



Multivariate Assessment of Surface Water Quality and Pollution Dynamics Using Physicochemical and Nutrient Indicators

Abodunde Halleluyah Joshua¹, Oladele Olabisi Excel²

¹Ladoke Akintola university of technology. Email: abodundehalleluyah614@gmail.com

²LAUTECH. ORCID ID: 0009-0004-6944-7553. Email: rufusexcel@gmail.com

Article History

Received Date: 06/11/2025

Revised Date: 21/01/2026

Accepted Date: 10/02/2026

Published Date: 17/02/2026

Abstract

Surface-water systems integrate natural geochemical processes, seasonal hydrology, catchment inputs, and anthropogenic pressures, making multivariate evaluation essential for identifying dominant water-quality patterns. This study assessed physicochemical, nutrient, organic-pollution, and microbiological indicators across three monitoring locations and three seasonal classes. Twelve primary indicators were evaluated through descriptive statistics, Kruskal-Wallis tests with Benjamini-Hochberg adjustment, epsilon-squared effect sizes, Spearman rank correlation, and standardized principal component analysis (PCA). Spatial differences were significant for all twelve indicators after multiple-testing correction, with particularly large effects for dissolved oxygen, phosphate, ammonia nitrogen, nitrite nitrogen, temperature, and nitrate nitrogen. Seasonal variation was comparatively weak; only specific conductivity remained significant after false-discovery-rate correction. Strong positive correlations linked ammonia with nitrite, nitrite with phosphate, conductivity with total coliform, and biochemical oxygen demand with total coliform. PCA showed that the first three components explained 73.03% of total standardized variance, with the first two accounting for 61.53%. The results demonstrate that spatial heterogeneity was the dominant organizing feature of water quality, while nutrient and organic-pollution variables formed coherent multivariate gradients. The approach provides a reproducible framework for prioritizing monitoring indicators and interpreting pollution dynamics in surface-water systems.

Keywords:

- surface water quality,
- physicochemical indicators
- nutrient pollution
- multivariate analysis
- principal component analysis

1. Introduction

Surface-water quality is integral to ecological integrity, public health, agriculture, and regional development and can be dynamic in response to multiple natural and anthropogenic stressors. River chemistry is dependent on catchment geology, rainfall, runoff, wastewater discharge, agro-industry inputs and urban development, and assessment should consider physical, chemical, nutrient and biological signals. Use of both physicochemical and microbiological monitoring is able to differentiate between relatively stable conditions and the presence of new contamination pressures (Debassi et al., 2025). Physicochemical variables are especially useful as changes in pH, conductivity, turbidity, oxygen and related parameters frequently indicate general disturbance of the environment (Adestina et al., 2024). Spatial and temporal heterogeneity also hinders interpretation: an overall mean can mask the presence of local pollution regimes.

Multivariate analysis has revealed that water quality can be affected in various ways by site and season. The research in Lake Ziway revealed that spatial and temporal assessment can uncover patterns not captured by univariate summaries (Tibebe et al., 2022) and research in the Vietnamese Mekong Delta highlighted the importance of integrated statistical analysis for addressing the geographic and temporal variability (Hong, Viet, et al., 2022). Giao et al. (2022) also found significant spatio-temporal variation in the physicochemical and biological characteristics in Vinh Long Province. In addition, assessment of the tropical river at the Southern Western Ghats revealed that the spatial and temporal change can be combined and understood through a combination of the physicochemical indices and multivariate techniques (Saha et al., 2022). These results lend themselves to analyses that explicitly compare locations and seasons, as opposed to one or the other being dominant.

Nutrient enrichment is a key component of surface-water pollution, as the nutrients ammonia, nitrite, nitrate, and phosphate can be related to decomposition, fertilizer inputs, sediment interactions, runoff and sewage. The distribution of nutrient elements has been documented in surface waters of Tibet in both spatial and temporal dimensions, suggesting the nutrient elements should be assessed along with other parameters of water pollution (Hong, Zhang, et al., 2022). Multivariate patterns have also been found to show environmental change that is not evident from individual variables in seasonal water-quality studies (Hammoumi et al., 2024). Similar seasonal assessments of coastal waters showed that the interpretation of the physicochemical variability should be made in the hydrological context (Singh et al., 2022), and the seasonal pollution assessment of river and sediment in Northern Lebanon highlighted the value of using multivariate methods (Mansour et al., 2024).

Multivariate statistics offer a convenient way to reduce the number of correlated water-quality measurements. Physicochemical parameters and multivariate analysis can help differentiate the main pollution sources, and the interpretation can be enhanced (Zhang, Li, et al., 2022a). Common gradients can be summarized to fewer dimensions by PCA and related procedures without losing a lot of the information. Multivariate methods are used when assessing water quality of the Nanxi River (Zhang, Zhang, et al., 2022b); multivariate analysis has been used to distinguish water-quality structure in the Quang Binh River system (Do Thi et al., 2025). In Southern Nigeria, hydrochemical assessment also indicated that multivariate methods can be used to show a coherent surface-water dynamics in a complex terrain (Ubuoh et al., 2024).

Integrated analysis is particularly useful in situations where several hydrological, meteorological and anthropogenic drivers are in play. Multivariate analysis in a highly eutrophic reservoir revealed relationships between water-quality changes and hydrological and environmental inputs (Abdul Wahid & Arunbabu, 2025). Multifaceted statistical analysis has been able to provide evidence of anthropogenic influence, even in data poor catchments (Kanownik et al., 2024). Despite this advancement, many studies remain focused on either spatial contrasts, seasonal changes or isolated indices, but not all three in a single model—nested group comparisons, nutrient pollution relationships, and dimension reduction. The present study sought to examine therefore the levels of the physicochemical, nutrient, oxygen-demand and microbiological indicators at monitoring locations over the seasons to identify the major source of variation, quantify relationships between indicators, and identify multivariate pollution gradients using statistical techniques appropriate for a small environmental monitoring dataset. This incorporation of emphasis is crucial, as management decisions are typically based on the ability to determine recurring trends, not single measurements. A framework that compares group structure and then looks at correlated indicators would be able to separate the general change in the environment (an environmental organization) from the change in a particular parameter (parameter fluctuations), and give a better basis for future source-oriented monitoring.

Research objectives

- To characterize spatial and seasonal variations in physicochemical, nutrient, and pollution indicators of surface-water quality.
- To determine relationships among nutrient concentrations, physicochemical characteristics, and major indicators of organic and microbiological pollution.
- To identify dominant multivariate water-quality gradients and the variables contributing most strongly to observed pollution patterns.

2. Methodology

2.1 Research Design

Surface-water quality was evaluated using a quantitative observational design and in spatial and seasonal groupings. The analysis was based on repeated environmental measurements collected from three monitoring sites (Anuppur, Dindori and Jabalpur). The records were of three seasonal classes: Summer, Monsoon and Winter and included physicochemical, nutrient, organic-pollution, microbiological, and supporting hydrological variables. The comparative design was used since the primary goal was to describe differences between monitoring contexts and determine the most important variation structure. The observed records were analysed using all statistical procedures without synthetic augmentation, interpolation or imputation to maintain the original environmental signal and not artificially inflate the number of samples. The analytical approach was chosen in consideration for the small sample size and the variables measured. The study opting for robust descriptive summaries and non-parametric inference, rank based association analysis, and exploratory dimension reduction instead of fitting high parameter regression or machine learning models. These questions addressed four interrelated issues, the overall central water-quality profile; the spatial differences in indicator distributions; the seasonal differences in indicator distributions; and the multivariate gradients of water-quality variables correlated with each other (Dilip, 2024).

2.2 Water Quality Parameters

Twelve variables were identified as being primary indicators as they encompassed the main dimensions mentioned in the study title. The following core physicochemical parameters were studied: temperature, turbidity, pH and specific conductivity. Nutrient-related conditions included ammonia nitrogen, nitrite nitrogen, nitrate nitrogen and phosphate. Dissolved oxygen, biochemical oxygen demand (BOD), chemical oxygen demand (COD) were the measurements used to describe oxygen status and organic pollution, whereas total coliform was maintained as microbiological indicator of contamination pressure. This targeted effort minimized redundancy in the inferential analysis, yet maintained the broad range of environmental domains captured in the monitoring data. Other variables including total solids, dissolved solids, suspended solids, chloride, sulphate, alkalinity, hardness fractions, sodium, potassium, rainfall and water level were also included in the source data for contextual interpretation but were not included within the primary multivariate model. This limit was imposed due to the limited number of observations and variables. It would not have been wise to include all the available measures in PCA because this would have resulted in a less stable covariance structure and less interpretable loadings. The twelve variables selected therefore represented a compromise between environmental coverage and statistical simplicity, and in their direct connection to the physicochemical and nutrient markers of pollution dynamics, supported the focus of the article.

2.3 Data Preparation and Quality Screening

Data was entered into python and clean data was used for statistical test. The column names and types were checked, and categorical types were cleaned up by trimming spaces from the leading and trailing ends of the values. Seasonal

labels were standardized to Summer, Monsoon and Winter, and location labels were left as written, such as capitalization issues. There were 36 observations in the data set and 33 variables. The number of missing values were calculated for each column and duplicate rows were identified. There were no missing values and no duplication of records so no deletion / imputation procedure was performed. Measures of central tendency and dispersion, such as mean, standard deviation, median, minimum and maximum were used to inspect continuous variables. Several nutrient and microbiological measurements exhibited strong right-skew, and the number of observations per location was small, so distribution-free methods were used to make inferential comparisons. The variables selected were standardized to z-scores based on the sample mean and standard deviation for multivariate analysis. This was needed to standardise the different numerical scales of the variables measured, from low nutrient concentrations to conductivity and coliform. If variables were not scaled, those with larger numerical variations would have taken over the component solution.

2.4 Spatial and Seasonal Statistical Analysis

The Kruskal-Wallis H test was applied to check the spatial variation among Anuppur, Dindori, Jabalpur and seasonal variation among Summer, Monsoon, and Winter for 12 water quality parameters. It was reported that group medians would give a good estimate in skewed distributions. Spatial and seasonal tests were Bonferroni corrected for the False Discovery Rate (Benjamini-Hochberg) with adjusted $p < 0.05$ deemed statistically significant. Epsilon-squared was used to estimate effect magnitude, where higher values represent greater separation of groups. To assess the magnitude and direction of the spatial and seasonal variation, effect sizes were interpreted in conjunction with the adjusted p value and median difference. The spatial analysis and the seasonal analysis were performed separately to separate the relative contribution of the location and season. Because of the limited number of individuals in the groups and the interest in the overall water-quality heterogeneity in the study, post-hoc pairwise comparisons were not given a priority.

2.5 Correlation and Principal Component Analysis

Due to the non-normal distribution of several of the water-quality indicators, relationships among the twelve indicators were evaluated using Spearman rank-order correlation. Correlations were viewed both qualitatively and quantitatively, highlighting nutrient, pollution, oxygen-demand, conductivity and microbiological relations. Major patterns of association were illustrated using a correlation heatmap. The variables were standardized and analyzed by principal component analysis (PCA) to determine the major water-quality gradients. First three components were retained as they accounted for over 70% of the total variance. Influential indicators were identified using component loadings and spatial patterns using observation scores. PCA biplot was used to show the relationship between samples and variables by plotting scores and loading vectors. Data analysis was performed using python programming language and the following libraries: pandas, NumPy, SciPy, statsmodels, scikit-learn, and Matplotlib. To ensure analytical consistency, the same predefined variables were used throughout the workflow. Figures were created directly from statistical output at publication quality. Extreme observations were kept in order to keep the possibility of genuine environmental observations, and rank-based tests and standardized PCA reduced their effect. A reproducible workflow was created to ensure that the reported results could be easily tracked back to the statistical calculations.

2.6 Ethical Considerations

The study used publicly available environmental data and involved no human participants, personal information, or animal experimentation; therefore, ethical approval was not required. Data were analyzed objectively without manipulation, and all valid observations were retained to ensure transparency and reproducibility.

3. Results

3.1 Overall Water-Quality Characteristics

The 12 major water-quality indicators are summarized in Table 1. There was significant variation in the variables with respect to dispersion and scale. The average temperature was 20.686, the standard deviation was 9.756, and the range was 2.0 to 29.5, which shows a high standard of deviation among the observations and indicates that there was a considerable variation in temperature. Turbidity had a mean of 4.253 and a maximum of 16.3, while pH was fairly tightly clustered, with a mean of 7.637 and median of 7.680. Specific conductivity was more variable, and had a mean of 281.711, a standard deviation of 111.212, and a maximum of 829.0. Ammonia nitrogen, nitrate nitrogen, nitrite nitrogen and phosphate had high positive skews, with the mean being 0.155, 0.730, 0.087, and 0.077 respectively and the median being 0.013, 0.011, 0.010, and 0.004 respectively. These mean-median comparisons revealed that a small proportion of samples had significantly higher concentrations of nutrients. There was also a more mild, though still noticeable, right-skew for nitrate nitrogen with a mean of 0.419 and a median of 0.300. Dissolved oxygen was relatively stable (av. 7.456) while BOD and COD were more variable. The total coliform showed significant variation ranging from 1.8 to 63.0 with a mean of 29.033 and a median of 37.0. Overall, the descriptive results warranted the

use of robust rank-based comparisons, and indicated that nutrient and microbiological indicators were strong contributors to heterogeneity in the data set.

Table 1. Descriptive statistics of physicochemical, nutrient, and pollution indicators

Indicator	Mean	SD	Median	Minimum	Maximum
Temperature	20.686	9.756	25.1	2.0	29.5
Turbidity	4.253	2.874	3.9	1.0	16.3
pH	7.637	0.326	7.68	6.67	8.51
Sp. Conductivity	281.711	111.212	247.65	196.3	829.0
Ammonia Nitrogen	0.155	0.253	0.013	0.006	0.9
Nitrite Nitrogen	0.087	0.145	0.01	0.0	0.49
Nitrate Nitrogen	0.419	0.35	0.3	0.14	1.5
Phosphate	0.077	0.164	0.004	0.0	0.6
Dissolved Oxygen	7.456	0.609	7.6	5.7	8.5

BOD	1.214	0.624	1.4	0.1	2.6
COD	8.753	5.567	9.6	2.0	24.0
Total Coliform	29.033	20.135	37.0	1.8	63.0

3.2 Spatial Variation in Water Quality

After FDR correction, spatial analysis revealed that all twelve indicators varied significantly between locations (Table 2). Temperature exhibited a strong location effect ($H = 23.460$, $\epsilon^2 = 0.650$), while turbidity ($H = 14.016$, adjusted $p = 0.0012$, $\epsilon^2 = 0.364$) and pH also differed significantly. The smallest significant spatial effect was for conductivity ($\epsilon^2 = 0.181$). The highest differences were for the nutrient related indicators ammonia ($\epsilon^2 = 0.708$), nitrite ($\epsilon^2 = 0.691$), phosphate ($\epsilon^2 = 0.732$) and dissolved oxygen ($\epsilon^2 = 0.774$). The median COD level at Anuppur was 2.55, and at Dindori and Jabalpur, it was 8.20 and 14.50, respectively, whereas the total coliform level was 3.7 at Anuppur, 41.0 at Dindori and 44.5 at Jabalpur. Figure 1 shows that there was a definite spatial distinction in both nutrients, organic-pollution, and microbiological indicators, with very strong water-quality variability existing from place to place.

Table 2. Spatial differences in major water-quality indicators

Indicator	Anuppur median	Dindori median	Jabalpur median	H	Adj. p	ϵ^2	Sig.
Temperature	28.1	28.1	7.0	23.4598	0.0	0.6503	Yes
Turbidity	4.05	4.55	2.0	14.0162	0.0012	0.3641	Yes
pH	7.61	7.795	7.59	13.7191	0.0013	0.3551	Yes
Sp. Conductivity	226.35	261.4	283.5	7.9876	0.0184	0.1814	Yes
Ammonia Nitrogen	0.01	0.0125	0.49	25.3514	0.0	0.7076	Yes
Nitrite Nitrogen	0.0045	0.01	0.145	24.8169	0.0	0.6914	Yes
Nitrate Nitrogen	0.21	0.22	0.745	18.9468	0.0002	0.5135	Yes
Phosphate	0.0	0.004	0.15	26.156	0.0	0.732	Yes
Dissolved Oxygen	6.85	7.95	7.6	27.526	0.0	0.7735	Yes
BOD	0.3	1.5	1.45	12.3207	0.0023	0.3127	Yes
COD	2.55	8.2	14.5	17.4204	0.0003	0.4673	Yes
Total Coliform	3.7	41.0	44.5	16.5216	0.0004	0.44	Yes

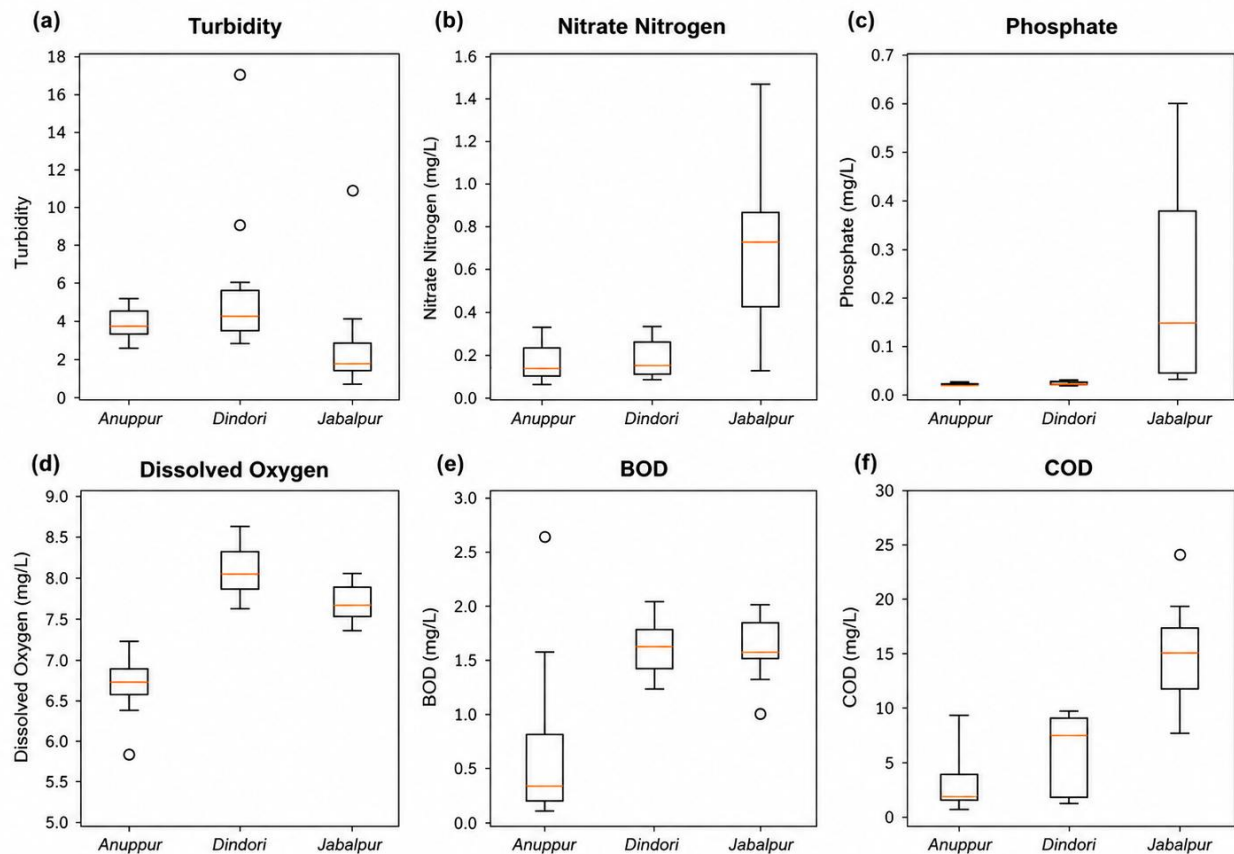


Figure 1. Spatial distribution of selected water-quality indicators across monitoring locations. (a) Turbidity, (b) Nitrate Nitrogen, (c) Phosphate, (d) Dissolved Oxygen, (e) BOD, and (f) COD. Boxes represent interquartile ranges, central lines indicate medians, whiskers represent dispersion, and circles denote outliers

3.3 Seasonal Dynamics of Water Quality

The seasonal variation was much smaller than the spatial variation. Table 3 indicates that only specific conductivity was found to be significant at the FDR corrected level ($H = 11.533$, FDR corrected $p = 0.0376$, $\epsilon^2 = 0.289$). The conductivity increased from 243.0 in Summer to 291.0 in Monsoon and decreased to 225.0 in Winter. Turbidity was borderline significant ($H = 5.896$, $p = 0.0525$) but after adjustment was nonsignificant. There were no significant differences in nutrient and pollution indicators by season. Median nitrate decreased from 0.380 in Summer to 0.180 in Winter, while COD increased from 5.3 to 10.0, but neither change was significant after correction. There was limited seasonal variation in the levels of dissolved oxygen and BOD, while total coliform was highest in Monsoon, but not significant. There is a high degree of overlap between seasonal distributions as shown in Figure 2. Overall results suggested that the spatial component accounted for a much higher proportion of the water-quality variability than did the seasonal component and that conductivity was the only major water-quality indicator that differed significantly between both spatial and seasonal factors.

Table 3. Seasonal differences in major water-quality indicators

Indicator	Summer median	Monsoon median	Winter median	H	Adj. p	ϵ^2	Sig.
Temperature	26.8	28.0	22.8	2.415	0.6128	0.0126	No

Turbidity	4.0	4.1	3.2	5.8955	0.3148	0.118	No
pH	7.64	7.68	7.8	0.8083	0.7291	0.0	No
Sp. Conductivity	243.0	291.0	225.0	11.5326	0.0376	0.2889	Yes
Ammonia Nitrogen	0.0125	0.018	0.012	0.8058	0.7291	0.0	No
Nitrite Nitrogen	0.012	0.01	0.01	2.0774	0.6128	0.0023	No
Nitrate Nitrogen	0.38	0.26	0.18	3.1808	0.6128	0.0358	No
Phosphate	0.005	0.006	0.002	1.6935	0.6128	0.0	No
Dissolved Oxygen	7.55	7.8	7.6	1.7876	0.6128	0.0	No
BOD	1.35	1.3	1.4	0.2455	0.8845	0.0	No
COD	5.3	9.2	10.0	2.3152	0.6128	0.0096	No
Total Coliform	27.5	43.0	33.0	1.5547	0.6128	0.0	No

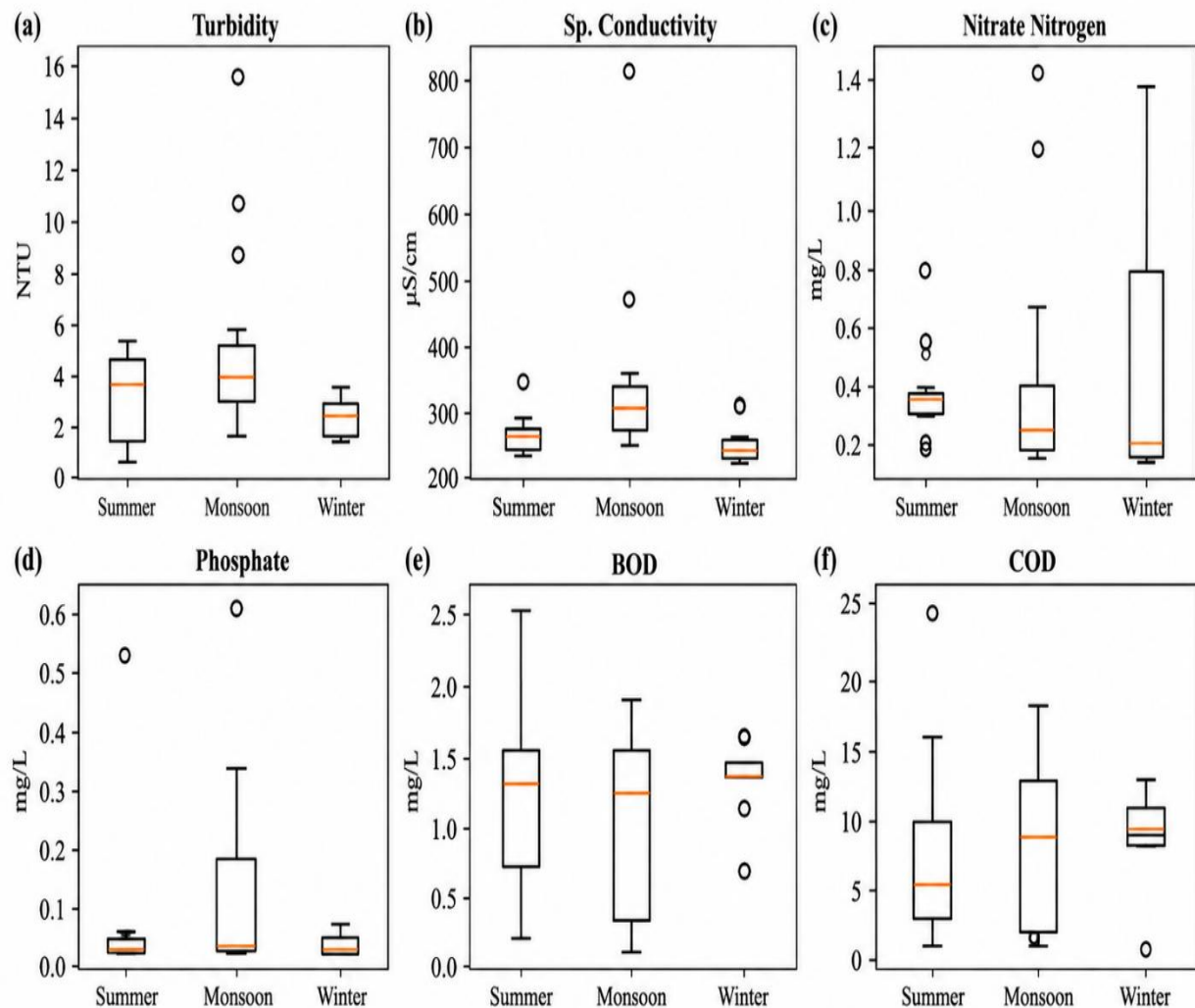


Figure 2. Seasonal variation in selected physicochemical, nutrient, and pollution indicators. (a) Turbidity, (b) Specific Conductivity, (c) Nitrate Nitrogen, (d) Phosphate, (e) BOD, and (f) COD. Boxes represent interquartile ranges, central lines indicate medians, whiskers represent dispersion, and circles denote outliers

3.4 Relationships Among Water-Quality Indicators

The nutrient indicators in Figure 3 have high positive correlations. There were positive correlations between ammonia nitrogen and nitrite nitrogen ($\rho = 0.837$) and between ammonia nitrogen and phosphate ($\rho = 0.799$) and between nitrite nitrogen and phosphate ($\rho = 0.827$) and between nitrite nitrogen and nitrate nitrogen ($\rho = 0.685$). These relationships suggested nutrient enrichment in a coordinated fashion across the sampling locations. There were also significant relationships for pollution indicators. Specific conductivity was correlated with the total coliform ($\rho = 0.779$), and BOD was positively correlated with the total coliform ($\rho = 0.703$). There was a positive relationship between COD and phosphate ($\rho = 0.634$) and between COD and nitrite nitrogen ($\rho = 0.623$); reflecting a relationship between nutrient enrichment and organic pollution. The relationship between temperature and COD was negative ($\rho = -0.730$) while it was positive between temperature and turbidity ($\rho = 0.721$). Figure 3 overall shows interconnected nutrient, organic, ionic and microbiological patterns, which will be used in the subsequent multivariate analysis.

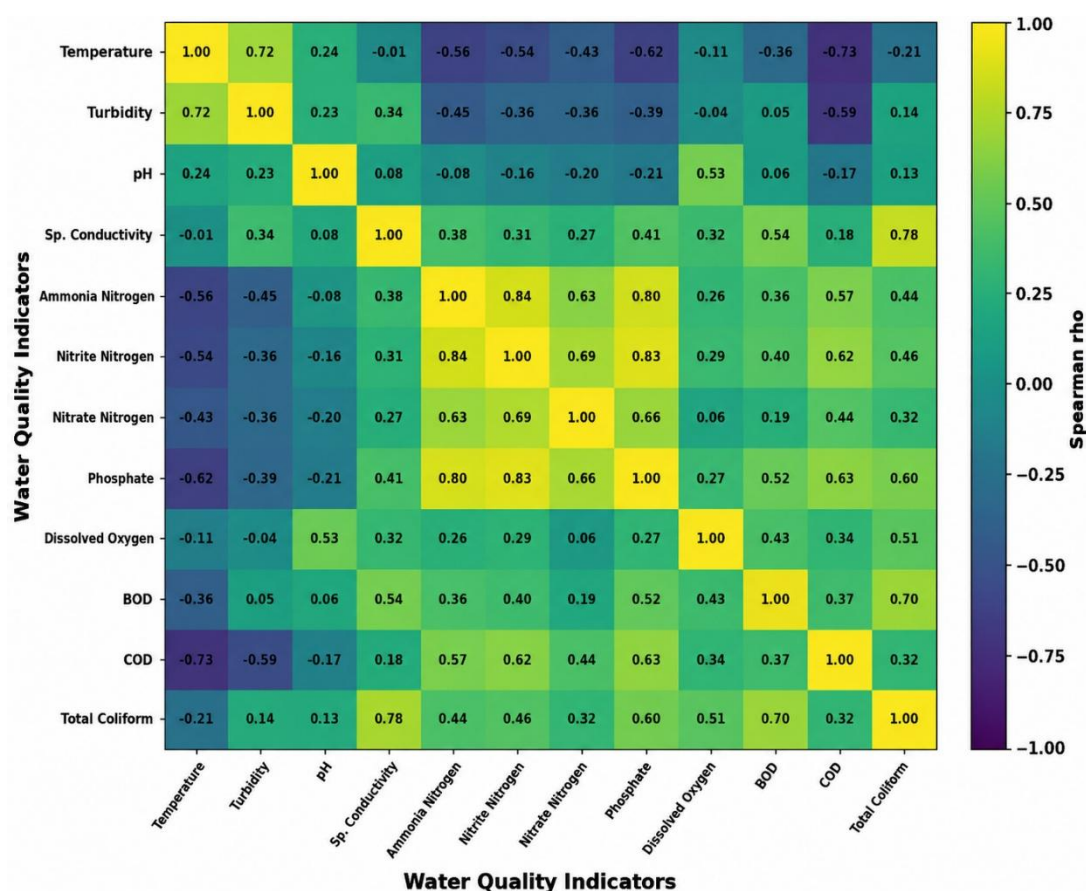


Figure 3. Spearman correlation matrix among physicochemical, nutrient, organic-pollution, and microbiological indicators

3.5 Multivariate Water-Quality Patterns

By PCA, the twelve standardized indicators were reduced to three principal components which accounted for 73.03% of total variance. PC1 accounted for 38.74%, PC2 for 22.79%, and PC3 for 11.50%, while the first two components together explained 61.53%. As seen in Table 4, PC1 was primarily correlated with temperature (0.419) and opposed to COD (-0.405), nitrite nitrogen (-0.391), phosphate (-0.364), ammonia nitrogen (-0.332) and nitrate nitrogen (-0.301) which represented a nutrient and organic-pollution gradient. Total coliform (0.468), turbidity (0.432), specific conductivity (0.411) and BOD (0.366) formed PC2, which represents a combined ionic, particulate, organic and microbiological gradient. The major factors affecting PC3 were pH (0.549) and dissolved oxygen (0.489) and opposite loadings for turbidity and conductivity. Figure 4 depicts good separation of observations in the PC1-PC2 space, in line with the strong spatial differences indicated previously. The nutrient and COD vectors were similar, while the turbidity, conductivity, BOD, and coliform vectors were more closely aligned with PC2. A PCA of all data suggested that water quality variation was due to several interacting pollution processes and not a single dominant process.

Table 4. Principal component loadings and three-component communality

Indicator	PC1	PC2	PC3	Communality (PC1-3)
Temperature	0.419	0.099	0.071	0.191
Turbidity	0.148	0.432	-0.393	0.362
pH	0.028	0.296	0.549	0.39
Sp. Conductivity	-0.138	0.411	-0.475	0.414
Ammonia Nitrogen	-0.332	-0.134	0.03	0.129
Nitrite Nitrogen	-0.391	-0.198	-0.054	0.195

Nitrate Nitrogen	-0.301	0.113	-0.234	0.158
Phosphate	-0.364	-0.18	-0.006	0.165
Dissolved Oxygen	-0.17	0.295	0.489	0.355
BOD	-0.25	0.366	0.081	0.203
COD	-0.405	-0.064	0.036	0.169
Total Coliform	-0.202	0.468	0.09	0.268

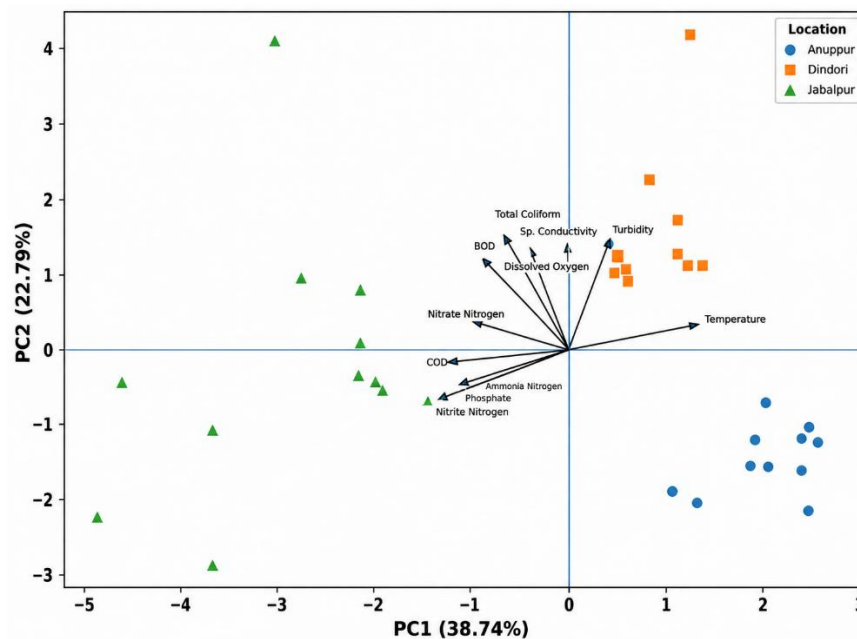


Figure 4. PCA biplot showing multivariate relationships among water-quality indicators and sampling observations

4. Discussion

Results indicated that spatial heterogeneity dominated the surface-water system being monitored. After false-discovery-rate correction, all 12 primary indicators were significantly different across locations, but differences among seasons were largely lacking. In contrast, longitudinal changes in physicochemical variables related to local inputs and hydrogeochemical conditions have been observed in river studies, with significant variations observed between different sections (Ahmad et al., 2024). Using the watershed-scale assessment, it was also found that different water-quality profiles may exist in time and space depending on the local environmental setting, as occurred in the Brazilian Cerrado (Castro et al., 2023). Site-level differences in the present analysis included basic physicochemical conditions, as well as pollution-sensitive variables, the strongest effects being seen for dissolved oxygen, phosphate, ammonia, nitrite, temperature, and nitrate.

The nutrient pattern was particularly coherent. Ammonia, nitrite, nitrate and phosphate were highly intercorrelated and had high location effects, indicating that enrichment was not a single chemical event. Runoff, wastewater influence, organic decomposition and sediment-water exchange have all the potential to influence nutrient behavior, and spatiotemporal studies of the Nokoue Lagoon have shown significant nutrient variability when the environmental conditions change locally (Socohou et al., 2024). Multivariate and water-quality-index assessment in Lake Ziway further revealed that the spatial hydrochemical structure contains information that can be lost if assessing indicators independently (Mechal et al., 2024). The much greater medians for a number of nutrient species in Jabalpur thus allow interpretation of a common enrichment gradient, although there is no specific nutrient source identified in the available data.

This spatial interpretation was confirmed by the organic-pollution and microbiological variables. BOD, COD and total coliform showed significant differences among locations and BOD had a strong relationship with coliform, and conductivity and coliform had a significant positive relationship. Similarly, in human-impacted river systems, the

hydrochemical and multivariate analysis identified linked pollution gradients, which indicated combined catchment and human stressed river system (Kumar et al., 2024). Common location structure may lead to higher correlations among variables; these relationships should not be interpreted as causal. They also demonstrate that the various components of nutrient enrichment and ionic conditions, oxygen demand, and microbiological contamination tended to co-varied and therefore should not be evaluated individually.

Seasonal influences were relatively weak, and only specific conductivity was significant following multiple-testing correction. The increased monsoon conductivity may relate to the transport of dissolved ions and catchment materials, but also could be influenced by site composition and local hydrology. Integrated lake analyses has demonstrated that indicators and environmental systems vary in their relative importance with respect to location and season (Subramaniam et al., 2023). The lack of seasonal significance for nutrients, BOD, COD, and coliform should thus be considered as a limited amount of evidence in the context of the current sample size, and not as evidence of the absence of seasonal processes. To capture shorter duration hydrological events, and site-site*season interactions, sampling frequency would need to be increased.

PCA gave a summarised description of observed structure in integrated manner. A compact component solution accounted for most information in the twelve indicators, with the first three components accounting for 73.03% standardized variance. Multivariate water-quality studies show that PCA can be a valuable exploratory tool for dimensionality reduction and for the identification of groups of related variables (Muniz & Oliveira-Filho, 2023). The first was temperature versus nutrient and COD inputs and the second was coliform, turbidity, conductivity, and BOD. Other multivariate applications have found groupings of biological and chemical indicators that are representative of larger environmental gradients (Fai et al., 2023). The clear spatial structure of scores in the biplot also confirmed the univariate result that the dominant organizing factor was location. The use of integrated multivariate and spatial assessment can also offer better monitoring prioritization (Azhari et al., 2022). Future efforts should therefore be expanded to include wider spatial and temporal sampling and incorporate land-use and discharge data to better relate the gradients identified to pollution sources and management actions. For management, this translates into resorting to uniform follow-up sampling without taking into account site history. The areas with the smallest nutrient and organic-pollution signatures should be sampled at a lower frequency, and broad seasonal sampling categories should be complemented by event-based sampling that is better able to capture the nutrient and organic-pollution pulses of runoff and the short-term effects of discharge.

5. Conclusion

The results of this study showed that the spatial variability had a greater effect than the seasonal variability on surface-water quality. All twelve physicochemical, nutrient, organic-pollution and microbiological indicators varied significantly among monitoring sites after false-discovery-rate correction, while specific conductivity was the only indicator that varied significantly between seasons. Spatial effect sizes of dissolved oxygen, phosphate, ammonia nitrogen, nitrite nitrogen, temperature and nitrate nitrogen were large, indicating sensitivity to site-level environmental differences. The high correlation between the nitrogen species suggested coordinated behavior of the nutrient species, whereas the positive correlation between BOD and conductivity and total coliform suggested a correlation between the organic loading and ionic conditions and total coliform microbiological contamination. Multivariate analysis further explained these patterns with the first three principal components accounting for 73.03% of total standardized variance. The main gradients of the first two were nutrient/COD and conductivity-turbidity-BOD-coliform. All these results taken together illustrate the merits of using non-parametric comparisons, effect-size estimation, correlation analysis and PCA in the assessment of complex surface-water quality patterns. The results, although limited by the relatively small number of observations (36), are a valuable platform for prioritization of environmental monitoring. Future investigations should expand spatial sampling and sampling frequency and include the land use, discharge characteristics and catchment pressures. Nutrients, conductivity, dissolved oxygen, BOD, COD, and coliform should be particularly emphasized as parameters for routine monitoring because they can help identify pollution sensitive areas and facilitate focused source investigations and effective surface-water management.

References

1. Abdul Wahid, A., & Arunbabu, E. (2025). Multivariate analysis of water quality dynamics in a highly eutrophic reservoir: Hydrological, meteorological, and environmental contributions. *Stochastic Environmental Research and Risk Assessment*, 39(6), 2373-2393.
2. Adestina, O. B., Paul, E. D., Nuhu, A. A., Onoyima, C. C., & Okibe, F. G. (2024). Dynamics of physicochemical parameters as indicator of water quality: A study of Ogun River, Nigeria. *Indonesian Journal of Chemical Analysis (IJCA)*, 7(2), 109-122.
3. Ahmad, T., Muhammad, S., Umar, M., Azhar, M. U., Ahmed, A., Ahmad, A., & Ullah, R. (2024). Spatial distribution of physicochemical parameters and drinking and irrigation water quality indices in the Jhelum River.

- Environmental Geochemistry and Health, 46(8), 263.
4. Azhari, H. E., Cherif, E. K., Sarti, O., Azzirgue, E. M., Dakak, H., Yachou, H., et al. (2022). Assessment of surface water quality using the water quality index (IWQ), multivariate statistical analysis (MSA) and geographic information system (GIS) in Oued Laou Mediterranean Watershed, Morocco. *Water*, 15(1), 130.
 5. Castro, R. M., Marcionilio, S. M. L. D. O., Morais, W. A., Alves, W. D. S., Oliveira, L. D., Teixeira, M. B., & Oliveira, D. M. D. S. (2023). Spatiotemporal analysis of water quality indicators in a watershed in Brazilian Cerrado. *Water Supply*, 23(3), 1161-1176.
 6. Debassi, B., Allaoua, N., Ghanem, N., Hafid, H., Benacherine, M., & Chenchouni, H. (2025). Assessment of water quality of groundwater, surface water, and wastewater using physicochemical parameters and microbiological indicators. *Science Progress*, 108(2), 00368504251348544.
 7. Dilip, M. (2024). Narmatha water quality [Data set]. Kaggle. <https://www.kaggle.com/datasets/ajithdari/narmatha-water-quality>
 8. Do Thi, V. H., Bui, T. T., & Nguyen, T. T. Q. (2025). Apply multivariate statistical analyses to assess surface water quality in the Quang Binh River System. *VNU Journal of Science: Earth and Environmental Sciences*, 41(2).
 9. Fai, P. B. A., Kenko, D. B. N., Tchamadeu, N. N., Mbida, M., Korejs, K., & Riegert, J. (2023). Use of multivariate analysis to identify phytoplankton bioindicators of stream water quality in the monomodal equatorial agroecological zone of Cameroon. *Environmental Monitoring and Assessment*, 195(6), 788.
 10. Giao, N. T., Dan, T. H., Ni, D. V., Anh, P. K., & Nhien, H. T. H. (2022). Spatiotemporal variations in physicochemical and biological properties of surface water using statistical analyses in Vinh Long Province, Vietnam. *Water*, 14(14), 2200.
 11. Hammoumi, D., Al-Aizari, H. S., Alaraidh, I. A., Okla, M. K., Assal, M. E., Al-Aizari, A. R., et al. (2024). Seasonal variations and assessment of surface water quality using water quality index (WQI) and principal component analysis (PCA): A case study. *Sustainability*, 16(13), 5644.
 12. Hong, J., Zhang, J., Song, Y., & Cao, X. (2022). Spatial and temporal distribution characteristics of nutrient elements and heavy metals in surface water of Tibet, China and their pollution assessment. *Water*, 14(22), 3664.
 13. Hong, T. T. K., Viet, L. H., & Giao, N. T. (2022). Analysis of spatial-temporal variations of surface water quality in the southern province of Vietnamese Mekong delta using multivariate statistical analysis. *Journal of Ecological Engineering*, 23(7).
 14. Kanownik, W., Policht-Latawiec, A., Możdżeń, M., & Dąbrowska, J. (2024). Deciphering anthropogenic impact: A multifaceted statistical analysis of physico-chemical parameters in a catchment with limited water quality data. *Desalination and Water Treatment*, 320, 100689.
 15. Kumar, V., Alam, A., Kumar, J., Thakur, V. R., Kumar, V., Srivastava, S. K., et al. (2024). Water quality assessment, possible pollution source identification from anthropogenically stressed River Yamuna, India using hydrochemical, water quality indices and multivariate statistics analysis. *Water, Air, & Soil Pollution*, 235(12), 820.
 16. Mansour, R., Halwani, J., El-Dakdouki, M. H., & Mina, S. (2024). Seasonal assessment of surface water and sediments pollution in Rachiine River, Northern Lebanon, using multivariate statistical analysis. *Heliyon*, 10(19).
 17. Mechal, A., Fekadu, D., & Abadi, B. (2024). Multivariate and water quality index approaches for spatial water quality assessment in Lake Ziway, Ethiopian Rift. *Water, Air, & Soil Pollution*, 235(1), 78.
 18. Muniz, D. H., & Oliveira-Filho, E. C. (2023). Multivariate statistical analysis for water quality assessment: A review of research published between 2001 and 2020. *Hydrology*, 10(10), 196.
 19. Saha, A., Paul, T. T., Sudheesan, D., Sharma, S. K., Suresh, V. R., Das, B. K., et al. (2022). Assessment of spatial and temporal changes in water quality of a tropical river in Southern Western Ghats, Kerala, India, using physicochemical quality indices and multivariate analysis. *Natural Resources Research*, 31(3), 1375-1401.
 20. Singh, J. K., Kumar, P., & Vishwakarma, S. (2022). Multivariate and statistical evaluation of coastal water quality and seasonal variation in the physicochemical properties of Gulf of Khambhat Region, Gujarat, India. *Water, Air, & Soil Pollution*, 233(9), 358.
 21. Socohou, A. A., Djihouessi, M. B., Tigo, B. A., Viainon, V. M., Adandedji, F. M., Adounpke, J., & Aina, M. P. (2024). Spatio-temporal dynamics of nutrients at the water-sediment interface: Case of the Nokoue lagoon in Southern Benin. *Environmental Monitoring and Assessment*, 196(1), 2.
 22. Subramaniam, P., Ahmed, A. N., Fai, C. M., Abdul Malek, M., Kumar, P., Huang, Y. F., et al. (2023). Integrated GIS and multivariate statistical approach for spatial and temporal variability analysis for lake water quality index. *Cogent Engineering*, 10(1), 2190490.
 23. Tibebe, D., Zewge, F., Lemma, B., & Kassa, Y. (2022). Assessment of spatio-temporal variations of selected water quality parameters of Lake Ziway, Ethiopia using multivariate techniques. *BMC Chemistry*, 16(1), 11.
 24. Ubuoh, E. A., Nwogu, F. U., Ossai-Abek, E. Q., Ikwuemesi, J. C., Oke, A. O., & Umoh, J. D. (2024). Evaluation of hydro-chemical facies and surface water quality dynamics using multivariate statistical approaches in Southern Nigeria. *Scientific Reports*, 14(1), 31600.
 25. Zhang, H., Li, H., Gao, D., & Yu, H. (2022a). Source identification of surface water pollution using multivariate statistics combined with physicochemical and socioeconomic parameters. *Science of the Total Environment*, 806, 151274.
 26. Zhang, Z. M., Zhang, F., Du, J. L., & Chen, D. C. (2022b). 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*, 14(5), 778.