Shahrood University of Technology

Journal of Hydraulic and Water Engineering (JHWE)

Journal homepage: https://jhwe.shahroodut.ac.ir

Iranian Hydraulic Association (IHA)

Using Multiple Correspondence Analysis to Identify Types of Pollution in Water Resources of Lukula Region

André Mampuya Nzita[1,2], Pascal Bakwita Nkoso[2,3], Jacob Gene Obi[4,5], Dan Kandala Muamba[2,4], André Kanjinga Ngoyi[2,4], Thierry Shima Ngoy[6], Guyh Dituba Ngoma[7], Clément N'zau Umba-di-Mbudi[1,4] Roger Vumilia Kizungu[8]

[1] President Joseph Kasa-Vubu University, Faculty of Engineering, Boma, Democratic Republic of the Congo
[2] Regional School of Water (ERE), University of Kinshasa (UNIKIN), Kinshasa, Democratic Republic of the Congo
[3] Higher Pedagogical Institute of Inongo, Agroveterinary Sciences, Exact Sciences stream, Inongo, Democratic Republic of the Congo
[4] University of Kinshasa, Faculty of Science and Technology, Kinshasa, Democratic Republic of the Congo
[5] Higher Institute of Agronomic Studies, Mangaly, and Ministry of Mines, Kinshasa, Democratic Republic of Congo
[6] Laboratory of Chemistry and Geochemistry, Center for Geological and Mining Research, B.P.898 Kinshasa Gombe, Democratic Republic of the Congo
[7] University of Quebec in Abitibi-Temiscamingue, School of Engineering, Rouyn-Noranda, Canada.
[8] University of Kinshasa, Faculty of Agronomic Sciences, Department of Phytotechnics, Kinshasa, Democratic Republic of the Congo

Article Info

Article history:
Received 25 August 2025
Received in revised form 11 June 2026
Accepted 25 June 2026
Published online 28 June 2026

DOI:
10.22044/jhwe.2026.16734.1078

Keywords
Multiple Correspondence Analysis
Multivariate Statistical Analysis:
Water Quality Assessment
Spring Water Pollution
Lukula Region

Abstract

Water quality degradation is a major environmental and public health issue, particularly in developing countries where communities depend on untreated groundwater and spring water for domestic use. In the Lukula region (Kongo Central Province, Democratic Republic of the Congo), limited access to safe drinking water increases the risk of exposure to physicochemical and microbiological contamination. This study aimed to characterize the quality of spring waters in the Lukula Health Zone and to identify the main types of pollution affecting these resources using multivariate statistical techniques. Water samples were collected from ten springs during two sampling campaigns conducted in April and December 2018. Physicochemical parameters (pH, temperature, electrical conductivity, total dissolved solids, turbidity, alkalinity, hardness, major ions, and nutrients) and microbiological indicators (total viable counts and total coliforms) were analyzed. Data were processed using Principal Component Analysis (PCA), Hierarchical Ascending Clustering (HAC), and Multiple Correspondence Analysis (MCA). The results revealed significant spatial variability in water quality. PCA indicated that the first two principal components explained 64.1% of the total variance, with the first component primarily associated with mineralization, alkalinity, hardness, bicarbonates, calcium, and magnesium, and the second primarily associated with chloride, sodium, potassium, and sulfate. HAC classified the springs into three hydrochemical groups reflecting differences in mineralization and salinity. MCA identified multiple pollution patterns, including mineral pollution, nutrient enrichment, microbiological contamination, sediment-related pollution, and anthropogenic impacts. Springs such as Kizinga, Kiyangi I, Kinduka, and Kilonde showed the highest contamination levels. These findings demonstrate the value of multivariate statistical methods for water-quality assessment and provide a scientific basis for monitoring, pollution control, and sustainable water-resource management in the Lukula region.

1. Introduction

Establishing a reliable and safe drinking water supply system is essential for meeting the diverse needs of patients, healthcare professionals, and local communities within the Lukula region. Due to the limited availability of potable water in this area, residents often rely on nearby water sources to meet their daily water needs. Consequently, a comprehensive assessment of the quality of these water resources is necessary to determine appropriate treatment requirements and ensure their suitability for human consumption. This issue is particularly important in the current context, where access to safe drinking water remains a fundamental public health concern (Shima Ngoy et al., 2020).

Numerous studies have investigated surface-water quality; however, several critical aspects related to sustainable water-resource management remain insufficiently addressed. Recent research has demonstrated that the assessment and characterization of water resources are essential for understanding hydrological processes and evaluating environmental impacts (Alam et al., 2021; Ajtai et al., 2025; Benkov et al., 2023; Bhatti et al., 2019; Gad et al., 2020; Giao et al., 2022; Gupta et al., 2023; Kumar et al., 2022; Mohammed & Al-Obaidi, 2021; Ustaoğlu et al., 2021). Nevertheless, significant knowledge gaps persist regarding the interactions among physicochemical parameters and their influence on environmental quality and human health. The present study seeks to address these gaps by integrating advanced multivariate statistical approaches to improve the understanding of spring-water characteristics and identify the major factors controlling water quality (Benkov et al., 2023; Xue et al., 2009; Zhang et al., 2022; Shil et al., 2019; Najah et al., 2021; Nnorom et al., 2019).

The primary objectives of this study were to characterize the water sources of the Lukula Health Zone and to identify the principal forms of pollution affecting these resources. To achieve these objectives, several multivariate statistical techniques were employed, including Principal Component Analysis (PCA), Hierarchical Ascending Clustering (HAC), and Multiple Correspondence Analysis (MCA). These methods facilitate data reduction, identify the most influential variables, and improve the interpretation of complex environmental datasets. The analyses were conducted using R software, enabling detailed evaluation of water-quality characteristics and the environmental factors influencing them (Alam et al., 2021; Benkov et al., 2023; Giao et al., 2022; Mampuya Nzita et al., 2024; Nzita et al., 2024).

This study contributes to a better understanding of water quality in the Lukula region. It provides valuable information to inform the development of effective management strategies to preserve these essential resources. The findings are expected to support evidence-based water management policies and improve access to safe drinking water for local populations.

2. Materials and Methods

2.1. Study Area

The study was conducted in the Lukula Health Zone, located in Kongo Central Province, Democratic Republic of the Congo (Figure 1). The investigated area comprises several villages whose water sources were selected for hydrochemical and microbiological assessment.

The geographical coordinates of the sampled springs are presented in Table 1.

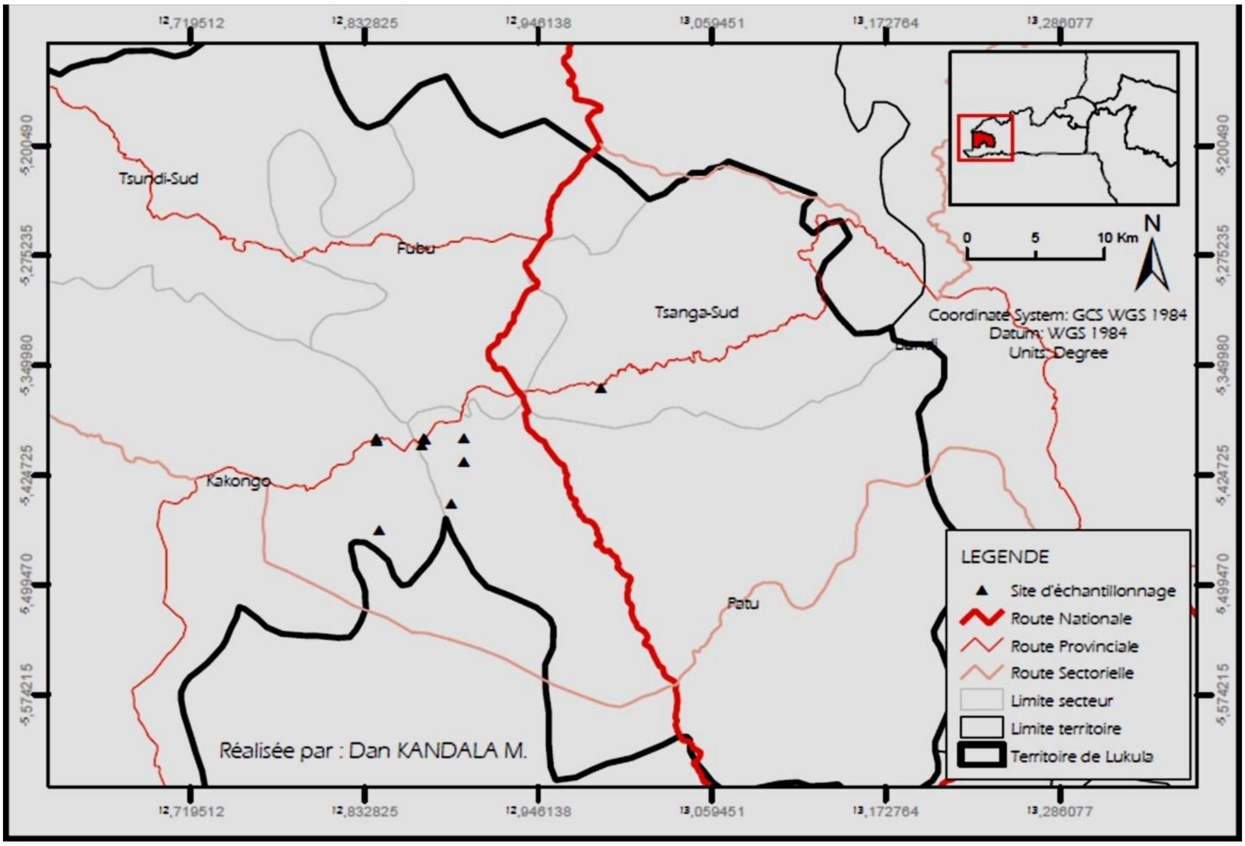

Figure 1. Location of the study site

Table 1. Geographical coordinates of village water points in the Lukula health zone

Sites	Latitude Sud	Longitude East
Kinduka	05°24’53.6’’	012°53’50.1’’
Kilonde	05°26’37.2’’	012°53’21.5’’
Kilonde I	05°27’40.5’’	012°50’31.4’’
Kiala Mongo I	05°24’02.7’’	012°50’26.9’’
Kibinonga	05°23’57.3’’	012°52’16.9’’
Kingola Fuka	05°24’14.9’’	012°52’12.9’’
Kiyangi I	05°23’59.5’’	012°52’20.4’’
Kizinga	05°23’57.9’’	012°53’50.1’’
Tuidi Yila	05°21’55.8’’	012°59’13.7’’
Kiala Mongo II	05°23’58.1’’	012°50’24.6’’

2.2. Data Collection Methods

Water sampling campaigns were conducted in April and December 2018 to account for potential seasonal variations in water-quality parameters (Shima Ngoy et al., 2020). Samples were collected from 10 spring water sources distributed across the study area.

2.3. Measurement Methods and Statistical Tests

Water samples were collected in sterile 1-L polyethylene bottles, properly labeled, and stored in insulated coolers at approximately 4°C during transportation to the laboratory of the Center for Geological and Mining Research (CRGM). Each sample was accompanied by detailed information regarding the sampling location, collection date, and sanitary conditions observed at the water source.

Three physicochemical parameters, namely pH, temperature, and electrical conductivity, were measured in situ immediately after sample collection (Ajtai et al., 2025). Laboratory analyses were subsequently performed to determine concentrations of chloride, sodium, potassium, sulfate, phosphate, nitrate, calcium, magnesium, bicarbonate, total dissolved solids, total alkalinity, total hardness, iron, suspended solids, and turbidity.

Microbiological analyses were conducted to quantify total coliforms, fecal coliforms, fecal streptococci, fungi, and yeasts using standard membrane filtration techniques. Total coliforms were determined by filtering 100 mL of water through a membrane filter, followed by incubation on MacConkey agar at 37°C. Fecal coliforms were analyzed by filtering 100 mL of water through a membrane filter and incubating the filter on lauryl sulfate broth at 37°C and 44°C for 48 hours. Fecal streptococci were enumerated by filtering 100 mL of water through a membrane filter and incubating the filter on Streptococcal KF agar. For the determination of fungi and yeasts, 250 mL of water was filtered through a membrane filter and incubated on Sabouraud agar.

In addition, total viable counts and total coliforms were determined using standardized membrane filtration procedures. Appropriate culture media and incubation conditions were applied in accordance with established analytical protocols. Microbiological results were expressed as colony-forming units per 100 mL of water (CFU/100 mL) (Ajtai et al., 2025).

The physicochemical analyses included the measurement of various parameters such as pH, chloride (Cl or Chl) in (mg/l), total iron (Fer) in (mg/l), sodium (Na) in (mg/l), potassium (K) in (mg/l), sulfate (Sul) in (mg/L), phosphate (Phos) in (mg/l), iron, suspended solids (Mo) in (mg/l), turbidity (Turb) in (NTU), total dissolved solids (Sols) in (mg/l), conductivity (Cond) in (µ/cm), calcium (Ca) in (mg/l), bicarbonate (Bic) in (mg/l), magnesium (Mg) in (mg/l), temperature (Temp) in (°C), total alkalinity (TAC) in (°F), total hardness (TH) in (°F), nitrate (Nit) in (mg/l), and for the analysis bacteriological we have the total germs (GeT) in (CFU/100ml) and total coliforms (ColiT) in (CFU/100ml) for individuals in the spring water samples respectively in tables 2 and 3. Standardized methods were applied to ensure the reliability of the results.

To ensure the reliability of the results, all analyses were performed in accordance with internationally recognized analytical procedures. Statistical analyses included descriptive statistics, Pearson correlation analysis, and Student's t-tests to evaluate relationships among the measured variables and identify factors influencing water quality (Benkov et al., 2023).

Table 2. The physicochemical parameters of spring water

Samples	Cond ($\mu s/c$)	Temp (°C)	pH	Sols (mg/l)	Turb (NTU)	TAC (°F)	Mo (mg/l)	Bic (mg/l)	Cl (mg/l)	Sul (mg/l)	Nit (mg/l)	Phos (mg/l)	Na (mg/l)	K (mg/l)	Ca (mg/l)	Mg (mg/l)	TH (°F)	Fer (mg/l)
Kinduka	178	26.1	7.67	89	0.157	8.2	3.5	100	32.2	2	6.05	0.29	18.01	11.25	23.21	5.41	7.22	0.08
Kilonde	180	26.2	7.46	90	0.142	7.8	3	95	39	2.81	8.31	0.31	22.04	5.01	23.71	7.22	8.91	0.12
Kilonde I	22	26.5	7.13	11	0.214	2.6	4.6	31	46	2.56	7.01	0.11	25.6	11.01	7.21	2.41	2.96	0.3
Kiala Mongo I	34	25.9	7.12	17	0.025	2.5	1.9	30.52	35	2.34	4.22	0.15	20	6.21	7.65	2.51	2.91	0.08
Kibinonga	24	26.9	7.14	12	0.605	2.4	4	29.21	28	3	2.25	0.21	16	3.7	7.34	2.41	2.81	0.15
Kingola Fuka	80	27.1	7.24	40	0.17	2.6	4.3	32	49	2.22	4.11	0.11	27.21	8.01	8.01	2.52	3.21	0.25
Kiyangi I	208	26.6	7.62	104	0.136	8.6	4.6	104	48	3.01	5.06	0.22	27.42	2.96	26.01	9.71	8.45	0.35
Kizinga	168	26.6	7.60	84	0.047	4.2	4.3	39	99	2.05	5.23	0.1	54.43	17.42	9.71	3.06	3.22	0.25
Tuidi Yila	42	27.1	7.50	21	0.016	2.5	2.9	30	49	2.51	3.12	0.12	27.33	8.01	7.51	2.41	2.87	0.07
Kiala Mongo II	43	27.5	7.42	21	0.036	2.4	2.5	29	35.07	2.8	4.11	0.09	19.74	6.3	7.31	2.41	2.82	0.1

Table 3. Biological parameters of different spring waters

Samples	Total germs (GeT) in CFU /100ml	Total coliforms (ColiT) in CFU /100ml
Kinduka	6.000	4.000
Kilonde	5.000	2.000
Kilonde I	4.000	0.000
Kiala Mongo I	6.000	3.000
Kibinonga	8.000	5.000
Kingola Fuka	7.000	5.000
Kiyangi I	8.000	6.000
Kizinga	9.000	5.000
Tuidi Yila	3.000	0.000
Kiala Mongo II	4.000	2.000

Values reported as 0.000 indicate that no microorganisms were detected in the water samples analyzed.

2.4. Multivariate Analysis Techniques

We used R software (Mampuya Nzita et al., 2025; Nzita et al., 2024) to conduct a multivariate analysis of parameters used to identify pollution types across different water sources, at a significance level ($\alpha = 0.05$). This software met the objectives of characterizing and identifying the types of pollution present in these water sources. To analyze the complex relationships between variables, three main statistical methods were used: Principal Component Analysis (PCA), Hierarchical Ascending Clustering (HAC), and Multiple Correspondence Factor Analysis (MCFA). PCA reduced the dimensionality of the data while identifying the most significant variables, thus facilitating the interpretation of the results. HAC was used to group water samples by similarity, while AFCM was used to explore relationships among water-source characteristics. These methods provide a comprehensive view of water quality and the environmental factors that influence it, thus contributing to informed water resource management decisions (Gad et al., 2020; Giao, 2022; Nath et al., 2024; Nzita et al., 2024; Tripathi & Singal, 2019).

3. Results and Discussion

3.1. Principal Component Analysis (PCA)

To analyze the results in Table 4 on variable loading coefficients from the Principal Component Analysis (PCA) of source waters, we will examine the variables' contributions to the first two principal components (PC1 and PC2) and the proportion of variance explained by these components.

Table 4. Absolute values of variable loading coefficients in PCA and the amount of information explained by the main factors of source waters

Variable	PC1	PC2
Sul (mg/l)	0.013	-0.276
Mo (mg/l)	-0.015	-0.194
Cl (mg/l)	-0.023	-0.459
K (mg/l)	-0.026	0.395
Na (mg/l)	-0.027	-0.458
Turb	0.033	0.169
Fer (mg/l)	0.076	-0.227
GeT (CFU /100ml)	-0.116	-0.214
ColiT (CFU /100ml)	-0.144	-0.133
Temp (°C)	-0.169	0.059
Nit (mg/l)	-0.201	-0.017
pH	-0.239	-0.163
Phos (mg/l)	0.264	-0.246
Mg (mg/l)	0.319	-0.075
Cond (mg/l)	-0.325	-0.137
Sols (mg/l)	-0.325	-0.137
TH (°F)	0.329	-0.118
Bic (mg/l)	0.334	-0.088
Ca (mg/l)	0.335	-0.089
TAC (°F)	0.339	-0.039
Standard Deviation	2.919	2.071
Proportion of Variance	0.426	0.214
Cumulative Proportion	0.426	0.641

The principal component analysis (PCA) revealed that the first two principal components accounted for a substantial proportion of the total variability observed in the spring water dataset. PC1 explained 42.6% of the total variance, whereas PC2 accounted for an additional 21.4%, resulting in a cumulative explained variance of 64.1%. These results indicate that the first two components capture most of the information contained in the original variables and therefore provide a reliable representation of the hydrochemical characteristics of the investigated spring waters.

The loading coefficients indicate the relative contribution of each variable to the formation of the principal components. For PC1, the highest loadings were observed for total alkalinity (TAC; 0.339), calcium (Ca; 0.335), bicarbonate (HCO_3^-; 0.334), total hardness (TH; 0.329), electrical conductivity (Cond.; 0.325), total dissolved solids (TDS; 0.325), and magnesium (Mg; 0.319). These variables are closely associated with groundwater mineralization processes and suggest that PC1 primarily reflects the mineral and carbonate chemistry of the spring waters. Variables such as pH (0.239) and nitrate (0.201) also contributed to PC1, although their influence was less pronounced.

For PC2, the largest loading coefficients were associated with chloride (Cl; 0.459), sodium (Na; 0.458), potassium (K; 0.395), sulfate (SO_4^{2-}; 0.276), and phosphate (0.246). This pattern suggests that PC2 is primarily associated with salinity and dissolved-ion enrichment, which may reflect lithological influences or anthropogenic inputs affecting water quality. In contrast, temperature (0.059), total coliforms (0.133), and total viable counts (0.214) contributed relatively weakly to this component.

The standard deviation values reported for PC1 (2.919) and PC2 (2.071) correspond to the square roots of the respective eigenvalues associated with these principal components. Their presentation is important because they provide a direct measure of the data's dispersion along each principal axis and facilitate interpretation of the components' relative importance. Squaring these standard deviations yields the eigenvalues, which quantify the amount of variance explained by each component. Consequently, the higher standard deviation associated with PC1 reflects its greater capacity to summarize the variability contained in the original dataset compared with PC2.

Overall, the PCA results indicate that the hydrochemical variability of the spring waters is primarily controlled by mineralization processes involving alkalinity, hardness, bicarbonate, calcium, and magnesium. At the same time, a secondary gradient is associated with salinity-related parameters such as chloride, sodium, and potassium. These findings provide valuable insight into the factors governing spring water quality and contribute to a better understanding of the hydrogeochemical processes operating within the study area.

Figure 2. PCA on 10 individuals and 20 variables from source waters.

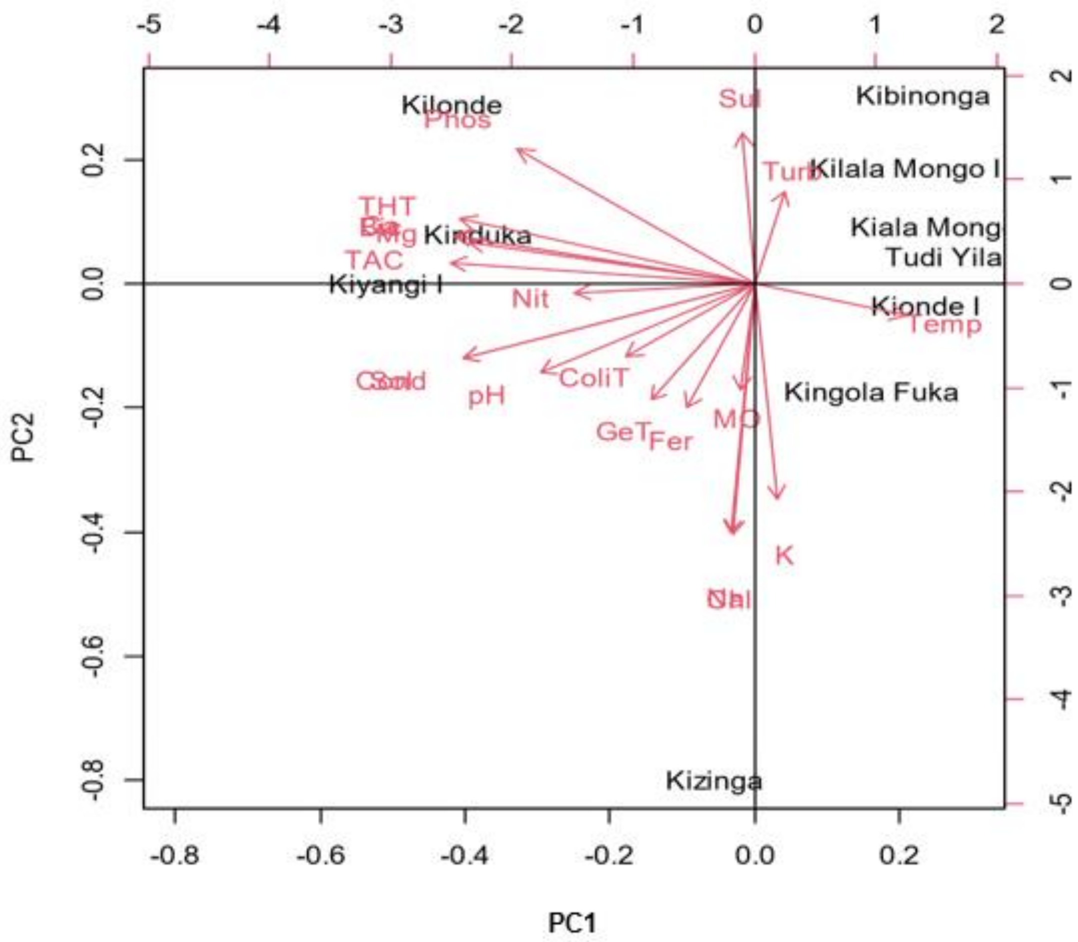

Figure 2. PCA on 10 individuals and 20 variables from the source waters

The projection of sampling sites onto the PC1–PC2 factorial plane revealed distinct hydrochemical characteristics among the investigated spring waters. Kizinga is positioned on the negative side of PC1 and is associated with elevated concentrations of sodium, chloride, and other dissolved constituents. This pattern suggests the influence of mineral enrichment and possible anthropogenic inputs, potentially related to agricultural activities or other sources of environmental contamination. Kiala Mongo I is located on the positive side of PC1. Although this site is associated with variables that contribute to this component, the turbidity value reported in Table 2 is very low (0.025 NTU), indicating good water clarity. The position of Kiala Mongo I on the factorial plane is more likely related to the combined influence of other physicochemical parameters associated with PC1 rather than suspended particulate matter.

Kibinonga is characterized by its association with variables contributing to the positive side of PC1, reflecting hydrochemical conditions similar to those observed at neighbouring springs. Kinduka is situated on the negative side of PC2 and appears to be influenced by bicarbonate concentrations and alkalinity, which may reflect natural water–rock interactions or, to a lesser extent, external environmental inputs. Kiyangi I is associated with nitrate and coliform contamination, suggesting the influence of agricultural runoff and microbiological pollution. The concomitant increase in conductivity and iron concentrations further indicates possible anthropogenic impacts on water quality. Kilonde is associated with phosphorus and sulfate, which may originate from agricultural practices, domestic wastewater, or natural geochemical processes. Elevated phosphorus concentrations may contribute to eutrophication in receiving aquatic environments. Kilonde I, located on the positive side of PC2, is associated with higher temperature and potassium levels, reflecting specific hydrogeochemical conditions that distinguish it from the other sampling sites.

Tuidi Yila and Kiala Mongo II are grouped with variables related to suspended matter and turbidity, suggesting a greater influence of sediment inputs than that observed at Kiala Mongo I. Such conditions may result from soil erosion, surface runoff, or other environmental disturbances affecting water clarity.

Overall, the PCA highlights significant spatial variability among the spring waters, reflecting differences in hydrochemical composition and potential sources of contamination. These findings emphasize the importance of continuous monitoring and appropriate management strategies to preserve water quality and protect public health.

3.1.1. Hierarchical Ascending Classification (HAC)

Hierarchical ascending classification (HAC) was used to group waters with similar salinity and hardness characteristics, as shown in Figure 3.

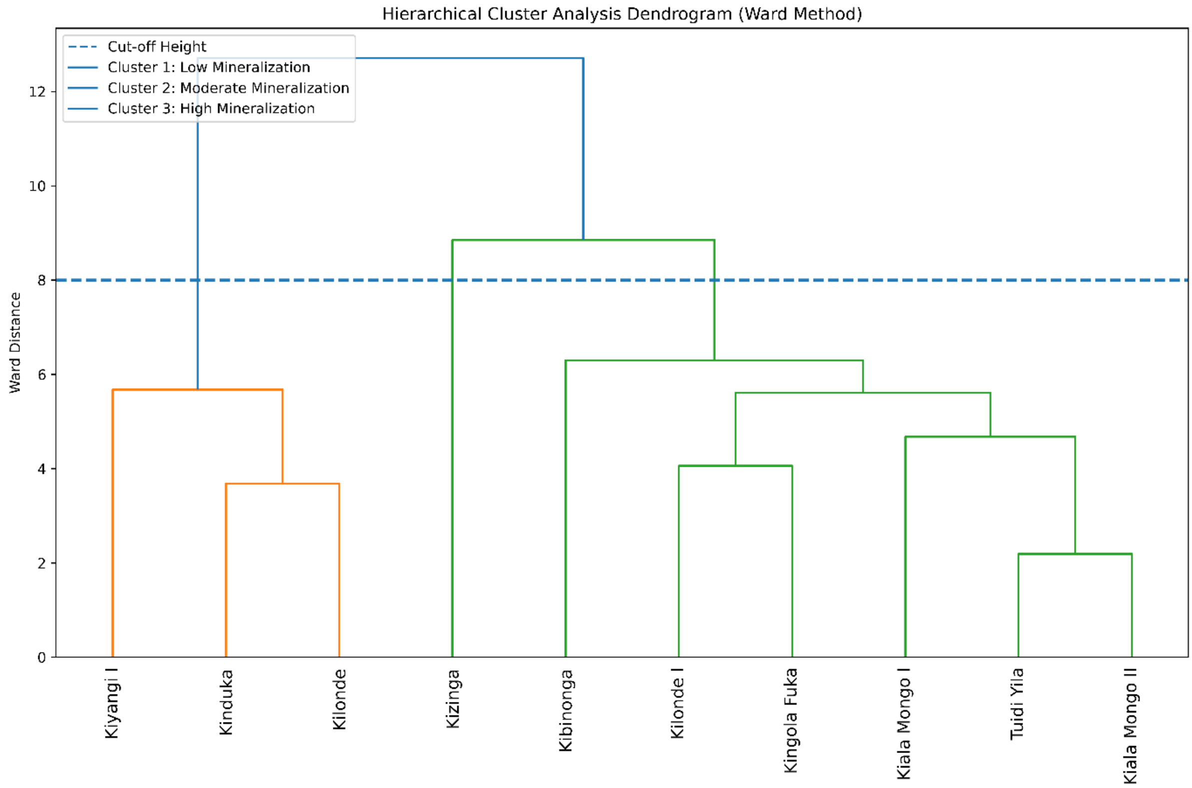

Figure 3. Dendrogram classification for different water sources

Hierarchical analysis revealed three main types of spring water. The first group comprises the most mineralized waters (Kinduka, Kilonde, and Kiyangi I), the second group includes waters with low to medium mineralization (six springs: Kilonde I, Kiala Mongo I, Kibinonga, Kingola Fuka, Tuidi Yila, and Kiala Mongo II), while the Kizinga spring constitutes a third group due to its atypical chemical composition. This classification highlights the heterogeneity of spring water quality and suggests distinct natural mechanisms that govern their physicochemical composition.

3.1.2. Multiple Correspondence Factor Analysis (MCFA)

Multiple correspondence factor analysis (MCFA) was performed, revealing relationships among the classes in Table 5. To analyze the types of pollution encountered at the water source sampling sites of Kinduka, Kilonde, Kilonde I, Kilala Mongo I, Kibinonga, and Kingola Fuka, it is essential to consider the results of the pH, Temp, Sols, and Turb classes. These classes allow an assessment of the conditions of each site, based on precise numerical values:pH-- (low pH), pH- (slightly low), pH+ (slightly high), pH++ (high), Temp-- (low temperature), Temp- (slightly low), Temp+ (slightly high temperature), Temp++ (high temperature), Sol-- (low dissolved solids), Sol- (slightly low dissolved solids), Sol+ (slightly high dissolved solids), Sol++ (high dissolved solids), Turb-- (low turbidity), Turb- (slightly low turbidity), Turb+ (slightly high turbidity), and Turb++ (high turbidity).

Table 5. Physicochemical characteristics of spring water samples

Samples of spring waters	pHcla	TempCla	SolCla	TurCla
Kinduka	pH+ + 3	Temp- - 3	Sol+ + 3	Turb+ 2
Kilonde	pH+ 2	Temp- - 3	Sol+ + 3	Turb+ 2
Kilonde I	pH- - 3	Temp- 3	Sol- - 3	Turb+ + 3
Kiala Mongo I	pH- - 3	Temp- - 3	Sol- - 3	Turb- - 3
Kibinonga	pH- - 3	Temp+ 1	Sol- - 3	Turb+ + 3
Kingola Fuka	pH- 2	Temp+ + 3	Sol+ 2	Turb+ + 3

Kinduka indicates that a high pH could signal increased alkalinity, potentially due to agricultural drainage water inputs or industrial discharges. The low temperature and slightly elevated turbidity may also suggest sediment inputs, often associated with soil erosion. Slightly increased dissolved solids may affect the water's potability. Kilonde's results are similar to Kinduka's, suggesting potentially similar pollution. This may be due to agricultural practices in the area, where fertilizer use can increase pH and dissolved solids. Kilonde I indicates that a low pH may indicate acidification, perhaps due to acid fallout or acidic wastewater discharges. High turbidity suggests a significant presence of sediment or organic matter, which can harm aquatic life by reducing water clarity. Kilala Mongo I exhibits the least favorable water quality characteristics, with low pH and low turbidity. This could indicate good water clarity, but low dissolved solids could harm aquatic biodiversity. The source of pollution could be less anthropogenic and more natural. Kibinonga notes that increased temperature, combined with low pH, may indicate thermal pollution, possibly due to hot-water discharges from nearby industries. High turbidity may result from degradation of surrounding soils, affecting water quality. As with Kibinonga, low pH and high temperature may signal thermal or acidic pollution. Slightly elevated dissolved solids and turbidity suggest sediment inputs that may be linked to nearby agricultural or urban activities. The results show significant variation between sites, reflecting the impacts of pollution. Sites with low pH and high turbidity pose risks to aquatic life, including affecting photosynthesis and respiration in aquatic organisms. Furthermore, high levels of dissolved solids can degrade water quality, making it unsafe for consumption. Temperature fluctuations can influence the distribution of aquatic species, favoring those that tolerate warmer conditions. It is essential to continue monitoring these sites to detect trends in water quality and assess the impact of human activities and natural phenomena. Appropriate management measures must be implemented to reduce pollution and protect these precious water resources.

Figure 4 assesses the condition of each source water based on the signs associated with physicochemical parameters.

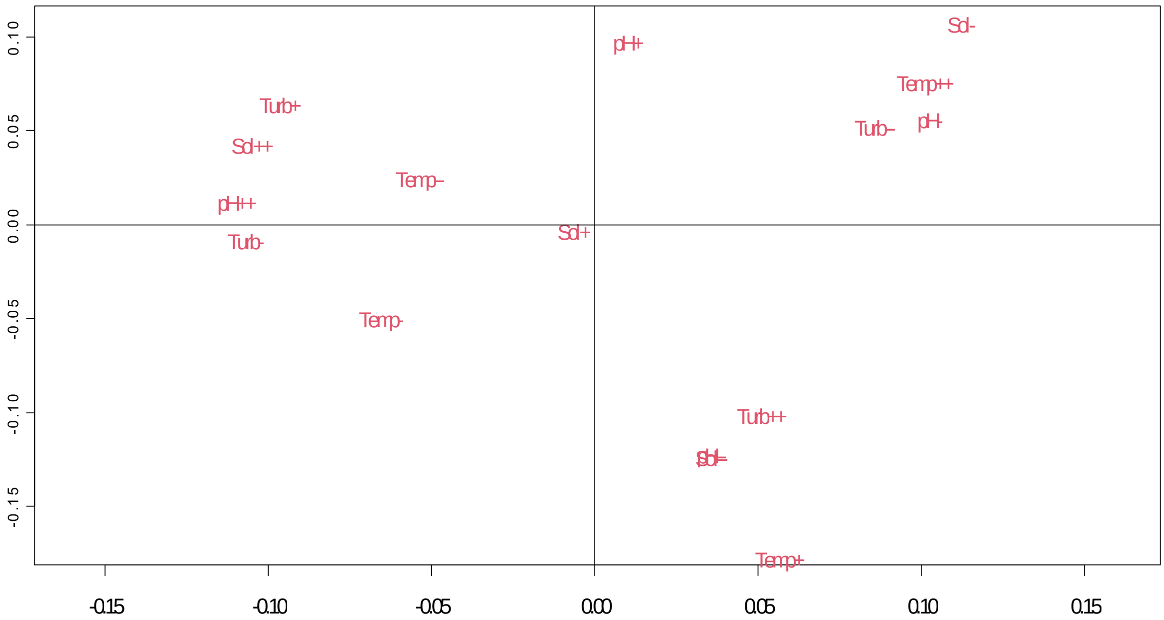

Figure 4. Multiple Correspondence Analysis of Physicochemical Parameters of Spring Water

In Figure 4, the symbols for each physicochemical parameter represent different levels of that variable. The symbol "++" denotes a high value, "+" indicates a moderately high value, "–" corresponds to a moderately low value, and "––" represents a low value. Thus, pH++ denotes a high pH, whereas pH–– denotes a low pH. Similarly, Sol++ and Sol–– correspond to high and low concentrations of dissolved solids, respectively; Temp++ and Temp–– indicate high and low water temperatures, respectively; and Turb++ and Turb–– represent high and low turbidity levels, respectively. These classifications facilitate the interpretation of the associations revealed by the Multiple Correspondence Analysis (MCA) and enable the identification of the dominant physicochemical characteristics of each group of sampling sites.

To analyze the results in Figure 4, which illustrates the results of a Multiple Correspondence Analysis.

Analysis (MCA), we will examine each quadrant and the associated physicochemical parameters. Each quadrant represents correlations between different parameter classes. In the first quadrant, we find parameters with the following signs: pH+ (slightly high), Sol- (slightly low dissolved solids), Temp++ (high temperature), and Turb-- (low turbidity). This means that in this area, the sites exhibit moderate alkalinity, relatively low dissolved solids, high temperatures, and low turbidity. This profile could indicate a relatively healthy environment, with good water clarity, which is favorable for aquatic life. However, high temperatures could signal a potential impact of thermal discharges. The second quadrant shows that the signs include pH++ (high), Soil++ (high dissolved solids), Turb+ (slightly elevated turbidity), and Temp- (low temperature). This indicates that these sites have high alkalinity, high dissolved solids levels, and slightly higher turbidity, while the temperature is relatively low. This combination could suggest significant inputs of agricultural drainage water or industrial discharge, increasing pH and dissolved solids.

The increased turbidity could impair water clarity, thus affecting photosynthesis in aquatic organisms. The third quadrant shows Turb- (slightly low turbidity), Soil+ (slightly elevated dissolved solids), and Temp- (slightly low temperature). This suggests sites where turbidity is slightly below normal, dissolved solids are slightly elevated, and temperature is relatively low. This profile could indicate a less disturbed environment, where water clarity is acceptable, but sediment inputs may still occur, potentially from agricultural activities or natural processes. Finally, in the fourth quadrant, we have Turb++ (high turbidity), pH-- (low pH), and Sol-- (low dissolved solids), as well as Temp+ (slightly elevated temperature). The results here show sites with high turbidity, indicating a high sediment load; low pH, suggesting acidification; and low dissolved solids. This could signal adverse environmental conditions in which water quality is compromised. High turbidity can impair water clarity, thereby affecting aquatic life, while low pH may result from acidic fallout or sewage discharges. Figure 4 provides a clear view of the conditions at different sites based on physicochemical parameters. Each quadrant reveals distinct characteristics that may be related to anthropogenic or natural impacts. Analyzing these results is essential for a better understanding of water quality and the measures needed to protect these resources. Variations in the observed parameters indicate the need for continued monitoring to assess the effects of human activities and environmental phenomena.

3.1.3. Validation of PCA results

To further explore the relationships between microbial contamination indicators and the main gradients identified by principal component analysis (PCA), correlations between the water quality variables and the first two principal components (PC1 and PC2) were calculated. The results are presented in Table 6, which summarizes the correlation coefficients and their associated p-values for total viable counts, total coliforms, and nitrate concentrations.

Table 6. Correlations between principal component analysis (PCA) and water quality indicators

Variable	Correlation with PC1 (r)	p-value (PC1)	Correlation with PC2 (r)	p-value (PC2)
Total viable count (GeT)	0.3573	0.3107	0.4707	0.1698
Total coliforms (ColiT)	0.4331	0.2112	0.2877	0.4203
Nitrate (Nit)	0.5868	0.0745	0.0145	0.9684

To assess the reliability of the PCA results, correlations were calculated between PCA scores and independent indicators of water quality. PC1 showed a moderate positive correlation with total germs ($r = 0.357$) and total coliforms ($r = 0.433$), while the strongest relationship was observed with nitrate concentration ($r = 0.587$, $p = 0.075$). Although the correlations were not statistically significant at the 5% level, likely due to the limited sample size ($n = 10$), the results suggest that PC1 reflects a physicochemical gradient associated with nitrate enrichment and microbiological contamination. PC2 exhibited weaker relationships with the external indicators, indicating a lower contribution to water-quality variability.

3.1.4. Student's t-tests and Pearson correlation coefficients

The p-values obtained from Student's *t*-test revealed several statistically significant relationships among the physicochemical parameters of the spring waters. Electrical conductivity (Cond.) showed highly significant associations with total dissolved solids (TDS) ($p = 2.2 \times 10^{-16}$), total hardness (TH) ($p = 0.002$), calcium (Ca) ($p = 0.001$), total alkalinity (TAC) ($p < 0.001$), bicarbonate (HCO_3^-) ($p = 0.001$), and magnesium (Mg) ($p = 0.003$). Similarly, bicarbonate concentrations were significantly associated with TDS ($p = 0.001$), TH ($p = 6.36 \times 10^{-7}$), conductivity ($p = 0.001$), calcium ($p = 9.27 \times 10^{-5}$), TAC ($p = 8.36 \times 10^{-9}$), magnesium ($p = 7.62 \times 10^{-11}$), and phosphate ($p = 0.003$). Total alkalinity exhibited significant relationships with TDS ($p < 0.001$), TH ($p = 4.04 \times 10^{-6}$), conductivity ($p < 0.001$), calcium ($p = 1.85 \times 10^{-8}$), bicarbonate ($p = 8.36 \times 10^{-9}$), magnesium ($p < 0.001$), and phosphate ($p = 0.005$). Calcium was significantly correlated with TDS ($p = 0.001$), TH ($p = 9.92 \times 10^{-8}$), conductivity ($p = 0.001$), TAC ($p = 1.85 \times 10^{-8}$), bicarbonate ($p = 9.27 \times 10^{-5}$), magnesium ($p = 2.75 \times 10^{-5}$), and phosphate ($p = 0.003$). Likewise, magnesium displayed significant associations with TDS ($p = 0.003$), TH ($p = 3.37 \times 10^{-5}$), conductivity ($p = 0.003$), calcium ($p = 2.75 \times 10^{-5}$), TAC ($p < 0.001$), bicarbonate ($p = 7.62 \times 10^{-11}$), and phosphate ($p = 0.024$). Total hardness was significantly related to TDS ($p < 0.001$), conductivity ($p = 0.002$), calcium ($p = 9.92 \times 10^{-8}$), TAC ($p = 4.04 \times 10^{-6}$), bicarbonate ($p = 6.36 \times 10^{-7}$), magnesium ($p = 3.37 \times 10^{-5}$), and phosphate ($p = 0.002$). Furthermore, pH showed significant relationships with TDS ($p = 0.005$), conductivity ($p = 0.005$), calcium ($p = 0.039$), TAC ($p = 0.022$), and bicarbonate ($p = 0.037$). Significant differences were also observed between phosphate and TAC ($p = 0.005$), bicarbonate ($p = 0.003$), calcium ($p = 0.003$), magnesium ($p = 0.024$), and TH ($p = 0.002$). In contrast, sodium, molybdenum, sulfate, iron, and chloride generally showed non-significant relationships with most other parameters ($p > 0.05$), indicating weaker interactions within the hydrochemical system. To better understand the relationships among the physicochemical characteristics of the spring waters, a Pearson correlation analysis was performed. This analysis enabled the identification of the degree of association among the water quality parameters and highlighted potential hydrochemical interactions. The correlation coefficients obtained are presented in Table 7.

Table 7. Pearson correlation analysis of physicochemical parameters of different spring waters

Variables	Cond.	T	pH	Sols	Turb.	TAC	Mo	Bic	Chl	Sul	Nit	Phos	Na	K	Ca	Mg	TH	Fer
Cond.	1																	
T	-0.37	1																
pH	0.80	-0.02	1															
Sols	0.99	-0.38	0.802	1														
Turb.	-0.22	0.07	-0.42	-0.22	1													
TAC	0.91	-0.49	0.708	0.915	-0.09	1												
Mo	0.13	0.54	0.142	0.13	0.26	0.026	1											
Bic	0.87	-0.48	0.66	0.87	-0.06	0.99	-0.002	1										
Chl	0.34	0.06	0.36	0.24	-0.37	-0.02	0.28	-0.13	1									
Sul	-0.13	0.29	-0.19	-0.13	0.14	0.05	0.20	0.098	-0.44	1								
Nit	0.52	-0.55	0.245	0.53	-0.26	0.587	0.023	0.58	0.104	-0.12	1							
Phos	0.59	-0.57	0.33	0.59	0.298	0.803	-0.28	0.83	-0.42	0.19	0.44	1						
Na	0.35	0.06	0.368	0.35	-0.37	-0.01	0.28	-0.12	0.999	-0.42	-0.10	-0.4	1					
K	0.14	-0.12	0.268	0.14	-0.33	-0.10	0.17	-0.19	0.744	-0.79	- 0.23	-0.36	0.73	1				
Ca	0.88	-0.47	0.657	0.88	-0.07	0.992	-0.001	-0.998	-0.12	0.14	0.57	0.83	-0.04	-0.2	1			
Mg	0.83	-0.36	0.59	0.83	-0.07	0.93	0.09	0.93	-0.06	0.33	0.48	0.70	-0.11	-0.35	0.95	1		
TH	0.84	-0.46	0.59	0.835	-0.06	0.969	-0.04	0.98	-0.15	0.20	0.62	0.85	-0.14	-0.28	0.988	0.95	1	
Fer	0.28	0.07	-0.00	0.28	0.14	0.16	0.72	0.12	0.44	0.15	0.20	-0.21	0.45	0.12	0.14	0.32	0.12	1

On the other hand, some parameters, such as sodium (Na) and iron (Fe), have higher p-values, indicating less significant relationships. Turning to Table 7, which presents Pearson correlations among the parameters, the coefficients reveal marked positive and negative relationships. For example, conductivity (Cond.) shows a very strong correlation with other variables such as calcium (Ca) (0.88) and bicarbonate (Bic) (0.869), indicating that an increase in conductivity is often associated with increases in these parameters.

3.2. Discussion

3.2.1 Principal Component Analysis (PCA)

In addition to assessing the quality of water resources in the Lukula region, the study presented above meets the assessment requirements as reported by the following researchers. In this analysis, we used principal component analysis (PCA) to assess spring water quality in the Lukula region. The results show that the first principal component (PC1) explains 42.6% of the total variance in the data, making it the most informative component. The highest loading coefficient on PC1 is for total alkalinity (TAC), at 0.339, indicating that it plays a crucial role in spring water quality. High alkalinity can influence the water's ability to buffer pH variations. These findings are supported by Alam et al. (2021), who highlighted the importance of alkalinity in water resource management, and by Benkov et al. (2023), who also found a significant relationship between alkalinity and water quality. Other influential variables on PC1 include calcium (Ca) and soils, with coefficients of 0.335 and 0.325, respectively. This highlights the importance of minerals in water, which are essential for the health of aquatic ecosystems. Bicarbonates (Bic) and total hydrocarbon traces (THT), with coefficients of 0.334 and 0.329, also show a significant influence on the first component. Variables such as pH (0.239) and nitrates (Nit) (0.201) contribute to PC1, but to a lesser extent, indicating a less pronounced influence on the variance. These observations are consistent with the work of Giao (2022), who demonstrated that minerals play a key role in water quality, and Mohammed and Al-Obaidi (2021), who noted similar interactions between physicochemical parameters. The second principal component (PC2) explains 21.4% of the total variance and provides complementary information. The most significant variable for PC2 is chlorine (Chl), with a coefficient of 0.459, indicating that chlorine is a determining factor in water quality, possibly due to anthropogenic inputs. Sodium (Na), with a coefficient of 0.458, also exerts a strong influence, highlighting salinity as a key criterion for PC2. Potassium (K), with a coefficient of 0.395, also suggests its importance in water source classification.

This finding is supported by the research of Ustaoğlu et al. (2021), who also found that chlorine and sodium are important indicators of water quality. The standard deviation values (2.919 for PC1 and 2.071 for PC2) indicate that PC1 captures greater variation in the data than PC2. The proportions of cumulative variance, 42.6% for PC1 and 64.1% for PC2, indicate that these two principal components capture a significant portion of the information about the quality of water sources. This suggests that the mineral and chemical composition of water is mainly influenced by environmental and anthropogenic factors. Previous work, such as that of Giao et al. (2022), has also highlighted that chemical composition is essential for assessing water quality, thus reinforcing our results. Analysis of individuals' projections onto PC1 and PC2 revealed concerning characteristics of the water sources. Kizinga is located on the negative axis of

PC1, where it is strongly associated with high levels of sodium, chlorine, and other contaminants. This combination indicates pollution of water sources by mineral salts and microbiological contaminants, often linked to intensive agricultural practices and industrial discharges. These factors raise concerns about water quality and public health. Previous studies, such as that of Alam et al. (2021), confirm that contaminants from human activities can seriously affect water quality, corroborating the results observed in Kizinga. Kiala Mongo I, positioned on the positive axis of PC1, exhibits high turbidity, indicating degradation of water quality in its sources. This high turbidity could be caused by sediment inputs from erosion and human activities, which affect photosynthesis and aquatic life. Research by Giao et al. (2022) also shows that turbidity is a key indicator of degraded water quality, reinforcing the importance of monitoring this variable. Kibinonga shares similar characteristics to Kiala Mongo I, with high turbidity levels. Excessive turbidity in spring waters can harm aquatic life by reducing water clarity and disrupting the food chain. Results from a study by Bhatti et al. (2019) also demonstrate that turbidity is often associated with negative impacts on aquatic ecosystems, supporting the findings for Kibinonga. Kinduka, located on the negative axis of PC2, suggests potential hydrocarbon and mineral pollution of spring waters. The presence of bicarbonates and high alkalinity may indicate inputs of agricultural drainage or industrial discharges, impacting water quality. Studies such as Kumar et al. (2018) highlight that agricultural pollutants can severely affect water resources, consistent with observations at Kinduka. Kiyangi I is also linked to more complex pollution, involving nitrates and coliforms, resulting from the use of agricultural fertilizers and microbiological contamination. This compromises the potability of source water and poses risks to public health. Research by Gad et al. (2020) corroborates these findings by showing that nitrates from agricultural activities are a major problem in many water sources. Kilonde, associated with phosphorus and sulfur, indicates potential pollution from fertilizers or industrial discharges. High phosphorus concentrations can lead to eutrophication of water bodies, affecting aquatic biodiversity. Kilonde I, on the other hand, is associated with elevated temperatures and potassium levels, suggesting thermal pollution likely due to hot industrial water discharges. The work of Xue et al. (2009) points out that high temperatures can decrease oxygen solubility, thereby harming aquatic life, which is consistent with observations at Kilonde I. Tuidi Yila and Kiala Mongo II are also associated with high turbidity, suggesting a high concentration of suspended sediment. This could result from agricultural activities, soil erosion, or wastewater discharges, affecting water clarity and the health of aquatic ecosystems. The results of Okello et al. (2015) reinforce this idea by showing that turbidity is a reliable indicator of environmental impacts related to human activities.

3.2.2 Hierarchical Ascending Classification (HAC)

Analysis of the dendrogram results reveals distinct groupings of water sources, illustrating their relationships. The first group includes the Kingola Fuka, Tuidi Yila, Kiala Mongo I, Kiala Mongo II, Kilonde I, and Kibinonga water sources. Relatively low salinity and hardness levels characterize these springs. This indicates that the water from these springs is softer and less mineralized, which is generally favorable for human consumption and aquatic ecosystem health. The proximity of these springs on the dendrogram suggests that they share similar hydrological characteristics, potentially influenced by common environmental factors, such as soil type or water resource management practices. Studies such as that of Nnorom et al. (2019) corroborate these observations, highlighting that lower salinity levels are often associated with better drinking water quality. The second group is represented by the Kizinga water source, which is distinguished by higher salinity. This characteristic indicates that the water from Kizinga is more mineralized, which may have implications for its use. The increased salinity may make this water less suitable for human consumption without prior treatment. This group is then linked to other springs such as Kiyangi I, Kinduka, and Kilonde, which also have high hardness levels. This suggests that these springs are subject to similar mineral influences, probably due to geological features or agricultural practices in the area. The results of Ajtai et al. (2025) support this idea, showing that higher salinity levels can adversely affect water quality, particularly for domestic uses. The dendrogram clearly shows two main groups of water sources based on salinity and hardness. The first group, consisting of freshwater sources, is potentially more suitable for human consumption, while the second group, with more saline and hard water, could present challenges for its use. This classification can guide water resource management efforts and inform decisions regarding the treatment and use of these different sources. Research such as that of Benkov et al. (2023) reinforces the importance of these classifications for sustainable water resource management, demonstrating how statistical approaches can help assess water quality.

3.2.3 Multiple Correspondence Factor Analysis (MCFA)

Analysis of the physicochemical characteristics of the spring water samples reveals significant variations, indicating potential pollution impacts on water quality. The results for Kinduka show a high pH, which could signal increased alkalinity, potentially due to agricultural drainage water inputs or industrial discharges. The low temperature and slightly elevated turbidity also suggest sediment inputs, often associated with soil erosion. The slight increase in dissolved solids may affect the water's potability. These observations are corroborated by work such as that of Giao et al. (2022), which highlights the impact of agricultural practices on water quality. The results from Kilonde are similar to those from Kinduka, suggesting comparable potential pollution. This could be attributed to agricultural practices in the region, where fertilizer use can increase pH and dissolved solids. The studies by Gupta et al. (2023) support this hypothesis by demonstrating that agricultural practices can adversely affect water quality in similar settings. Kilonde I has a low pH, which could indicate acidification, potentially caused by acid fallout or acidic wastewater discharges. The high turbidity suggests a significant presence of sediment or organic matter, which can harm aquatic life by reducing water clarity. The results of Ustaoğlu et al. (2021) reinforce this idea, showing that water acidification can have serious consequences for aquatic ecosystems. Kiala Mongo I exhibits the least favorable water quality characteristics, with low pH and turbidity. This could indicate good water clarity, but low dissolved solids levels could harm aquatic biodiversity. The source of pollution could be less anthropogenic and more natural. The results of Benkov et al. (2023) suggest that some water sources can maintain good quality despite moderate pollution levels, depending on environmental influences. Kibinonga indicates that elevated temperatures, combined with low pH, may signal thermal pollution, likely due to hot-water discharges from nearby industries. High turbidity could result from degradation of surrounding soils, affecting water quality. The work of Okello et al. (2015) confirms that thermal pollution can severely affect aquatic populations, favoring some species while harming others. The results show significant variation between sites, reflecting the impacts of pollution. Sites with low pH and high turbidity pose risks to aquatic life, affecting the photosynthesis and respiration of aquatic organisms. In addition, high levels of dissolved solids can degrade water quality, making it unsafe for consumption. Temperature fluctuations can influence the distribution of aquatic species, favoring those that tolerate warmer conditions. Studies such as Shil et al. (2019) highlight the importance of monitoring these variables to maintain the health of aquatic ecosystems. Analysis of Multiple Correspondence Analysis (MCA) results provides insight into the correlations between different classes of physicochemical parameters. Each quadrant reveals distinct characteristics of the sampling sites. In the first quadrant, parameters include slightly elevated pH (pH+), low dissolved solids (Sol-), elevated temperature (Temp++), and low turbidity (Turb--). This suggests that sites in this area exhibit moderate alkalinity, good water clarity, and elevated temperatures, potentially due to thermal discharges. This combination indicates a relatively healthy environment, favorable to aquatic life. Research such as Giao et al. (2022) shows that low turbidity levels and high temperatures may indicate favorable conditions but still require attention to thermal impacts. The second quadrant presents parameters such as high pH (pH++), high dissolved solids (Soil++), slightly elevated turbidity (Turb+), and relatively low temperature (Temp-). This profile could indicate significant inputs of agricultural drainage water or industrial discharges, leading to increases in pH and dissolved solids. This increase in turbidity can impair water clarity, thereby affecting photosynthesis in aquatic organisms. Studies such as those by Najah et al. (2021) confirm that industrial discharges often contribute to increased dissolved solids and turbidity in aquatic ecosystems. The third quadrant shows slightly low turbidity (Turb-), slightly high dissolved solids (Soil+), and relatively low temperature (Temp-). This indicates sites where water clarity is acceptable but sediment inputs may be present, likely due to agricultural activities or natural processes. This pattern could reflect a less disturbed environment. The results by Kumar et al. (2022) support the idea that moderate levels of sedimentation can often be tolerated without seriously harming aquatic ecosystems. The fourth quadrant exhibits high turbidity (Turb++), low pH (pH--), low dissolved solids (Soil--), and slightly elevated temperature (Temp+). This profile indicates sites with high sediment presence, potential acidification, and compromised water quality. High turbidity can impair water clarity, affecting aquatic life, while low pH could result from acidic fallout or wastewater discharge. The work of Okello et al. (2015) highlights that such conditions can adversely affect aquatic biodiversity, making the management of these sites crucial.

3.2.4 Validation of PCA results

The validation of the PCA results was performed by examining the relationships between the principal components and selected water-quality indicators, namely total viable counts, total coliforms, and nitrate concentrations. The analysis revealed that PC1 was moderately correlated with total viable counts ($r = 0.357$) and total coliforms ($r = 0.433$), while the strongest association was observed with nitrate concentration ($r = 0.587$; $p = 0.075$). Although these correlations did not reach the conventional significance threshold of 5%, they suggest that PC1 captures a gradient related to nutrient enrichment and microbiological contamination. In contrast, PC2 exhibited weaker correlations with all tested indicators, indicating a more limited contribution to the variability in water quality degradation. These findings are consistent with previous studies reporting that the first principal component often reflects the dominant pollution processes affecting water resources. For example, Tripathi and Singal (2019) demonstrated that PCA effectively identifies the key parameters controlling water quality and that the first component generally represents the major sources of environmental variation. Similarly, Gad et al. (2020) found that PC1 was strongly associated with indicators of mineralization and contamination, confirming its usefulness for describing water-quality gradients in surface waters. Comparable observations were reported by Nnorom et al. (2019) for spring waters in southeastern Nigeria, where the first principal component was linked to anthropogenic influences and indicators of water deterioration. The positive relationship observed between PC1 and nitrate concentration in the present study supports the interpretation that nutrient enrichment contributes substantially to the hydrochemical variability of the investigated spring waters. Although the correlation was not statistically significant, its relatively high magnitude suggests an underlying environmental relationship. Similar associations between PCA-derived components and nitrate contamination have been reported by Xue et al. (2009), who emphasized the importance of nitrate as an indicator of agricultural and domestic pollution sources, and by Zhang et al. (2022), who identified nutrient enrichment as a major factor influencing water-quality variability in river systems. The moderate correlations between PC1 and microbiological indicators further reinforce this interpretation. Studies conducted by Giao (2022) and Giao et al. (2022) showed that PCA frequently groups microbiological contamination variables with nutrient-related parameters, reflecting common pollution sources such as agricultural runoff, wastewater discharge, and inadequate sanitation practices. Likewise, Alam et al. (2021) reported that bacterial contamination and nutrient enrichment often contribute simultaneously to the first principal component in water-quality assessments around landfill environments. The absence of statistically significant correlations may be explained by the relatively small number of sampling sites included in the present investigation. This limitation has been discussed by Berglund et al. (2017) and Worley and Powers (2016), who noted that PCA validation procedures may produce weaker statistical significance when applied to small datasets, even when meaningful environmental relationships exist. Nevertheless, the observed correlation patterns remain informative, as they are consistent with the hydrochemical and microbiological characteristics identified by PCA. The weaker relationships observed for PC2 suggest that this component represents secondary sources of variability that are less directly related to microbial contamination and nitrate enrichment. Similar conclusions were reached by Benkov et al. (2023), Ajtai et al. (2025), and Nath Roy et al. (2024), who found that higher-order principal components generally describe localized or less influential environmental processes compared with the dominant gradients captured by PC1. Overall, the validation analysis supports interpreting PC1 as the principal axis controlling water-quality variation in the studied springs. The associations observed with nitrate concentration, total viable counts, and total coliforms are in agreement with findings reported in previous water-quality studies and provide additional evidence that the

PCA model successfully captured the main environmental factors influencing the hydrochemical and microbiological characteristics of the investigated water sources.

3.2.5 Student's t-Tests and Pearson Correlation Coefficients

The analysis of the p-values from Student's t-tests for the physicochemical parameters presented highlights significant differences across several comparisons. Conductivity (Cond.) has an extremely low p-value, approaching 2.2 × 10-16 when compared with other parameters such as calcium (Ca) and total alkali content (TAC). This indicates that the observed differences are highly significant, suggesting a strong correlation between these variables. This observation is crucial, as high conductivity can indicate the presence of dissolved ions, as noted in studies such as Gad et al. (2020), where conductivity is used as a key indicator of water quality. Furthermore, the p-value for bicarbonates (HCO_3) and total alkali content (TAC) is also very low, reinforcing the idea that these elements are closely related in water samples. This is consistent with the results of other research, such as Kumar et al. (2018), which shows that bicarbonates play an important role in regulating water alkalinity, thereby influencing overall water quality. On the other hand, some parameters such as sodium (Na) and iron (Fe) have higher p-values, indicating less significant relationships. This could suggest that these elements are less strongly correlated with other parameters than those mentioned above. Studies, such as Nnorom et al. (2019), indicate that the presence of sodium and iron can vary considerably depending on water sources and environmental conditions, which could explain the less significant results observed here. Analysis of p-values and Pearson correlations between the different physicochemical parameters highlights significant and varied relationships. Some parameters, such as sodium (Na) and iron (Fe), have higher p-values, indicating less significant relationships. This suggests that these elements are not strongly correlated with other measured parameters. This observation is consistent with the results of Nnorom et al. (2019), who show that the variability of sodium and iron concentrations in water sources may depend on specific environmental factors, thereby complicating their interpretation. On the other hand, in Table 7, conductivity (Cond.) shows a very strong correlation with other variables, including calcium (Ca) (coefficient = 0.88) and bicarbonate (Bic) (coefficient = 0.869). This indicates that an increase in conductivity is often associated with increases in these parameters. These results are supported by studies such as those by Gad et al. (2020), which highlight conductivity as a key indicator of water quality and the presence of dissolved ions. Temperature (T) also shows positive correlations with other parameters, although to a lesser extent, such as pH (0.02). This may suggest interconnected environmental influences that affect both water temperature and acidity. Research by Shil et al. (2019) argues that temperature can influence the solubility of gases and nutrients in water, thereby affecting the dynamics of aquatic ecosystems. In contrast, turbidity (Turb.) displays weak correlations with other variables, suggesting the other measured parameters do not strongly influence it. This may indicate that turbidity is affected by distinct factors, such as sediment inputs or human activities, which is also discussed in Bhatti et al. (2019), emphasizing the impacts of anthropogenic activities on stream turbidity. Another notable aspect is the strong correlation between total alkali content (TAC) and bicarbonate (Bic), with a coefficient of 0.993. This confirms that these two variables are closely related within the context of spring waters. The studies of Kumar et al. (2022) reinforce this idea, stating that bicarbonate plays a crucial role in regulating alkalinity and can serve as an indicator of water quality. This study provides an in-depth understanding of water resource quality in the Lukula region, employing multivariate analysis methods to identify the types of pollution present. By integrating advanced statistical techniques, such as principal component analysis and multiple correspondence analysis, this research highlights the complex interactions among various physicochemical parameters and their impacts on public health and the environment. The results provide a solid basis for sustainable water resource management strategies, emphasizing the importance of continuous monitoring and evaluation of management practices. The discussion reveals that some water sources, notably Kizinga, are heavily polluted with mineral salts and microbiological contaminants. Other sources, such as Kinduka and Kilonde, also exhibit concerning levels of pollution, often linked to agricultural and industrial activities. These findings underscore the urgent need for action to preserve water resource quality in the Lukula region. Targeted interventions are essential to reduce pollution and ensure access to quality drinking water for the local population. A limitation of the present study should nevertheless be acknowledged. The interpretation of the statistical relationships identified through Student's *t*-tests, Pearson correlations, PCA, HAC, and MCA was performed primarily based on the dataset collected from the investigated springs and comparisons with findings reported in the literature. Although multivariate statistical techniques are widely recognized as powerful tools for identifying hydrochemical patterns and pollution sources (Tripathi & Singal, 2019; Gad et al., 2020; Ustaoğlu et al., 2021; Benkov et al., 2023), their interpretation is strengthened when supported by long-term monitoring data and regional reference datasets. In the absence of external datasets from neighboring areas and multi-season observations, it was not possible to directly validate the spatial patterns observed or evaluate their temporal persistence. Similar limitations have been highlighted in studies emphasizing the need for spatiotemporal monitoring to improve the reliability of water quality assessments (Okello et al., 2015; Giao et al., 2022; Gupta et al., 2023; Ajtai et al., 2025). Consequently, the relationships identified among physicochemical and microbiological parameters should be interpreted with caution, as they may partly reflect local environmental conditions prevailing during the sampling period. Furthermore, the absence of complementary source-tracing approaches, such as isotopic techniques, limits the precise identification of nitrate origins and contamination pathways (Xue et al., 2009). Future investigations should therefore incorporate larger datasets, seasonal and multi-year monitoring campaigns, and regional comparative analyses to strengthen the robustness of the interpretations and improve the understanding of the factors controlling groundwater quality in the Lukula region. Such an approach is recommended in recent studies applying multivariate methods to water quality assessment (Nnorom et al., 2019; Nath Roy et al., 2024; Zhang et al., 2022). To this end, it is recommended to strengthen water quality monitoring programs. This includes establishing regular monitoring protocols for the physicochemical and microbiological parameters of water sources. At the same time, awareness campaigns should be initiated among farmers to promote sustainable agricultural practices, thereby helping to minimize negative impacts on water resources. Furthermore, future research should explore the impact of climate change on water quality and the efficiency of wastewater treatment systems. These studies will help improve water resource management and protect aquatic ecosystems. The findings of this study, corroborated by previous research, including Gad et al. (2020) and Nnorom et al. (2019), reinforce the need for an integrated approach to ensure a safe and sustainable drinking water supply for the population.

4. Conclusion

This study employed Principal Component Analysis (PCA), Hierarchical Ascending Clustering (HAC), and Multiple Correspondence Analysis (MCA) to characterize spring water quality and identify the dominant types of pollution in the Lukula region. PCA showed that the first two principal components explained 64.1% of the total variance, with PC1 accounting for 42.6% and mainly associated with mineralization parameters, including total alkalinity, bicarbonate, calcium, magnesium, total hardness, conductivity, and dissolved solids. PC2 explained 21.4% of the variance and was primarily related to chloride, sodium, potassium, sulfate, and phosphate concentrations. The HAC distinguished three groups of springs according to their hydrochemical characteristics: a group of highly mineralized waters (Kinduka, Kilonde, and Kiyangi I), a group of low- to moderately mineralized waters (Kilonde I, Kiala Mongo I, Kibinonga, Kingola Fuka, Tuidi Yila, and Kiala Mongo II), and the Kizinga spring, which formed a separate cluster because of its atypical chemical composition. The MCA identified different pollution patterns across the study area, including mineral enrichment, acidification, sediment-related pollution, and microbiological contamination. Among the investigated sites, Kizinga, Kiyangi I, Kinduka, and Kilonde exhibited the highest levels of pollution indicators, characterized by elevated concentrations of dissolved ions, nutrients, and microbiological contaminants. Overall, the results demonstrate significant spatial variability in spring water quality and confirm that both natural hydrogeochemical processes and anthropogenic activities influence the physicochemical and microbiological characteristics of water resources in the Lukula Health Zone. The lack of external datasets and long-term observations should be acknowledged as a limitation, as it prevents independent validation of the identified hydrochemical patterns and assessment of their long-term persistence. These findings provide a scientific basis for improving the monitoring and management of drinking water resources in the region. These findings highlight the need for continuous physicochemical and microbiological monitoring, protection of spring catchments, improved sanitation measures, and sustainable agricultural practices. Future studies should expand monitoring efforts and incorporate additional analyses to support effective and sustainable water resource management in the region.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgements

We want to express our deep gratitude to the Lukula Territorial Chief for his invaluable support throughout this study. His collaboration was essential in facilitating access to the research sites and mobilizing the local community. We also thank the local authorities and residents of the region for their commitment and cooperation, which greatly contributed to the success of our work.

References

Ajtai, I., Anton, A., Roba, C., Botezan, C., Piștea, I., Oprea, M., & Baciu, C. (2025). Impact of wastewater on surface water quality and relevance of water reuse in agriculture using a comprehensive methodology based on PCA and specific indices. *Water*, 17(13), 2011.

Alam, R., Ahmed, Z., Seefat, S. M., & Nahin, K. T. K. (2021). Assessment of surface water quality around a landfill using a multivariate statistical method, Sylhet, Bangladesh. *Nanotechnology for Environmental Engineering*, 15, 100422.

Benkov, I., Varbanov, M., Venelinov, T., & Tsakovski, S. (2023). Principal component analysis and the water quality index—A powerful tool for surface water quality assessment: A case study on Struma river catchment, Bulgaria. *Water*, 15, 1961.

Berglund, A. E., Welsh, E. A., & Eschrich, S. A. (2017). Characteristics and validation techniques of gene expression signatures based on PCA. *International Journal of Genomics*, 2017(1), 2354564.

Bhatti, E.-H., Khan, M. M., Shah, S. A. R., Raza, S. S., Shoaib, M., & Adnan, M. (2019). Water quality dynamics: An impact assessment process for water resources management. *Processes*, 7, 102.

Gad, M., Elsayed, S., Moghanm, F. S., Almarshadi, M. H., Alshammari, A. S., Khedher, K. M., Aïd, E. M., & Hussein, H. (2020). Combining water quality indices and multivariate modeling to assess surface water quality in the northern Nile Delta, Egypt. *Water*, 12, 2142.

Giao, N. (2022). Analyzing surface water quality using multivariate statistical approaches: A case study in Ca Mau Peninsula, Vietnam. *Pollution*, 8, 463–477.

Giao, N., Dan, T., Ni, D., Anh, P., & Nhien, H. (2022). Spatiotemporal variations in physicochemical and biological properties of surface water using statistical analyses in Vinh Long Province, Vietnam. *Water*, 14, 2200.

Gupta, A. K., Kumar, A., Maurya, U. K., Singh, D., Islam, S., Rathore, A. C., & Madhu, M. (2023). Comprehensive spatiotemporal calibration of surface water quality in the Hindon River, a tributary of the Yamuna River, India: Adopting a multivariate statistical approach. *Environmental Science and Pollution Research*, 30(55), 116804–116830.

Kumar, R., Dutt, V., Raina, A., & Sharma, N. (2022). Spatial assessment of water quality of a mountain stream in northwest India using multivariate statistical techniques. *Environmental Monitoring and Assessment*, 194, 785.

Kumar, V., Sharma, A., Kumar, R., Bhardwaj, R., Thukral, A. K., & Rodrigo-Comino, J. (2018). Assessment of heavy metal pollution in three different Indian water bodies by combining multivariate analysis and water pollution indices. *Human and Ecological Risk Assessment*, 26, 1–16.

Mampuya Nzita, A., Musanga Matondo, J., Mwanda Mizengi, L., Biabia Mumpele, P., Nzadi Bilonda, J., Dituba Ngoma, G., & Tangou Tabou, T. (2025). Identification of pollution in the Lukaya River and its tributaries: An innovative approach. *Journal of Hydraulic and Water Engineering*, 2(2), 37–50.

Mohammed, R., & Al-Obaidi, B. (2021). Influence of municipal wastewater effluent treatability on surface water quality assessment based on the Nemerow pollution index using an artificial neural network. *IOP Conference Series: Earth and Environmental Science*, 877, 012008.

Najah, A., Teo, F. Y., Chow, M. F., Huang, Y. F., Latif, S. D., Abdullah, S., Ismail, M., & El-Shafie, A. (2021). Surface water quality status and prediction during the movement control operation order under the COVID-19 pandemic: Case studies from Malaysia. *International Journal of Environmental Science and Technology*, 18, 1009–1018.

Nath Roy, B., Roy, H., Rahman, K. S., Mahmud, F., Bhuiyan, M. M. K., Hasan, M., Bhuiyan, A.-A. K., Hasan, M., Mahbub, M. S., & Jahedi, R. M. (2024). Principal component analysis integrated water quality index modeling for Dhaka-based rivers. *City and Environment Interactions*, 23, 100150.

Nnorom, I. C., Ewuzie, U. S. A., & Eze, S. O. (2019). Multivariate statistical approach and water quality assessment of natural springs and other drinking water sources in southeastern Nigeria. *Heliyon*, 5, e01123.

Nzita, A. M., Mizengi, L. M., Ndiwulu, G. W., Ngoma, G. D., Umba-di-Mbudi, C. N. Z., Ngimbi, R. M., & Phuati, E. P. (2024). Statistical analysis of electrical component aging in group 6 of the Inga 2 hydroelectric plant following abrupt shutdowns. *Computational and Experimental Research in Materials and Renewable Energy*, 7(2), 78–92.

Nzita, A. M., Phuati, E. P., Ngimbi, R. M., Ngoma, G. D., & Mavungu, N. M. (2024). Investigation of lateral vibrations in turbine-generator unit 5 of the Inga 2 hydroelectric power plant. *IgMin Research*, 2(4), 255–260.

Okello, C., Tomasello, B., Greggio, N., Wambiji, N., & Antonellini, M. (2015). Impact of population growth and climate change on freshwater resources in Lamu Island, Kenya. *Water*, 7, 1264–1290.

Shil, S., Singh, U. K., & Mehta, P. (2019). Water quality assessment of a tropical river using water quality index (WQI), multivariate statistical techniques and GIS. *Applied Water Science*, 9, 168.

Shima Ngoy, T., Monsengo N'solo, R., Kongota Isasi, E., & Misamu Ndandu, S. (2020). Physicochemical, bacteriological and hydrochemical facies characterization of water from developed springs in the Lukula health zone, Kongo Central Province, Congo. *Congo Science*, 8(3), 154–160.

Tripathi, M., & Singal, S. K. (2019). Use of principal component analysis for parameter selection for development of a novel water quality index: A case study of river Ganga India. *Ecological Indicators*, 96, 430–436.