Asian Journal of Water, Environment and Pollution. Vol. 23, No. 1 (2026), pp. 211-221.
doi: 10.36922/AJWEP025320248

ORIGINAL RESEARCH ARTICLE

Spatial assessment of water quality in a hierarchically structured river system using stream order classification and multivariate statistical techniques: A case study from Tunduma, Tanzania

Matungwa William[1]*, Zacharia Katambara[2], and Omari Shegwando[2]

[1]Department of Earth Sciences, College of Science and Technical Education, Mbeya University of Science and Technology, Mbeya, Tanzania

[2]Department of Civil Engineering, College of Engineering and Technology, Mbeya University of Science and Technology, Mbeya, Tanzania

Corresponding author: Matungwa William (matungwa.william@must.ac.tz)

Received: August 8, 2025; 1st revised: September 4, 2025; 2nd revised: September 25, 2025; 3rd revised: October 12, 2025; Accepted: October 20, 2025; Published online: November 17, 2025

Abstract: Water quality assessment is essential for understanding pollutant dynamics, supporting evidence-based watershed management, and protecting public health. While numerous studies have utilized statistical and modeling approaches, limited attention has been paid to how stream order influences water quality variability, particularly in urban catchments of sub-Saharan Africa. This study investigates spatial patterns of water quality in a hierarchically structured stream network in Tunduma, Tanzania, by integrating Strahler stream order classification with multivariate statistical techniques, based on monthly monitoring of six surface water points over 12 months ($n = 72$) during both wet and dry seasons to analyze physicochemical, nutrient, and microbial parameters. Hierarchical cluster analysis, combined with Pearson correlation matrices and significance testing, was employed to assess pollutant similarity and accumulation patterns across different stream orders. Results revealed that phosphate (PO_4^{3-}) concentrations ranged from 0.42 to 1.49 mg/L and nitrate (NO_3^-) levels ranged from 4.3 to 13.2 mg/L. Strong positive correlations ($r > 0.95$) were observed among ion-derived parameters, such as electrical conductivity, total dissolved solids, and the concentrations of calcium (Ca^{2+}) and magnesium (Mg^{2+}). Third-order stream segments exhibited elevated concentrations of total suspended solids (0.990), biochemical oxygen demand (0.982), and microbial indicators, with fecal coliforms of 0–5 CFU/100 mL and total coliforms of 0–18 CFU/100 mL, reflecting cumulative pollutant loading in downstream reaches. The integration of Strahler stream ordering and cluster-based analytics enabled the identification of pollution hotspots and revealed the critical role of hydrological connectivity in shaping water quality trends. This research contributes a novel spatial–statistical framework for stream-based water quality assessment in East African urban contexts, offering practical insights for catchment-scale pollution control and resource management.

Keywords: Water quality assessment; Hierarchical cluster analysis; Strahler stream order; Nutrient and microbial pollution; Catchment management

1. Introduction

Water quality monitoring plays a crucial role in evaluating ecosystem health, safeguarding public health, and enabling sustainable water resource management.[1] The spatial variation of river water quality is critical for designing effective treatment strategies and formulating evidence-based policy decisions, and water quality management has become a global priority, particularly in the context of the United Nations Sustainable Development Goal 6, which aims to ensure universal access to clean water and sustainable sanitation.[2] Degraded water quality can severely impact public health, aquatic ecosystems, and socioeconomic development, underscoring the urgent need for continuous monitoring and integrated management approaches.[2] In sub-Saharan Africa, particularly in rapidly urbanizing areas, water bodies are increasingly affected by land use change, population pressure, and limited wastewater infrastructure.[3] Urban streams in such contexts often receive complex mixtures of pollutants from both point sources (e.g., domestic effluent) and non-point sources (e.g., agricultural runoff), leading to significant spatial variability in water quality across stream networks.[4-6]

In Tanzania, research on water quality has traditionally focused on site-specific or point-based assessments[7,8] with limited emphasis on how hydrological structure, particularly stream order, modulates pollution accumulation. However, stream networks are inherently hierarchical, and higher-order streams often integrate cumulative upstream influences.[9] Studies elsewhere have demonstrated how solute transport, catchment memory, and nutrient buildup correlate with stream order.[10] This hierarchical perspective remains underexplored in Tanzanian urban catchments. This study addresses this critical research gap by focusing on Tunduma, a fast-growing transboundary municipality in southwestern Tanzania. The area presents a dendritic stream network, subject to multiple anthropogenic pressures, including informal settlements, poorly managed sanitation, and agricultural intensification. Despite these challenges, there has been no systematic assessment of how spatial stream hierarchy interacts with land use, pollutant distribution, and water quality gradients across stream segments. To fill this gap, we propose a novel spatial–statistical framework that integrates the Strahler stream ordering method with multivariate analysis techniques, including hierarchical cluster analysis (HCA) and correlation heat mapping. This framework allows us to examine physicochemical, nutrient, and microbial parameters across six spatially distributed sampling points from first-to third-order streams in Tunduma. We hypothesize that higher-order streams accumulate greater concentrations of ionic and particulate pollutants due to upstream connectivity, land use pressures, and convergence of tributary flows. Our approach extends prior Tanzanian work by focusing on stream hierarchy and by applying advanced multivariate clustering to delineate spatial water quality patterns. This method is supported by recent advancements in explainable machine learning for environmental assessment[11] in screening drivers of water quality. It reflects a growing emphasis on coupling spatial hydrology with statistical diagnostics for targeted water quality management. This study introduces a combined spatial–statistical framework, integrating Strahler stream ordering with HCA, to evaluate nutrient, microbial, and physicochemical parameter variations to reveal spatial drivers of pollution.

We aim to contribute to a novel application of stream order theory, multivariate clustering, and spatial mapping in a Tanzanian urban watershed, with implications for future hydrological modeling, pollution tracking, and integrated catchment management.

By linking hydrological structure, landscape pressures, and multivariate patterns, this study contributes valuable insights for evidence-based interventions, such as pollution hotspot mapping, riparian buffer restoration, and stream-specific mitigation strategies. Furthermore, our findings serve as a foundation for future modeling efforts that incorporate non-linear hydrologic responses and catchment memory.[12,13]

By integrating multivariate statistical techniques on correlation analysis and the Strahler stream ordering method, the research aims to identify pollution hotspots and priority intervention zones. Ultimately, the findings support catchment-level decision-making for water quality protection and sustainable resource planning in the transboundary urban context of Tunduma as well as other related Tanzanian urban areas.

2. Materials and methods

2.1. Study area

This study was carried out in Tunduma, located in the southern highlands zone of Tanzania within the Songwe region, at the Tanzania–Zambia border (Figure 1). The town covers an estimated area of 87.5 km^2 and has a rapidly growing population of 219,309 inhabitants, reflecting an exceptional annual growth rate of 13% that is exerting increasing pressure on local water resources. The topography varies from just below

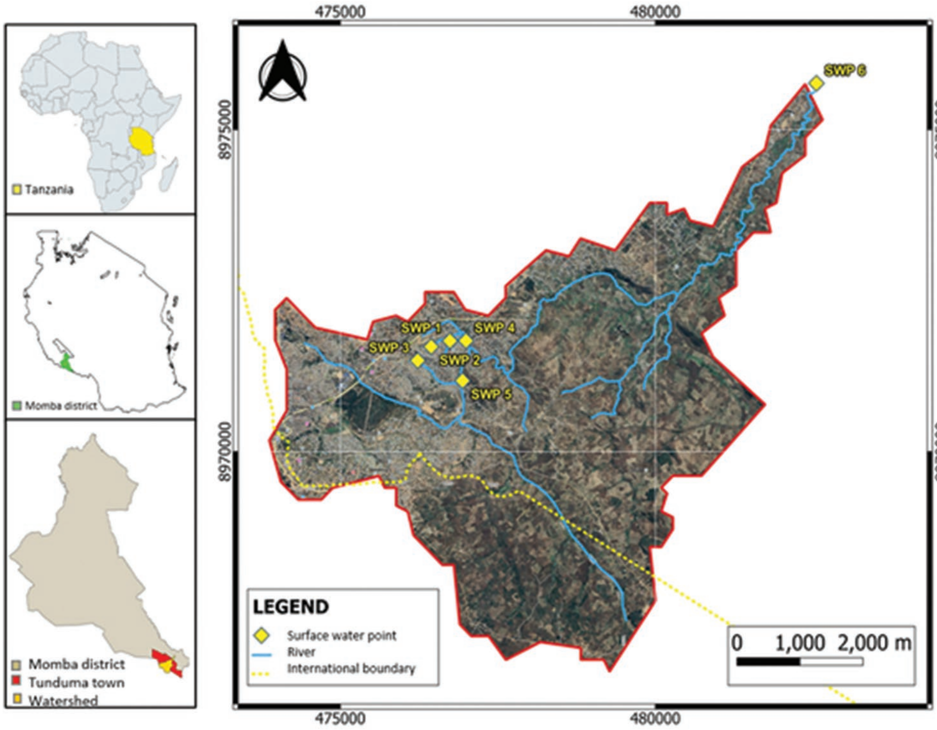

Figure 1. Location of the study area (Half-London ward) in Tunduma, Tanzania

1,500 m to above 1,600 m above mean sea level, producing a diverse landscape that influences drainage characteristics and hydrological pathways. The focal point of the investigation was the Half-London ward, a sub-catchment that is strategically positioned within the river network and experiences significant human–environment interactions. Climatically, the area is characterized by a unimodal rainfall regime extending from November to mid-May, with peak precipitation typically occurring in January and February.[14] The annual mean precipitation is approximately 1,006 mm, while the mean annual temperature stands at 20.5°C, ranging from a minimum of 6.5°C in October to a maximum of 29.0°C in July.[15]

2.2. River network classification

The hierarchical configuration of the Tunduma stream network was delineated using the Strahler stream ordering method, a widely recognized hydrological classification system[10] that defines stream hierarchy based on confluence patterns.[16] In this approach, first-order streams represent unbranched tributaries; when two streams of the same order merge, they form a stream of the next higher order. This methodology facilitates understanding of flow convergence, hydrological connectivity, and cumulative pollutant loading potential across a catchment.[17]

Within the study area, surface water point (SWP) 1, SWP 2, and SWP 3 were classified as first-order streams, each representing isolated upstream catchments with no tributary inflow. SWP 4, located at the confluence of SWP 1 and SWP 2, was designated as a second-order stream. SWP 5, downstream of SWP 4 and SWP 3, represented a third-order stream, capturing the cumulative influence of the previous tributaries. Finally, SWP 6, situated at the basin outlet, integrates the flow contributions of the entire network, thereby functioning as the terminal node in the stream hierarchy.

A schematic stream diagram (Figure 2) illustrates this dendritic drainage pattern, typical of semi-urban catchments in Tunduma, Tanzania. The diagram includes flow direction arrows and node labels to clarify the stream order transitions and sampling locations.

This structured spatial approach enhanced not only the selection of sampling points but also the interpretation of pollutant transport dynamics, particularly how hydrological connectivity and stream order influence nutrient accumulation, microbial loads, and physicochemical dispersion. Such classification is essential for targeted water quality monitoring and aligns with the best practices for integrated watershed assessment.[11,18] A methodological overview of sampling, analysis, and statistics is shown in Figure 3.

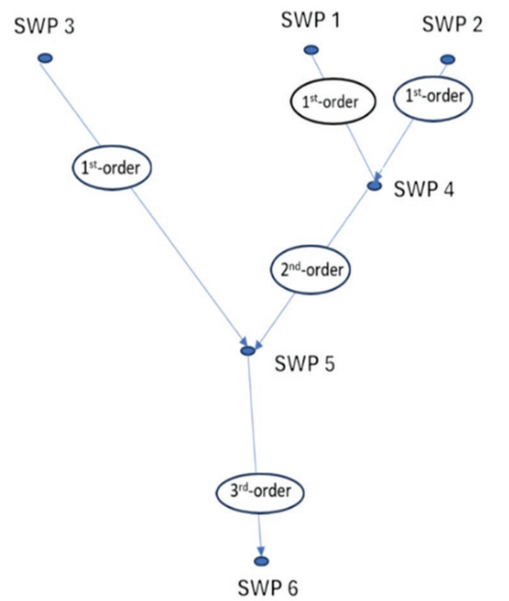

Figure 2. The schematic diagram of the Strahler stream order

Abbreviation: SWP: Surface water point.

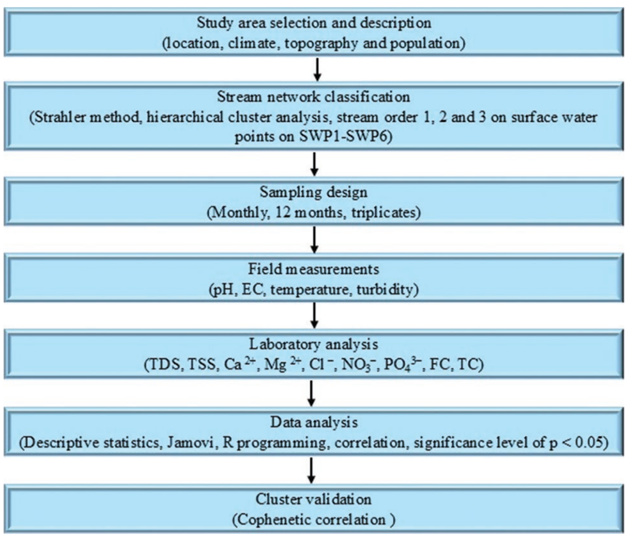

Figure 3. Methodology flow diagram

Abbreviations: DO: Dissolved oxygen; EC: Electrical conductivity; HCA: Hierarchical clustering analysis; SWP: Surface water point; TDS: Total dissolved solids

2.3. Sampling procedure and sample preservation

Before the commencement of the study, ethical approval and authorization for surface water sampling were obtained from the Tunduma Town Council. Water samples were collected monthly over 12 months (from July 2022 to May 2023) using grab sampling techniques. Before sample collection, all containers were acid-washed and rinsed with deionized water. On-site, bottles were rinsed three times with stream water to avoid contamination. Samples intended for microbial analysis were stored in sterile 500 mL polypropylene containers, immediately placed in ice-cooled boxes at 4°C, and transported to the laboratory for analysis within 6 h.[7] For physicochemical and nutrient analysis, water samples were collected in acid-cleaned polyethylene bottles. Nutrient samples were filtered on-site using Whatman GF/C microfiber filters (1.2 μm). Nitric acid (HNO_3) was added until a pH of <2 was achieved for total dissolved solids (TDSs) and metal preservation, following the protocols outlined in the 2017 American Public Health Association (APHA) report.

2.4. Instrumentation and analytical methods

In-situ parameters pH, electrical conductivity (EC), and TDS were measured using a Hanna Instruments (United States) HI 98194 portable multiparameter meter, calibrated daily with standard buffer solutions (pH 4.01, 7.01, and 10.01) and conductivity standard (1,413 μS/cm) per manufacturer guidelines. Turbidity was measured using a HACH 2100Q portable nephelometer (Hach Company, United States), calibrated with Formazin primary standards and operated in accordance with Environmental Protection Agency-approved nephelometry procedures. Nutrient concentrations, specifically phosphate (PO_4^{3-}) and nitrate (NO_3^-), were determined spectrophotometrically using a Shimadzu (Japan) UV-1800 ultraviolet–visible (UV–VIS) spectrophotometer. Phosphate was analyzed through the ascorbic acid method (APHA Method 4500-P), while nitrate was assessed using the cadmium reduction method (APHA Method 4500-NO_3), as outlined in APHA. Microbial water quality was evaluated by quantifying fecal coliforms (FC) and total coliforms (TC) through the membrane filtration technique (APHA Methods 9222 B and E), employing 0.45 μm cellulose nitrate membranes. Filtered samples were incubated on m-Endo agar for TC at 35°C and m-FC agar for FC at 44.5°C. Colony counts were recorded as colony-forming units (CFU) per 100 mL. These methods ensured analytical rigor and comparability with global water quality monitoring standards.[19]

2.5. Quality assurance and quality control

To ensure analytical reliability, 10% of all samples were collected in triplicate. Field blanks and equipment blanks were also included in each sampling round. Calibration was verified with control standards before and after analysis. Limits of detection (LOD) and limits of quantification were determined based on three and ten times the standard deviation of blank measurements, respectively. The LODs were as follows: PO_4^{3-} = 0.01 mg/L, NO_3^- = 0.05 mg/L, and FC/TC = 1 CFU/100 mL. Relative standard deviation for replicate samples remained below 10% for all tested parameters.

2.6. Data analysis and software

All data analyses were conducted using R version 4.3.1 and Jamovi version 2.6.44. R was used for correlation analysis, hierarchical clustering, and heatmap visualizations. Jamovi was used for normality testing, transformations, and statistical summaries.

To ensure statistical robustness, all raw water quality data (Table S1) were initially assessed for normality using the Shapiro–Wilk test at a significance level of $p<0.05$ in Jamovi v2.6.44. Parameters that failed normality assumptions were $\log_{10}$-transformed before further analysis to normalize the distribution and reduce skewness. Pearson correlation coefficients were computed using R v4.3.1 to evaluate interrelationships among the measured parameters. To control for false discovery rate (FDR) arising from multiple comparisons, the Benjamini–Hochberg procedure was applied to adjust p-values, ensuring reliable interpretation of significance levels. A correlation heatmap annotated with significance levels ($p<0.05$ and $p<0.01$).

For pattern recognition, HCA was performed on standardized (z-score) data to assess spatial similarity of water quality across stream orders. The analysis employed the Mahalanobis distance metric, which accounts for correlations among variables, and Ward's minimum variance method for linkage.[20] Cluster validity was evaluated using two complementary approaches:

i. The Cophenetic correlation coefficient (CCC), which measured the correlation between the original distance matrix and the dendrogram (CCC = 0.87, indicating strong cluster reliability)
ii. The average silhouette width (ASW), which quantified the degree of cohesion and separation among clusters (ASW = 0.62, reflecting moderately strong clustering structure).

The HCA analysis was performed to group sampling points according to similarities in water quality profiles, providing insight into spatial relationships and patterns of water quality variation.[21]

All analyses were conducted in alignment with stream order classifications to evaluate how hierarchical hydrological structure influences pollutant distribution. Graphical representations, including dendrograms and heatmaps, were interpreted in the context of spatial positioning, stream convergence, and land use gradients.

3. Results

3.1. Pearson correlation patterns across stream orders

To assess spatial relationships among water quality parameters, a correlation analysis was conducted using Pearson's correlation coefficients across five stream order transitions: first-order (SWPs 1–4, 2–4, 3–5), second-order (SWPs 4–5), and third-order (SWPs 5–6). Figure 4 presents the correlation heatmap, summarizing relationships among 20 key physicochemical and microbial parameters, including pH, EC, TDS, turbidity, total suspended solids (TSS), nutrients (PO_4^{3-}, NO_3^-), major ions (Ca^{2+}, Mg^{2+}, Fe^{2+}, Cl^-, SO_4^{2-}), microbial indicators (FC, TC), and biochemical oxygen demand.

Strong positive correlations ($r > 0.95$) were observed among ionic parameters such as EC, TDS, Ca^{2+}, and salinity, particularly in the third-order stream segment (SWP 5–6). For example, EC–TDS correlation reached $r = 0.997$, while EC–Ca^{2+} reached $r = 0.964$, indicating consistent ionic accumulation downstream. This reflects not only cumulative contributions from upstream tributaries but also hydrological connectivity and potential groundwater-surface water exchange, commonly reported in urban tropical catchments.[2,22]

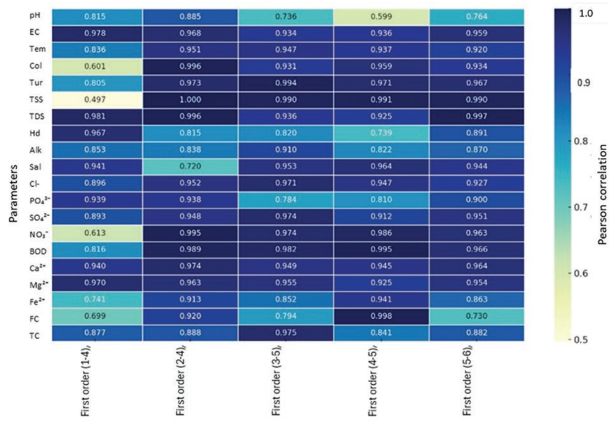

Figure 4. Heatmap of Pearson correlation coefficients (r) across five stream order transitions. Darker colors represent stronger correlations. Significance levels ($p<0.05$) were used to determine statistical robustness.

Abbreviations: Alk: Alkalinity; BOD: Biochemical oxygen demand; Col: Color; DO: Dissolved oxygen; EC: Electrical conductivity; FC: Fecal coliform; HCA: Hierarchical clustering analysis; Hd: Hardness; Sal: Salinity; SWP: Surface water point; TC: Total coliform; TDS: Total dissolved solids; Tem: Temperature; TSS: Total suspended solids; Tur: Turbidity.

Similarly, turbidity and TSS displayed very high correlation in higher-order streams (r = 0.990–0.994), signifying particulate loading likely derived from erosion, stormwater input, and unpaved surface runoff in densely populated zones. For instance, in the second-order stream (SWP 4–5), turbidity and TSS correlations peaked at r = 0.991, highlighting sediment transport linked to land-use disturbance.

Interestingly, nutrient correlations varied by stream order. Nitrate (NO_3^-) and phosphate (PO_4^{3-}) showed slightly lower correlations (r = 0.784–0.810) in higher-order streams, possibly due to dilution from tributary confluences or in-stream biogeochemical processes, including denitrification and plant uptake. These patterns align with the river continuum concept, which suggests that nutrient concentrations may fluctuate longitudinally due to both source inputs and in-stream assimilation.[23,24]

Microbial indicators such as FC and TC showed moderate to strong correlations in downstream sites (e.g., FC–TC = r = 0.882 at SWPs 5–6), reflecting human and animal waste inflows from upstream communities. The FC correlation dropped slightly (r = 0.730) at the third order, which could be attributed to environmental inactivation or sediment adsorption.

The data confirm that stream order is a key spatial driver of water quality variability. Downstream reaches (SWP 5 and SWP 6), acting as integrators, exhibited higher clustering of correlated parameters, reflecting pollutant accumulation, as previously noted in other structured river systems.[11,12,20]

3.2. HCA

3.2.1. Nutrients cluster analysis

The HCA, framed within the watershed's hydrological structure, provided deeper insights into the spatial variability of nutrient dynamics across the six surface water sampling points (SWP 1–SWP 6). The dendrogram (Figure 5) and clustered heatmap (Figure 6) show clear grouping patterns based on phosphate (PO_4^{3-}) and nitrate (NO_3^-) concentrations. Phosphate levels ranged from 0.42 mg/L at SWP 1 to 1.49 mg/L at SWP 6, while nitrate concentrations varied between 4.3 mg/L (SWP 1) and 13.2 mg/L (SWP 5). Elevated phosphate concentrations were notably recorded at SWP 3 (1.21 mg/L), SWP 5 (1.42 mg/L), and SWP 6 (1.49 mg/L), whereas relatively low concentrations were observed at SWP 1 (0.44 mg/L) and SWP 2 (0.47 mg/L), reflecting minimal upstream anthropogenic influence.

Similarly, nitrate concentrations peaked at SWP 5 (13.2 mg/L) and SWP 6 (12.8 mg/L), while significantly

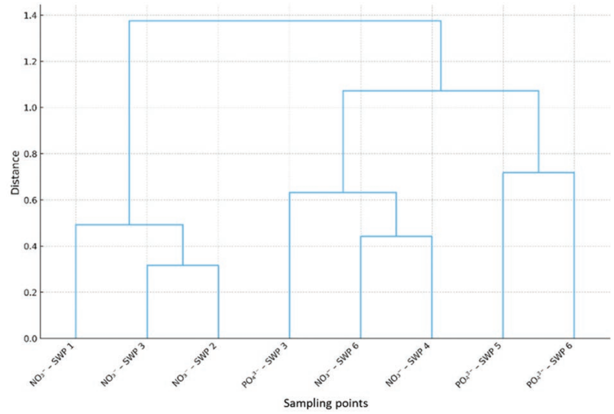

Figure 5. Dendrogram showing hierarchical clustering of phosphate (PO_4^{3-}) and nitrate (NO_3^-) concentrations across six surface water points (SWP 1–SWP 6) in the Tunduma watershed

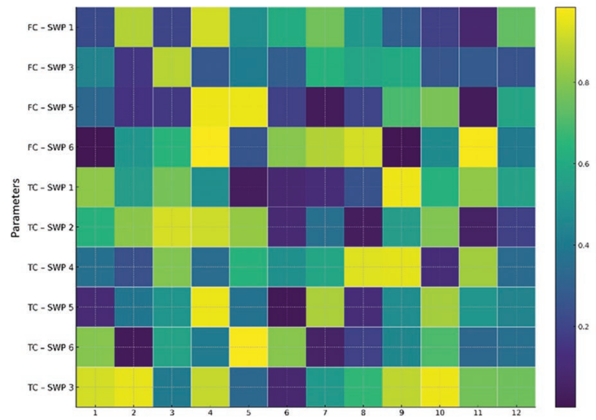

Figure 6. Clustered heatmap of normalized phosphate (PO_4^{3-}) and nitrate (NO_3^-) concentrations for six sampling surface water points (SWP 1–SWP 6)

lower levels were measured at SWP 1 (4.3 mg/L) and SWP 2 (4.7 mg/L), demonstrating nutrient accumulation downstream due to agricultural runoff, domestic effluent discharges, and surface washout from informal settlements. Based on the dendrogram, HCA clustered SWP 3, SWP 5, and SWP 6 together, forming Cluster 1, representing high-nutrient enrichment zones, while SWP 1, SWP 2, and SWP 4 formed Cluster 2, characterized by lower nutrient concentrations (≤0.70 mg/L phosphate; ≤7.5 mg/L nitrate).

Cluster validity metrics confirmed robust results, with a CCC of 0.87 and an ASW of 0.62, supporting the reliability and interpretability of the clustering. Pearson correlation analysis supported these findings, showing strong positive correlations ($r \geq 0.95$, p<0.01) among phosphate, nitrate, and ionic parameters (EC, TDS, and Ca^{2+}) in downstream reaches (SWP 5 and SWP 6). In contrast, weaker associations ($r \leq 0.65$) were detected in headwater streams (SWP 1 and SWP 2). These results highlight the integrator role of downstream sites in cumulative pollutant loading and the influence of stream order on nutrient clustering. By identifying high-risk nutrient hotspots, this study directly informs local strategies aimed at achieving cleaner and safer freshwater systems under the Sustainable Development Goal 6 framework.

These results are consistent with findings from other tropical and subtropical river systems.[20,25] Du *et al*.[20] and Withers and Jarvie[25] similarly observed downstream phosphate levels surpassing 1.5 mg/L in China's Liao River Basin, attributed to intensive agriculture and wastewater inputs, while Hellar-Kihampa *et al*.[7] and Alavaisha *et al*.[26] documented phosphate levels up to 1.7 mg/L and nitrate above 10 mg/L in Tanzania's Pangani River Basin, particularly near confluences. In addition, Makumbura *et al*.[11] applied clustering and machine learning in Tanzanian urban catchments and affirmed that stream order facilitates nutrient concentration and clustering through cumulative tributary effects. In international contexts, such as the Han River basin in South Korea, Beaulieu *et al*.[27] found that urban and agricultural land uses exerted pronounced negative impacts on phosphate and nitrate levels. Clustering and vegetation buffering effects were evident at both micro- and sub-catchment scales, revealing similar spatial and anthropogenic drivers.

These similarities confirm that tributary convergence and hydrological connectivity are dominant drivers of nutrient enrichment across diverse regions, reinforcing the relevance of stream-order-based cluster analyses for catchment management.

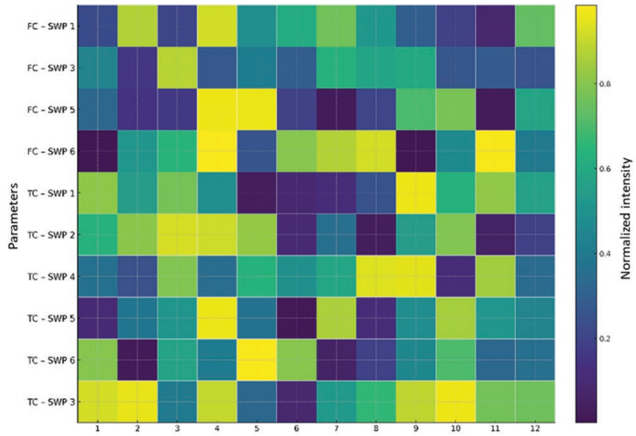

Figure 7. Heatmap for microbial (fecal coliform and total coliform) parameters

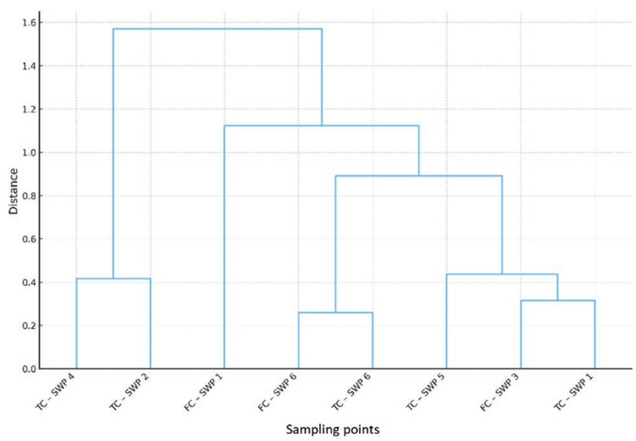

Figure 8. Dendrogram for microbial (FC and TC), showing the percentage of similarity among water quality parameters

Abbreviations: FC: Fecal coliform; SWP: Surface water point; TC: Total coliform.

3.2.2. Microbial cluster analysis

The HCA, supported by heatmap visualization (Figure 7) and dendrogram representation (Figure 8), revealed distinct spatial clustering patterns of FC and TC across SWP 1–SWP 6. The correlation data showed that FC exhibited its highest positive association in the second-order stream between SWP 4 and SWP 5 ($r = 0.998$), indicating significant downstream microbial accumulation, likely driven by wastewater inputs and inadequate sanitation. Similarly, strong correlations were observed in first-order segments, particularly between SWP 2 and SWP 4 ($r = 0.920$) and between SWP 3 and SWP 5 ($r = 0.794$).

Total coliforms displayed notable clustering patterns, with the highest correlation in first-order streams between SWP 3 and SWP 5 ($r = 0.975$), followed by SWP 1 and SWP 4 ($r = 0.877$) and SWP 2 and SWP 4 ($r = 0.888$). In contrast, the lowest TC similarity was observed at the SWP 4–SWP 5 segment ($r = 0.841$), suggesting dilution effects or localized treatment in this midstream reach.

HCA grouped SWP 3, SWP 4, and SWP 5 into Cluster 1, characterized by higher microbial contamination, where FC levels exceeded 5 CFU/100 mL and TC levels reached above 21 CFU/100 mL. Cluster 2 included SWP 1, SWP 2, and SWP 6, showing comparatively lower microbial loads, with FC averaging <2 CFU/100 mL and TC <15 CFU/100 mL. The dendrogram separated sampling points into two clusters: Cluster 1 (SWP 3, SWP 4, SWP 5) as high-contamination sites, and Cluster 2 (SWP 1, SWP 2, SWP 6) as relatively low-contamination sites, in which cluster validity measures supported these results (CCC = 0.84; ASW = 0.59). These patterns indicate that microbial contamination is intensely concentrated in midstream and downstream regions, corresponding to areas with dense human settlements, livestock activity, and improper waste management practices.

Comparisons with previous studies support these findings. In Tanzania's Pangani Basin, Hellar-Kihampa *et al*.[7] reported FC levels above 5 CFU/100 mL in downstream reaches due to similar domestic and agricultural pressures. Likewise, Lam *et al*.[28] identified clustering of coliforms in urban-impacted streams in coastal waters, consistent with the high correlations observed in this study. Globally, Fang *et al*.[29] found similar microbial clustering in Singaporean urban catchments, linking it to storm-driven surface washouts and poor wastewater infrastructure.

Overall, the strong correlations ($r \geq 0.95$) in downstream reaches indicate that hydrological connectivity amplifies pollutant integration, making SWP 4 and SWP 5 critical microbial hotspots. These findings highlight the need for localized sanitation upgrades, livestock management interventions, and riparian buffer restoration to minimize microbial risks and improve downstream water quality.

3.2.3. Physicochemical cluster analysis

The HCA, combined with heatmap visualization (Figure 9) and dendrogram grouping (Figure 10), revealed evident spatial heterogeneity in the distribution of four key physicochemical parameters: EC, TDS, turbidity, and TSS across the six sampling locations (SWP 1–SWP 6).

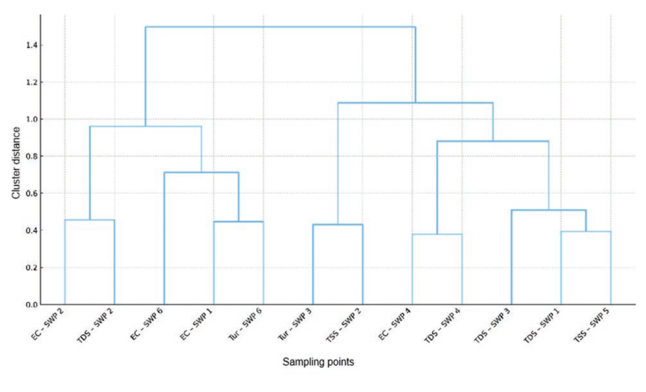

Figure 9. Dendrogram for physicochemical (EC, TDS, Tur, and TSS) showing the percentage of similarity among water quality parameters

Abbreviations: EC: Electrical conductivity; SWP: Surface water point; TDS: Total dissolved solids; TSS: Total suspended solids; Tur: Turbidity.

The heatmap illustrates distinct clustering patterns, where EC ranged from 142 µS/cm at SWP 1 to 524 µS/cm at SWP 5, and TDS varied between 92 mg/L (SWP 1) and 328 mg/L (SWP 5). These elevated values at SWP 5 and SWP 6 reflect ionic enrichment in downstream sites, likely influenced by geochemical leaching, wastewater infiltration, and agricultural return flows. The strong correlation between EC and TDS ($r \geq 0.97$, $p<0.01$) indicates that dissolved ionic loads predominantly drive salinity patterns within the watershed. The dendrogram divided variables into two clusters: dissolved ionic group (EC and TDS) and particulate group (TSS and turbidity), in which cluster validation indicated consistent results (CCC = 0.89; ASW = 0.64), confirming clear partitioning between dissolved and particulate processes. These findings highlight how ionic variables represent relatively stable geochemical signals, whereas particulate loads are more variable, episodic, and linked to runoff and erosion processes. Similar findings were reported in the Fuji River Basin, Japan,[30] where EC and TDS co-varied due to natural weathering and fertilizer inputs.

In contrast, turbidity and TSS exhibited localized peaks, with maximum turbidity reaching 27.5 NTU and TSS exceeding 185 mg/L at SWP 3 and SWP 4. These elevated concentrations correspond to erosion-prone midstream zones, where increased sediment resuspension results from unvegetated banks, poor drainage, and storm-induced runoff. This pattern aligns with studies in tropical watersheds, where

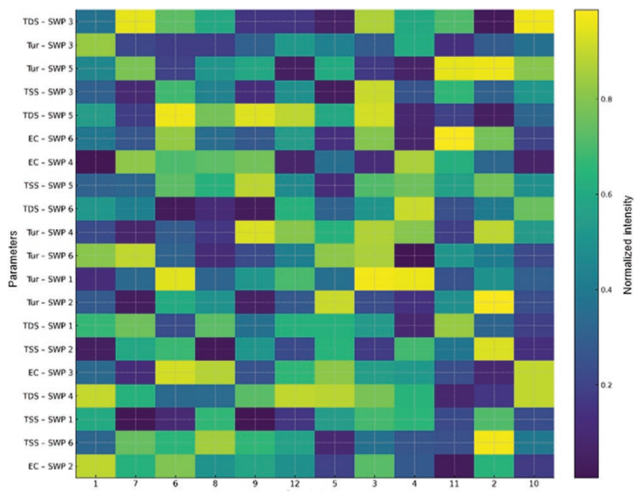

Figure 10. Heatmap for physicochemical (EC, TDS, Tur, and TSS) parameters

Abbreviations: EC: Electrical conductivity; SWP: Surface water point; TDS: Total dissolved solids; TSS: Total suspended solids; Tur: Turbidity.

rainfall-driven sediment mobilization leads to episodic spikes in particulate-associated pollutants.[29,31]

The dendrogram (Figure 10) separated the parameters into two distinct clusters. Cluster 1 contained EC and TDS, representing dissolved-phase indicators primarily influenced by groundwater seepage, mineral weathering, and ionic contributions from domestic sources. Cluster 2 grouped turbidity and TSS, reflecting particulate-associated dynamics linked to surface runoff, unpaved roads, and agricultural activity. Within these clusters, site-specific associations were evident; for instance, EC-SWP 2 and EC-SWP 4 exhibited strong similarity, while TSS-SWP 3 and turbidity-SWP 4 clustered together, indicating spatially constrained sediment transport hotspots.

These findings highlight a fundamental distinction between dissolved and particulate pollutants in the Tunduma watershed. While ionic indicators (EC and TDS) exhibited stable longitudinal trends reflecting geochemical and anthropogenic influence, particulate indicators (turbidity and TSS) were episodic and spatially variable, suggesting higher sensitivity to land use and storm-driven hydrological changes. Similar patterns have been documented in South Korea's Han River Basin, where downstream TDS exceeded 300 mg/L due to groundwater-mineral interactions, while sediment loads peaked during monsoonal floods.

4. Conclusion

This study demonstrated the strong influence of stream hierarchy and hydrological connectivity on water quality patterns within the Tunduma watershed. By integrating Strahler stream order classification with multivariate statistical techniques, we uncovered critical spatial trends in nutrient, microbial, and physicochemical parameters across six stream segments. Notably, phosphate concentrations (0.42–1.49 mg/L) and coliform counts (up to 1,040 CFU/100 mL) were significantly elevated in higher-order downstream reaches (SWP 5 and SWP 6), indicating cumulative impacts from upstream pollution sources. Nitrate concentrations (4.3–13.2 mg/L) showed a more diffuse spatial pattern, consistent with contributions from widespread agricultural runoff and shallow groundwater infiltration.

The HCA and correlation heatmaps confirmed distinct clustering of parameters: ionic solutes (e.g., EC and TDS) were strongly correlated and more stable across stream orders, while particulate-associated variables (e.g., TSS and turbidity) displayed greater spatial variability, particularly at midstream points such as SWP 3 and SWP 4. Microbial contamination was most severe in these mid- to downstream segments, underscoring the combined effect of inadequate sanitation infrastructure, land use intensity, and flow convergence.

These findings validate our hypothesis that stream order plays a pivotal role in shaping pollutant accumulation, transport, and transformation, reinforcing the utility of incorporating hydrological hierarchy into water quality assessments.

This study provides a transferable framework for evaluating water quality dynamics in hierarchically structured stream systems, particularly in urbanizing regions of sub-Saharan Africa. The integration of stream order theory with multivariate analysis can guide evidence-based decision-making, helping policymakers, local governments, and watershed stakeholders prioritize interventions based on spatial pollutant behavior and hydrological structure. Strengthening such data-driven approaches is crucial to achieving sustainable water governance, public health protection, and resilient urban development in line with Sustainable Development Goal 6 on Clean Water and Sanitation.

Acknowledgments

The authors gratefully acknowledge the support of the Mbeya University of Science and Technology (MUST) for facilitating field sampling, laboratory analyses, and data processing. We also thank the local authorities and residents of Tunduma Town for granting access to sampling sites and providing essential background information on land use and water resource practices.

Funding

None.

Conflict of interest

The authors declare no conflicts of interest. The funders had no role in the study design, data collection, data analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization: Matungwa William
Data curation: Omari Shegwando
Formal analysis: Matungwa William, Zacharia Katambara

Investigation: Omari Shegwando
Methodology: Matungwa William
Project administration: Zacharia Katambara
Supervision: Zacharia Katambara
Validation: Omari Shegwando
Visualization: Matungwa William
Writing–original draft: Matungwa William
Writing–review & editing: Zacharia Katambara, Omari Shegwando

Availability of data

The datasets generated and analyzed during the present study are available from the corresponding author on reasonable request.

References

1. Zheng LY, Yu HB, Wang QS. Assessment of temporal and spatial variations in surface water quality using multivariate statistical techniques: A case study of Nenjiang River basin, China. *J Cent South Univ*. 2015;22(10):3770-3780. doi: 10.1007/s11771-015-2921-z
2. Siriwardhana KD, Jayaneththi DI, Herath RD, *et al.* A simplified equation for calculating the water quality index (WQI), Kalu River, Sri Lanka. *Sustainability*. 2023;15(15):12012. doi: 10.3390/su151512012
3. Meyer JL, Paul MJ, Taulbee WK. Stream ecosystem function in urbanizing landscapes. *J North Am Benthol Soc*. 2005;24(3):602-612. doi: 10.1899/04-021.1
4. Duan W, He B, Nover D, *et al.* Water quality assessment and pollution source identification of the eastern poyang lake basin using multivariate statistical methods. *Sustainability*. 2016;8(2):133. doi: 10.3390/su8020133
5. Wu Y, Giri S, Qiu Z. Understanding the spatial distribution of hydrologic sensitive areas in the landscape using soil topographic index approach. *Int Soil Water Conserv Res*. 2016;4(4):278-283. doi: 10.1016/j.iswcr.2016.10.002
6. Payus C, Ann Huey L, Adnan F, *et al*. Impact of extreme drought climate on water security in north borneo: case study of sabah. *Water (Basel)*. 2020;12(4):1135. doi: 10.3390/w12041135
7. Hellar-Kihampa H, De Wael K, Lugwisha E, Van Grieken R. Water quality assessment in the Pangani River basin, Tanzania: Natural and anthropogenic influences on the concentrations of nutrients and inorganic ions. *Int J River Basin Manag*. 2013;11(1):55-75. doi: 10.1080/15715124.2012.759119
8. Chen L, Han L, Tan J, *et al*. Water environmental capacity calculated based on point and non-point source pollution emission intensity under water quality assurance rates in a Tidal river network area. *Int J Environ Res Public Health*. 2019;16(3):428. doi: 10.3390/ijerph16030428
9. Alexander RB, Boyer EW, Smith RA, Schwarz GE, Moore RB. The role of headwater streams in downstream water quality. *J Am Water Resour Assoc*. 2007;43(1):41-59. doi: 10.1111/j.1752-1688.2007.00005.x
10. Ensign SH, Doyle MW. Nutrient spiraling in streams and river networks. *J Geophys Res Biogeosci*. 2006;111(G4). doi: 10.1029/2005JG000114
11. Makumbura RK, Mampitiya L, Rathnayake N, *et al*. Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature. *Results Eng*. 2024;23:102831. doi: 10.1016/j.rineng.2024.102831
12. Rathnayake N, Rathnayake U, Chathuranika I, Dang TL, Hoshino Y. Cascaded-ANFIS to simulate nonlinear rainfall-runoff relationship. *Appl Soft Comput*. 2023;147:110722. doi: 10.1016/j.asoc.2023.110722
13. Kiss T, Fiala K, Sipos G, Szatmári G. Long-term hydrological changes after various river regulation measures: Are we responsible for flow extremes? *Hydrol Res*. 2019;50(2):417-430. doi: 10.2166/nh.2019.095
14. Mbululo Y, Nyihirani F. Climate characteristics over Southern highlands Tanzania. *Atmos Clim Sci*. 2012;2(4):454-463. doi: 10.4236/acs.2012.24039
15. William M, Katambara Z. Assessment of spatial water quality variations in shallow wells using principal component analysis in half London Ward, Tanzania. *J Water Resour Prot*. 2025;17(2):108-143. doi: 10.4236/jwarp.2025.172007
16. VanDerWal J, Shoo LP, Graham C, Williams SE. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecol Modell*. 2009;220(4):589-594. doi: 10.1016/j.ecolmodel.2008.11.010
17. Goyal VC, Jain SK, Pareek N. Water logging and drainage Assessment in Ravi-Tawi irrigation command (J&K) using remote sensing approach. *J Indian Soc Remote Sens*. 2005;33(1):7-15. doi: 10.1007/BF02989986
18. Cabral S, Campos JNB, Silveira C. hydrologic and hydraulic modelling integrated with gis: A study of the Acaraú River basin – CE. *J Urban Environ Eng*. 2015;8:167-174. doi: 10.4090/juee.2014.v8n2.167-174

19. Edokpayi JN, Rogawski ET, Kahler DM, *et al*. Challenges to sustainable safe drinking water: A case study of water quality and use across seasons in rural communities in Limpopo Province, South Africa. *Water*. 2018;10(2):159. doi: 10.3390/w10020159
20. Du X, Shao F, Wu S, Zhang H, Xu S. Water quality assessment with hierarchical cluster analysis based on Mahalanobis distance. *Environ Monit Assess*. 2017;189(7):335. doi: 10.1007/s10661-017-6035-y
21. Athamena A, Gaagai A, Aouissi HA, *et al*. Chemometrics of the environment: Hydrochemical characterization of groundwater in Lioua Plain (North Africa) using time series and multivariate statistical analysis. *Sustainability*. 2022;15(1):20. doi: 10.3390/su15010020
22. Lapworth DJ, Nkhuwa DCW, Okotto-Okotto J, *et al*. Urban groundwater quality in sub-Saharan Africa: Current status and implications for water security and public health. *Hydrogeol J*. 2017;25(4):1093-1116. doi: 10.1007/s10040-016-1516-6
23. Vannote RL, Minshall GW, Cummins KW, Sedell JR, Cushing CE. The river continuum concept. *Can J Fish Aquat Sci*. 1980;37(1):130-137. doi: 10.1139/f80-017
24. Uygur A, Kargı F. Biological nutrient removal from pre-treated landfill leachate in a sequencing batch reactor. *J Environ Manage*. 2004;71(1):9-14. doi: 10.1016/j.jenvman.2004.01.002
25. Withers PJA, Jarvie HP. Delivery and cycling of phosphorus in rivers: A review. *Sci Total Environ*. 2008;400(1-3):379-395. doi: 10.1016/j.scitotenv.2008.08.002
26. Alavaisha E, Lyon SW, Lindborg R. Assessment of water quality across irrigation schemes: A case study ofwetland agriculture impacts in Kilombero Valley, Tanzania. *Water (Switzerland)*. 2019;11(4):671. doi: 10.3390/w11040671
27. Beaulieu JJ, Tank JL, Hamilton SK, *et al.* Nitrous oxide emission from denitrification in stream and river networks. *Proc Natl Acad Sci*. 2011;108(1):214-219. doi: 10.1073/pnas.1011464108
28. Lam QD, Schmalz B, Fohrer N. Assessing the spatial and temporal variations of water quality in lowland areas, Northern Germany. *J Hydrol (Amst)*. 2012;438-439:137-147. doi: 10.1016/j.jhydrol.2012.03.011
29. Fang H, Vergara GGR, Goh SG, Ang CYL, Gu X, Gin KY. Effect of rainfall on the microbial water quality of a tropical Urban catchment. *J Environ Qual*. 2018;47(5):1242-1248. doi: 10.2134/jeq2018.03.0099
30. Shrestha S, Kazama F. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environ Modell Softw*. 2007;22(4):464-475. doi: 10.1016/j.envsoft.2006.02.001
31. Dodds WK, Oakes RM. Headwater influences on downstream water quality. *Environ Manage*. 2008;41(3):367-377. doi: 10.1007/s00267-007-9033-y