Statistical Yearbook. Available online: <http://www.liyang.gov.cn/class/KPAQQDDJ#ly4> (accessed on 11 February 2022).
34. Technical Specifications Requirements for Monitoring of Surface Water and Waste Water. Available online: https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/jcffbz/200301/t20030101_66890.htm (accessed on 11 February 2022).
35. Environmental Quality Standards for Surface Water. Available online: https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/shjbh/shjzlbz/200206/t20020601_66497.shtml (accessed on 11 February 2022).
36. Zhao, Y.; Song, Y.; Cui, J.; Gan, S.; Yang, X.; Wu, R.; Guo, P. Assessment of water quality evolution in the Pearl River estuary (South Guangzhou) from 2008 to 2017. *Water* **2020**, *12*, 59. [CrossRef]
37. Debels, P.; Figueroa, R.; Urrutia, R.; Barra, R.; Niell, X. Evaluation of water quality in the Chillán river (Central Chile) using physicochemical parameters and a modified water quality index. *Environ. Monit. Assess.* **2005**, *110*, 301–322. [CrossRef] [PubMed]
38. Kocer, M.A.T.; Sevgil, H. Parameters selection for water quality index in the assessment of the environmental impacts of land-based trout farms. *Ecol. Indic.* **2014**, *36*, 672–681. [CrossRef]
39. Zhao, P.; Tang, X.; Tang, J.; Wang, C. Assessing water quality of three gorges reservoir, China, over a five-year period from 2006 to 2011. *Water Resour. Manag.* **2013**, *27*, 4545–4558. [CrossRef]
40. Li, B.; Wan, R.; Yang, G.; Wang, S.; Wagner, P. Exploring the spatiotemporal water quality variations and their influencing factors in a large floodplain lake in China. *Ecol. Indic.* **2020**, *115*, 106454. [CrossRef]

41. 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](#)]
42. Changzhou “263” Special Action Topic. Available online: http://www.changzhou.gov.cn/ns_class/cz263zt (accessed on 12 January 2022).
43. Bouza-Deao, R.; Ternero-Rodríguez, M.; Fernández-Espinosa, A.J. Trend study and assessment of surface water quality in the ebro river (spain). *J. Hydrol.* **2008**, *361*, 227–239. [[CrossRef](#)]
44. Ogwueleka, T.C. Use of multivariate statistical techniques for the evaluation of temporal and spatial variations in water quality of the Kaduna River, Nigeria. *Environ. Monit. Assess.* **2015**, *187*, 137. [[CrossRef](#)]
45. Varol, M.; Gökot, B.; Bekleyen, A.; Şen, B. Water quality assessment and apportionment of pollution sources of Tigris River (Turkey) using multivariate statistical techniques—A case study. *River Res. Appl.* **2012**, *28*, 1428–1438. [[CrossRef](#)]
46. Wang, Y.; Wang, P.; Bai, Y.; Tian, Z.; Li, J.; Shao, X.; Mustavich, L.; Li, B. Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China. *J. Hydro-Environ. Res.* **2013**, *7*, 30–40. [[CrossRef](#)]