Nigerian Journal of Physics (NJP)
ISSN online: 3027-0936
ISSN print: 1595-0611
DOI: https://doi.org/10.62292/njp.v35(s).2026.659
Volume 35(S), 2026

A Computational Investigation of Physicochemical Water Quality Parameters through Condensed Matter Physics and Statistical Modelling: A Case Study of Groundwater and Surface Water Sources in a Basement Complex Terrain, North-Central Nigeria

[1]Silas A. Idowu, *[2]Ozuem V. Edward, [2]Tunde D. Akinnubi and [1]Victor A. Ogunsuyi

[1]Federal University of Technology, Akure, Nigeria.
[2]Anchor University, Lagos, Nigeria.

*Corresponding Author's Email: eozuem@gmail.com

Keywords:
Water quality index,
Principal component analysis,
Hierarchical cluster analysis,
Hydrogen-bond network,
Ion hydration,
Basement complex aquifer,
Groundwater–surface water contrast,
Nigeria.

Submitted: 20 Jul. 2026

Accepted: 3 Aug. 2026

Published: 10 Aug. 2026

ABSTRACT

Water quality assessment in basement complex terrains in sub-Saharan Africa has traditionally relied on descriptive, standards-compliance reporting, with limited investigation of underlying geochemical mechanisms. This study reanalyses physicochemical data — pH, temperature, total dissolved solids (TDS), total hardness, total alkalinity, and lead — from two boreholes and three surface-water streams at the Federal University Lokoja campus, Kogi State, Nigeria, within a condensed matter and statistical physics framework. Data were analysed using descriptive statistics, Pearson correlation, principal component analysis (PCA), hierarchical cluster analysis (HCA), and a weighted arithmetic water quality index (WQI). The first PCA component accounted for 69.1% of total variance and separated groundwater from surface water, a partition corroborated by Ward-linkage HCA. Groundwater exhibited markedly higher TDS (392.5 mg/L) and total hardness (185.0 mg/L) than surface water, consistent with prolonged water–rock interaction and mineral accumulation through transport-limited weathering of crystalline basement aquifers. The only significant correlation was between hardness and alkalinity ($r = -0.95$, $p = 0.014$), suggesting that elevated hardness at one borehole (200 mg/L, above the NSDWQ limit of 150 mg/L) is predominantly non-carbonate. The WQI was highly sensitive to lead, which carries a unit weight exceeding 99% owing to its very low permissible limit (0.01 mg/L); absent lead, all sites would be rated excellent-to-good. The measured parameters are interpreted through hydrogen-bond network reorganisation, Debye–Hückel screening, ion hydration thermodynamics, and diffusion-limited mineral dissolution, providing a framework through which condensed matter physics can offer mechanistic interpretations unavailable within descriptive hydrogeochemistry. The principal limitation is the small dataset ($n = 5$); multivariate outputs are therefore exploratory.

INTRODUCTION

One of the biggest health problems in sub-Saharan Africa is the lack of safe water, and the groundwater from shallow boreholes and surface water that has not been treated are the main domestic water sources in peri-urban and institutional settings. In much of north-central Nigeria, water chemistry is dominated by basement complex terrains composed of granites, gneisses and schists, and controlled more by the kinetics of the interaction between the minerals and the water in the fracture and weathered-zone aquifer than by simple dilution of water. It is thus not only a problem of regulatory compliance that should be addressed in characterising this chemistry, but also a problem of the physics of transport, dissolution and ionic equilibria within a condensed medium of a heterogeneous structure. The traditional approach to water quality characterisation in this area is to sample, analyse in a laboratory and then report the water quality data as a result of the analysis and compare that data to World Health Organisation (WHO) and national water quality regulatory guidelines, typically represented in table format. It's helpful for compliance screening and, by nature, it's a descriptive technique. It is a statement that the parameter is high or

low with respect to a limiting parameter, but no explanation, in mechanistic terms, is given as to how the observed concentration is attained, nor why one parameter is related to another. A borehole with a high hardness and a stream with low alkalinity are typically reported as two separate facts, rather than as a single physical process – water-rock interaction in various conditions – residence-time and thermodynamic boundaries.

The physicochemical nature of matter can be understood in a complementary way from condensed matter and from statistical physics. Liquid water is not a passive solvent but a dynamically hydrogen-bonded network whose local tetrahedral order, dielectric screening, and diffusive relaxation directly govern how dissolved ions are stabilised, transported, and exchanged with surrounding mineral surfaces (Debenedetti and Stanley, 2003; Agmon, 2012). Electrolyte behaviour in natural waters is described by Debye–Hückel screening theory, in which the ionic strength of the solution determines the effective range of electrostatic interactions between charged species; ion hydration thermodynamics determine the free-energy cost of stripping water molecules from a mineral surface into solution during weathering; and diffusion-limited or transport-limited kinetics govern the rate at which dissolved constituents accumulate along a groundwater flow path. None of these are exotic imports into hydrogeochemistry; they are the physical substrate on which hydrogeochemical processes already operate, but they are rarely made explicit in applied water-quality reporting, particularly in resource-limited academic settings where computational tools beyond spreadsheet arithmetic are seldom deployed.

A second limitation of much of the existing literature is computational: physicochemical datasets are frequently reported as raw tables and simple bar charts without multivariate statistical treatment. Principal component analysis (PCA) and hierarchical cluster analysis (HCA) have been used productively elsewhere to resolve latent geochemical structure in groundwater datasets (Taşan *et al.,* 2022; Arıman *et al.,* 2024), and weighted arithmetic water quality indices (WQI) are widely applied to condense multi-parameter datasets into a single interpretable score (Arıman et al., 2024). However, these techniques are seldom combined with an explicit physical interpretation of the resulting components, nor is their statistical fragility at small sample sizes routinely acknowledged, an omission with direct consequences for how confidently their outputs should be read.

This study addresses both gaps using a compact yet well-documented physicochemical dataset from two boreholes and three surface-water streams sampled around the Federal University Lokoja campus, a basement-complex setting in Kogi State, north-central Nigeria. The specific objectives are threefold: (i) to apply a coherent computational workflow descriptive statistics, Pearson correlation, PCA, HCA, and weighted arithmetic WQI to the physicochemical dataset and to report its outputs with appropriate statistical caveats given the small sample size; (ii) to interpret each measured parameter (pH, temperature, total dissolved solids, hardness, alkalinity, and lead) mechanistically through condensed matter and statistical physics concepts, including hydrogen-bond network reorganisation, Debye–Hückel electrostatics, ion hydration thermodynamics, and diffusion-limited weathering kinetics; and (iii) to critically evaluate the robustness of the computational outputs themselves, including a demonstration of how the weighted arithmetic WQI formulation is disproportionately sensitive to trace parameters with very low regulatory thresholds a methodological finding with implications beyond this specific dataset.

The novelty of this contribution is not the water quality dataset itself, which is limited in scope, but the demonstration that even a modest, standards-compliance-oriented dataset can be re-read through a physically grounded, computationally explicit framework that yields testable mechanistic hypotheses for example, that the non-carbonate character of hardness at the most mineralised borehole reflects a specific class of weathering reaction rather than a generic 'hard water' label. We make no claim that the present dataset alone can resolve these hypotheses; rather, we use the computational and physical analysis to identify precisely which additional measurements (major ions, a full heavy-metal panel, and repeated seasonal sampling) would be required to test them, and we state this limitation explicitly throughout rather than allowing the interpretive framework to overstate the evidentiary weight of five data points.

MATERIALS AND METHODS

Study Area

The study area is situated in and around the Federal University Lokoja (FULokoja) campus, Lokoja, Kogi State, north-central Nigeria (approximately 7°48′ N, 6°44′ E), at the confluence of the River Niger and River Benue. The climate is tropical wet-and-dry (savanna) and the terrain is basement complex consisting of granites, gneisses and schists, with local relief of around 110-270 m above sea level. The land use is a combination of institutional, residential, woodland and farmland cover, which can offer diffuse anthropogenic inputs to groundwater and surface water bodies.

Sampling Design and Field Procedures

For the university community, five sampling points were selected to represent the main groundwater and surface water sources; two from the borehole group (BH, BOREHOLE) and three from the stream channel group (Stream 1, Stream 2, Stream 3) located throughout the northern, southern and central-east of the study area.

Samples were all taken from a single sampling campaign (28th of June 2025), with sampling methods followed according to APHA (2017). Borehole samples were taken to represent water from the aquifer after 5 minutes of pumping and stream samples were taken mid-channel, as far as the flow went, to minimise the disturbance of sediment in the water. The collection was done in acid-washed high-density polyethylene bottles, preserved on ice (~4°C) and acidified with concentrated HNO_3 for lead analysis to pH < 2 and then sent to the analytical laboratory within 24 hours.

Laboratory Analysis

Laboratory analyses was conducted at the Environmental Health Sciences Laboratory, Faculty of Public Health, College of Medicine, University of Ibadan, using the methods recommended by Standard Methods for the Examination of Water and Wastewater (APHA, 2017). The total dissolved solids (TDS) were gravimetrically determined with cross checking with the calibrated TDS meter, the total hardness was determined by acid-base titrimetry with EDTA titrating to the methyl orange endpoint, total alkalinity was determined by acid-base titrimetry to the methyl orange endpoint, and lead (Pb) was determined by atomic absorption spectrophotometry after digestion in nitric acid. Each analytical batch was calibrated with certified standard and reagent blanks were run. The present laboratory campaign was restricted to these six parameters, and a wider range of ions and heavy metals (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, Cl^{-}, SO_4^{2-}, HCO_3^{-}, NO_3^{-}, Fe, Mn, Zn, Cu, Cr) was not measured and is specifically noted below as a goal for expanded future measurement. Laboratory results were entered in a structured master dataset (five sites × six parameters), checked for consistency and analysed using Python 3.12 (pandas, numpy, scipy for correlation and clustering, scikit-learn for PCA, matplotlib for plotting). All the measured values were compared with the WHO (2022) Guidelines for Drinking-Water Quality and the Nigerian Standard for Drinking Water Quality NSDWQ (2015). A detailed description of the computational procedures is given in Section 3. Due to the low number of independent sampling units (n = 5), all inferential statistics (correlation p-values, explained-variance partitioning) are reported with this limitation, and multivariate outputs are interpreted in an exploratory manner, not as confirmatory.

Computational Methodology

Descriptive and Correlation Statistics

All six parameters were calculated on a site-wise basis, together with standard deviations, range and coefficient of variation (CV), which were calculated after dividing the data set into groundwater (BH, BOREHOLE) and surface water (Stream 1–3). Pearson product-moment correlation coefficients were calculated between all parameters, and the two-tailed significance level was tested at the conventional level of $\alpha = 0.05$. With $n = 5$ (3 degrees of freedom for each pairwise test), only correlations with $|r| \gtrsim 0.88$ are considered nominally significant; correlations with smaller $|r|$ are not considered to be statistically significant but are reported descriptively as a direction of the correlation.

Principal Component Analysis (PCA)

Prior to PCA, all six variables were standardised (with mean set to zero and variance set to one) in order to eliminate bias from variables with larger absolute values (e.g., TDS) from dominating the decomposition of variance compared to those measured on different scales (e.g., lead, pH). The standardised data matrix was subjected to PCA using singular value decomposition (scikit-learn) to obtain the component loadings, explained-variance ratios, and site scores. In PCA, five samples allow for at most four non-trivial components (rank = n − 1), which is reported explicitly and not considered a free choice of the modelling, and only the first four components that account for the significant part of the variance can be interpreted.

Hierarchical Cluster Analysis (HCA)

Agglomerative HCA was performed on the standardised parameter matrix using Ward's minimum-variance linkage criterion and squared Euclidean distance, consistent with established groundwater multivariate-statistics practice (Taşan *et al.,* 2022; Arıman *et al*., 2024). The resulting dendrogram was cross-validated against the independent PCA site-score ordination to assess whether the two unsupervised methods converge on the same underlying groundwater/surface-water partition, which is a standard internal-consistency check in this literature.

Water Quality Index (WQI)

A weighted arithmetic WQI was computed following the classical formulation of Brown, McClelland, Deininger, and O'Connor as implemented in recent groundwater WQI studies (Arıman et al., 2024). For each parameter n, a unit weight $Wn = K/Sn$ was computed, where Sn is the NSDWQ (2015) permissible standard for that parameter and $K = 1/\Sigma(1/Sn)$ is a normalising constant ensuring $\Sigma Wn = 1$. A quality rating $Qn = [(Vn - V0)/(Sn - V0)] \times 100$ was computed for each parameter, where Vn is the measured concentration and V0 is the ideal value (zero for all parameters except pH, for which the pH-specific sub-index $QpH = |VpH - 7|/(Sn - 7) \times 100$ was used, with Sn = 8.5 (the upper bound of the NSDWQ permissible pH range 6.5–8.5), following the pH treatment established in the same reference methodology). This enabled an overall WQI to be calculated for each site, as $\Sigma(WnQn)/\Sigma Wn$. The five-tier classification of sites was based on the standard classification scheme: WQI < 25,

excellent; 25 – 50, good; 50 – 75, poor; 75 – 100, very poor; > 100, unsuitable for drinking.

This calculation has been done twice - once when all 5 parameters, which have NSDWQ standards, are included (pH, TDS, hardness, alkalinity, lead) and once without lead. This second calculation was not part of the analysis plan originally designed, and was added after the first WQI run was completed, but it was found that there was a methodological artefact present that should be reported explicitly, as it would be disproportionate to allow the large number of observations that are not relevant to the calculation of WQI to dominate the results, such was the case for this dataset with lead alone contributing 99.87% of the weight in the five-parameter index, which was effectively a rescaled lead sub-index, regardless of the values taken by the other four parameters. This is a well-known structural sensitivity when a few toxic parameters with low guideline values are combined with most of the parameters of the majority constituents with several orders of magnitude higher guideline values, and it is reported here as a substantive computational result, not as a nuisance result.

Recommended Extensions

Three computational extensions would materially strengthen future iterations of this analysis. First, an entropy-weighted or CCME-style WQI (Arıman *et al.,* 2024) would mitigate the trace-parameter weight-inflation problem identified in Section 3.4 by deriving weights from the empirical variance structure of the dataset rather than purely from the inverse of the regulatory standard. Second, with an expanded major-ion panel, hydrogeochemical facies diagrams (Piper trilinear, Gibbs, and Durov diagrams) would allow the PCA-resolved groundwater/surface-water axis identified in this study to be attributed to specific rock–water reaction pathways (e.g., silicate hydrolysis versus carbonate dissolution) rather than inferred indirectly from hardness–alkalinity behaviour. Third, given that the present PCA and HCA outputs are constrained by n = 5, a seasonally repeated sampling design (minimum n ≈ 15–20 site–time observations) is recommended before these multivariate outputs are treated as more than exploratory pattern recognition. We do not attempt any of these three extensions here, since doing so on the present dataset would not be supported by the available laboratory data; they are stated as concrete, falsifiable directions for subsequent work.

RESULTS AND DISCUSSION

Physicochemical Dataset

The measured physicochemical parameters for all five sampling locations, alongside WHO (2022) and NSDWQ (2015) permissible limits, are presented in Table 1.

Table 1: Physicochemical and Chemical Parameters of Sampled Water Sources Compared with WHO (2022) and NSDWQ (2015) Standards

Parameter	BH	Borehole	Stream 1	Stream 2	Stream 3	WHO (2022)	NSDWQ (2015)
pH	7.06	7.45	6.90	6.69	6.73	6.5–8.5	6.5–8.5
Temperature (°C)	29.1	28.6	29.4	28.0	28.0	—	—
TDS (mg/L)	446	339	270	264	150	500	500
Total Hardness (mg/L $CaCO_3$)	200*	170	50	37	44	500	150
Total Alkalinity (mg/L)	10	16	26	24	22	200	200
Lead, Pb (mg/L)	0.001	0.000	0.000	0.000	0.000	0.01	0.01

* Bold/asterisked value exceeds the NSDWQ (2015) permissible limit. Source: EHS Laboratory, University of Ibadan (2025); WHO (2022); NSDWQ (2015)

Descriptive Statistics

Table 2: Descriptive Statistics of Physicochemical Parameters (n = 5) and Groundwater/Surface-Water Subgroup Means

Parameter	Mean	SD	Min	Max	CV (%)	GW mean	SW mean
pH	6.97	0.31	6.69	7.45	4.4	7.26	6.77
Temperature (°C)	28.62	0.63	28.0	29.4	2.2	28.85	28.47
TDS (mg/L)	293.8	108.8	150	446	37.0	392.5	228.0
Hardness (mg/L)	100.2	78.3	37	200	78.1	185.0	43.7
Alkalinity (mg/L)	19.6	6.5	10	26	33.4	13.0	24.0
Lead (mg/L)	0.0002	0.0004	0.000	0.001	—†	0.0005	0.000

GW = groundwater subgroup (BH, BOREHOLE); SW = surface-water subgroup (Streams 1–3). † CV is not meaningfully interpretable for lead given a subgroup mean of zero and values at or near the analytical detection limit

Total hardness showed the largest relative dispersion among the non-trace parameters (CV = 78.1%), consistent with a bimodal groundwater/surface-water pattern rather than continuous variation around a single population mean. The groundwater subgroup carried appreciably higher TDS and hardness than the surface-water subgroup, while alkalinity showed the opposite pattern higher in surface water than in groundwater, a geochemically atypical relationship discussed in Section 5.

Correlation Structure

Figure 1 presents the full Pearson correlation matrix across all six parameters.

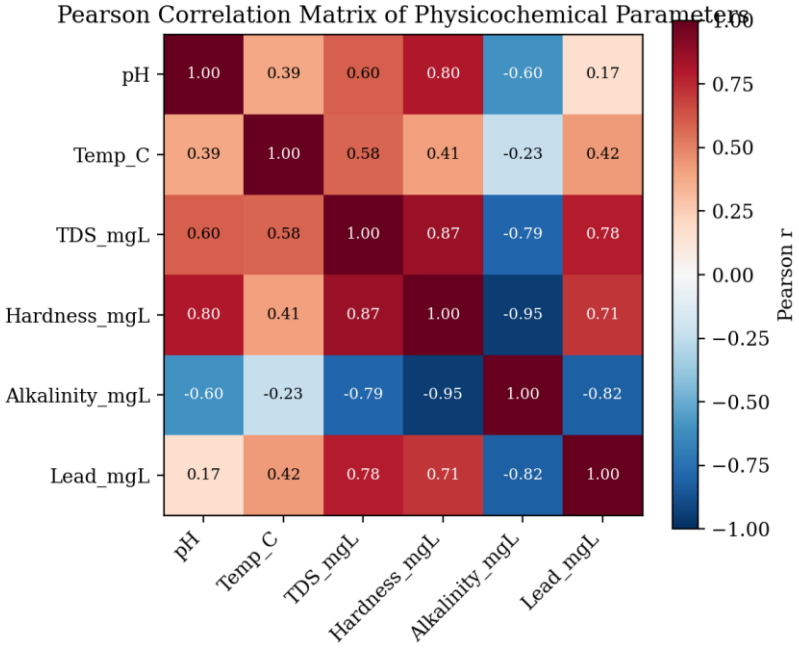

Figure 1: Pearson correlation matrix of physicochemical parameters (n = 5). Colour scale: red = positive correlation, blue = negative correlation

At n = 5, only the hardness–alkalinity pair reached conventional statistical significance ($r = -0.949$, $p = 0.014$). Strong but not formally significant associations ($0.75 < |r| < 0.90$, $p > 0.05$) were observed between TDS and hardness ($r = 0.866$, $p = 0.058$), alkalinity and lead ($r = -0.820$, $p = 0.089$), TDS and alkalinity ($r = -0.791$, $p = 0.111$), pH and hardness ($r = 0.801$, $p = 0.104$), and TDS and lead ($r = 0.782$, $p = 0.118$). These associations are reported as directional patterns motivating the mechanistic discussion in Section 5, not as confirmed statistical relationships.

Principal Component Analysis

PCA on the standardised six-parameter matrix yielded the variance decomposition and loadings summarised in Table 3.

Table 3: PCA Explained Variance and Component Loadings

Parameter	PC1 loading	PC2 loading	PC3 loading
pH	0.350	0.650	−0.378
Temperature	0.277	0.410	0.786
TDS	0.461	−0.030	0.152
Hardness	0.478	0.049	−0.241
Alkalinity	−0.450	0.272	0.308
Lead	0.396	−0.576	0.251
% Variance explained	69.1	14.6	14.1
Cumulative %	69.1	83.7	97.8

PC1 alone accounts for 69.1% of total variance and loads positively on TDS, hardness, lead, pH, and temperature, and negatively on alkalinity. This loading pattern identifies PC1 as a general mineralisation axis, separating highly mineralised, near-neutral, trace-lead-bearing groundwater from more dilute, mildly acidic, higher-alkalinity surface water. Figure 2 shows the PCA score–loading biplot.

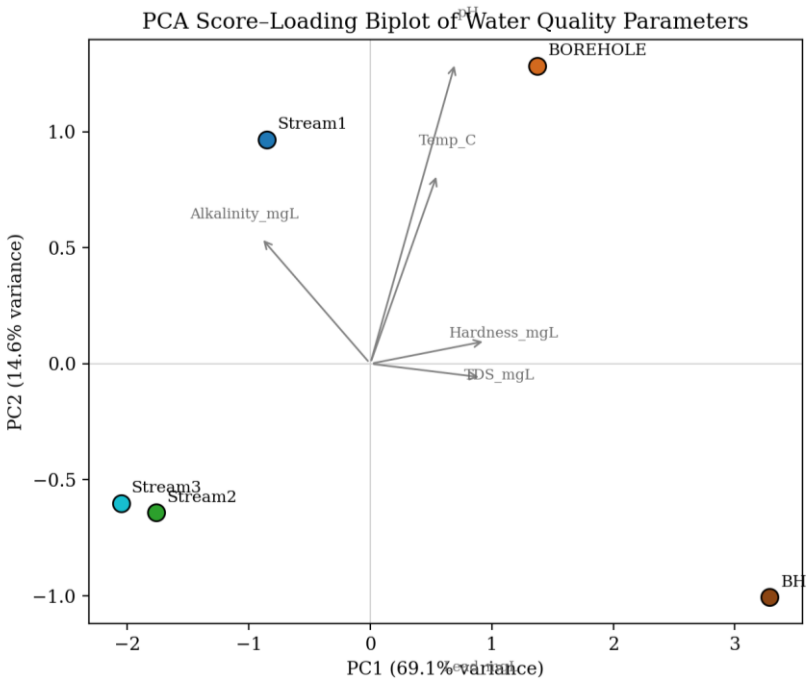

Figure 2: PCA score–loading biplot (PC1 vs PC2). Sites are coloured by source type; loading vectors indicate each parameter's contribution to the two leading components

Site scores along PC1 ordered the samples as BH (3.29) > BOREHOLE (1.38) > Stream 1 (−0.85) > Stream 2 (−1.76) > Stream 3 (−2.05), reproducing exactly the groundwater/surface-water dichotomy and, within each group, the relative mineralisation ranking already evident in the raw TDS and hardness data (Table 1). Given the rank-4 ceiling imposed by n = 5, PC1–PC3 together account for 97.8% of variance, leaving essentially no residual structure for higher components to resolve; PC4 and PC5 are not interpreted.

Hierarchical Cluster Analysis(HCA)

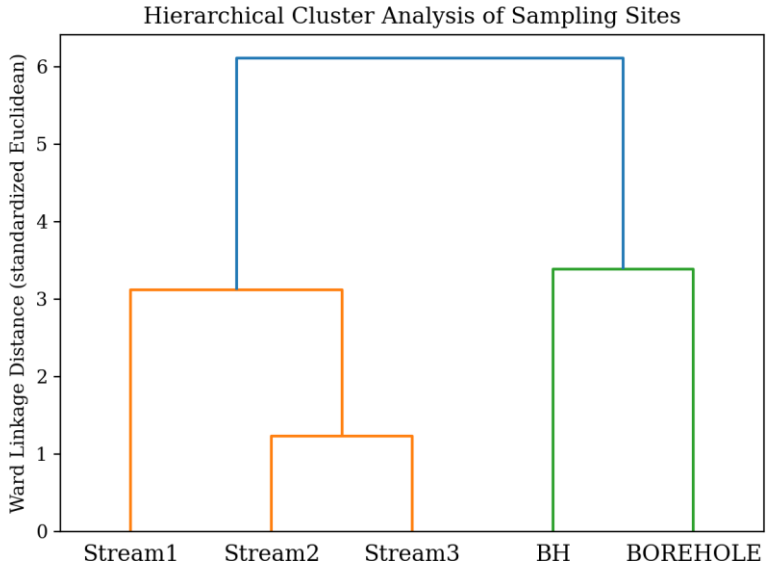

Figure 3: Ward-linkage dendrogram of the five sampling sites based on standardised physicochemical parameters (squared Euclidean distance)

HCA reproduced the same two-cluster partition identified by PCA: Streams 2 and 3 merged first (linkage distance 1.23), joined subsequently by Stream 1 (linkage distance 3.11) to form a surface-water cluster, while BH and BOREHOLE merged independently as a groundwater cluster (linkage distance 3.39); the two clusters joined only at the final, highest linkage distance (6.11). The convergence of two independent unsupervised methods (PCA ordination and Ward-linkage HCA) on an identical two-group partition is a standard internal-consistency check in the multivariate groundwater literature (Taşan *et al.,* 2022) and strengthens confidence that the groundwater/surface-water distinction, rather than any other grouping (e.g., by geographic sector), is the dominant structure in this dataset notwithstanding the caution warranted by n = 5.

Water Quality Index

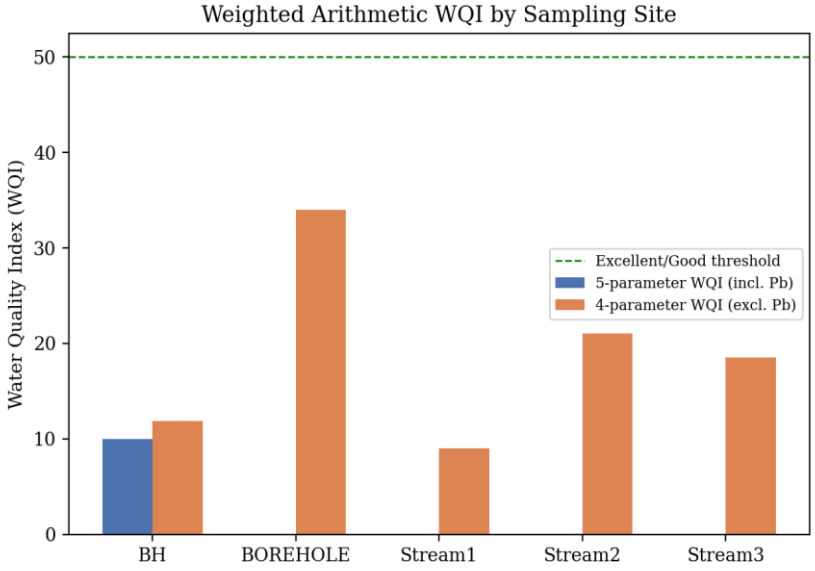

Figure 4: Weighted arithmetic WQI by sampling site, computed with and without lead in the parameter set

Table 4: Weighted Arithmetic WQI Results

Site	WQI (5-param, incl. Pb)	Class	WQI (4-param, excl. Pb)	Class
BH	10.00	Excellent	11.90	Excellent
BOREHOLE	0.04	Excellent	33.97	Good
Stream 1	0.01	Excellent	8.98	Excellent
Stream 2	0.03	Excellent	21.03	Excellent
Stream 3	0.02	Excellent	18.49	Excellent

In the five-parameter WQI, lead's unit weight (Wn = K/Sn = 0.9987, i.e., 99.87% of total weight) rendered the composite index almost numerically identical to the rescaled lead sub-index; BH, the only site with detectable lead, therefore returned the highest WQI (10.00) despite also being the only site exceeding a national hardness standard, while all lead-free sites collapsed to near-zero WQI values regardless of their pH, TDS, hardness, or alkalinity standing. The four-parameter WQI (excluding lead) also assigns very little weight to lead, instead focusing mainly on pH (89.6%) and the same pattern holds true as the permissible standard for pH is also very small compared to the other parameters, which again focuses the discriminating power onto a single variable. Despite this, all sites were still rated as excellent-to-good water quality, and as shown in Table 1, the parameter-by-parameter compliance picture would not differ significantly from that presented here. However, it is important to recognize that the WQI values shown here reflect this weighting sensitivity and not a composite signal from all measured constituents. This is covered further in Section 5.6.

Discussion

pH: Hydrogen-Bond Network Equilibria and Proton Transfer

Phenomenological pH values listed here (6.69–7.45) are in a narrow range around neutrality, but the pH of neutrality itself is temperature- and pressure-dependent: the autoionisation constant of water Kw is a consequence of the cooperative hydrogen-bond network, and excess protons are transported not by physical diffusion of a bare

H^+ ion, but by a relay of hydrogen-bond rearrangements along a chain of water molecules (Agmon, 2012). This mechanism is abnormally rapid relative to normal ionic diffusion and is extremely susceptible to the degree of hydrogen-bond order found in the water sample, and therefore to any process which affects the degree of order in the hydrogen-bond network that dissolves the ions, the CO_2, or the temperature gradient in the water sample, which in turn affects the effective proton mobility and hence the equilibrium pH towards which the sample will tend to evolve.

The observed pattern (borehole waters: 7.06 – 7.45, trending towards neutral to mild alkaline; stream waters: 6.69 – 6.90, trending mildly acidic) is consistent with two different proton-consuming and proton-donating regimes. The percolating water in the boreholes thus has a residence time in which the hydrolysis of silicate and carbonate minerals consume H+ ions from the borehole water, causing the water pH to increase during the aquifer's residence time. Dissolved atmospheric CO_2 quickly reacts with the shallow, well-mixed surface water to create carbonic acid, and organic acid inputs from the surrounding woodland and farmland catchment further remove protons to make the pH in the streams more acidic than neutral. This picture is consistent with the moderate positive correlation between pH and hardness, 0.801, with n = 5, which is not formally significant, but is consistently positive and in the direction of the picture, as sites with higher hardness would be the sites where there is more mineral hydrolysis, which also means the consumption of more H+ ions.

Temperature: Thermal Equilibration and Reaction Kinetics

Water temperatures (28.0–29.4 °C) were uniformly high, reflecting the tropical setting, with surface water showing marginally greater variability (28.0–29.4 °C) than groundwater (28.6–29.1 °C) a pattern consistent with the much smaller thermal mass and direct radiative coupling of shallow stream channels compared with the thermally buffered subsurface. From a statistical physics standpoint, temperature is the parameter that most directly modulates every other measured quantity, since it sets the Boltzmann-weighted rate of all dissolution, exchange, and diffusion processes contributing to the other five parameters. The Arrhenius-type acceleration of silicate and carbonate dissolution kinetics at these consistently warm temperatures is one contributing factor alongside residence time to the elevated mineralisation of the borehole sources; it also reduces dissolved-oxygen solubility (via the temperature dependence of Henry's law constants) and increases microbial metabolic rates, both of which have implications for water potability that are independent of the physicochemical parameters measured directly in this study.

Total Dissolved Solids: Diffusion-Limited Mineral Dissolution

TDS was measured as 150 mg/L (Stream 3) to 446 mg/L (BH), all of which were within WHO (2022) and NSDWQ (2015) limits and had a clearly defined groundwater/surface-water split (392.5 mg/L mean, BH compared to 228.0 mg/L mean, Stream 3). Surface reaction and diffusive transport steps are sequential reactions that occur at a solid–water interface, and for the common minerals of basement complex lithologies (silicate and aluminosilicate) the diffusion of reaction products away from the solid–water interface is generally the rate-limiting step with dissolved ion accumulation being proportional to about the square root of contact (residence) time under diffusion-limited conditions. Surface water in fast, turbulent, dilution-dominated stream flow, then, will have a shorter residence time in the rock, meaning that it will accumulate much less dissolved mass from that water–rock interaction than will groundwater slowly percolating through fractured basement rock which is the same physical asymmetry that accounts for the strong correlation between TDS and hardness throughout the dataset (r = 0.866) and that both are downstream products of the same water–rock interaction, but in different flow regimes.

Total Hardness: Ion Hydration Thermodynamics, and the Non-Carbonate Signature

The relative spread (CV) for total hardness (37-200 mg/L) was the greatest of any of the non-trace parameters in the data set (78.1%) and had the only exceedance identified in the data set for BH (200 mg/L) against the NSDWQ (2015) limit of 150 mg/L. Hardness is related to the total concentration of divalent cations, mainly Ca^{2+} and Mg2+, which behave in water in a completely different way, because they have a high charge-to-size ratio, and form a very stable and strongly bound hydration shell of oriented water dipoles around the divalent cation. This strong hydration free energy is precisely what allows Ca^{2+} and Mg^{2+} to be mobilised into solution from mineral lattices and to remain kinetically stable in the dissolved phase rather than re-precipitating, which is the physical basis for hardness accumulating cumulatively along a groundwater flow path in the way that TDS does.

The strong negative correlation between hardness and alkalinity ($r = -0.949$, $p = 0.014$ the only statistically significant relationship recovered in this dataset) is the most physically informative single result of the correlation analysis. Under classical hydrogeochemical convention, hardness and alkalinity in groundwater from carbonate-buffered aquifers typically co-vary positively, because Ca^{2+} and Mg^{2+} are frequently released to solution together with HCO_3^- during carbonate mineral dissolution ($CaCO_3 + CO_2 + H_2O \rightarrow Ca^{2+} + 2HCO_3^-$). The inverse relationship observed here BH exhibiting the

highest hardness (200 mg/L) but the lowest alkalinity (10 mg/L) of all five sites is inconsistent with a purely carbonate-hardness mechanism and instead points toward a non-carbonate hardness signature, in which the divalent cations at BH are more plausibly balanced by sulphate or chloride anions than by bicarbonate. This is a testable, falsifiable hypothesis rather than a settled conclusion: it can only be confirmed with the sulphate, chloride, calcium, and magnesium data that were not part of the present six-parameter laboratory panel, and we flag it here explicitly as the single most actionable target for extended analytical work at this specific site.

Total Alkalinity: Buffering Capacity and Carbonate Equilibria

Alkalinity (10–26 mg/L) was uniformly low across all sites relative to the 200 mg/L regulatory ceiling, indicating limited buffering capacity against acid inputs at every sampled location a condition of some concern given the mild acidity already observed in the surface-water sites. Alkalinity is a direct measure of the solution's ability to buffer incoming protons through the carbonate-bicarbonate-CO_2 equilibrium system, which is also controlled by the same hydrogen-bonded solvent network that is controlled by the carbonate-bicarbonate-CO_2 equilibrium system. The higher alkalinity of the surface water (mean 24.0 mg/L) than the groundwater (mean 13.0 mg/L) further supports the non-carbonate interpretation of the surface water's hardness discussed above: if the surface water was mainly carbonate derived, its alkalinity would be expected to be higher still than the groundwater's.

Lead: Trace-Ion Electrostatics, Substitution Chemistry, and a Weighting Artefact

Lead was detected only at BH (0.001 mg/L), an order of magnitude below the WHO (2022) and NSDWQ (2015) threshold of 0.01 mg/L. Even at trace concentration, Pb^{2+}'s behaviour in aqueous and biological systems is governed by condensed matter electrostatics: its ionic radius and charge density are sufficiently close to those of Ca^{2+} that it can substitute for calcium at protein and mineral binding sites, a substitution mechanism that underlies both its environmental mobility in calcium-bearing mineral phases and its neurotoxicity, since Pb^{2+} can cross the blood–brain barrier and disrupt calcium-dependent signalling once inside biological tissue (WHO, 2025). Given that lead is now understood to have no toxicologically safe threshold of exposure, particularly for children (WHO, 2025), the co-occurrence of detectable lead and non-carbonate hardness specifically at BH is a pattern worth monitoring even though both values individually fall within regulatory limits.

The computational finding reported in Section 4.6 that lead's very low permissible standard causes it to dominate the weighted arithmetic WQI almost completely is itself a condensed-matter-adjacent observation about how trace-constituent electrostatics interact with index construction: because Pb^{2+} is toxicologically potent at concentrations many orders of magnitude below those of the majority ionic constituents, any WQI formulation that weights parameters purely by the inverse of their regulatory standard will structurally amplify trace toxicant behaviour at the expense of majority-constituent information, regardless of whether that amplification reflects the analyst's intent. This is not a flaw specific to this dataset; it is a general property of the classical weighted arithmetic WQI whenever a trace parameter's standard differs from the majority parameters' standards by several orders of magnitude, and it argues for entropy-weighted or CCME-style alternatives (Section 3.5) whenever heavy metals are included alongside majority-ion parameters in future WQI work on this or comparable datasets.

Groundwater–Surface Water Geochemical Contrast: A Collective-Behaviour Interpretation

Taken together, the PCA and HCA results (Sections 4.4–4.5) indicate that a single latent axis essentially a residence-time-driven mineralisation gradient accounts for 69.1% of the total variance in this six-parameter dataset and cleanly separates the two water source types. This is consistent with a collective-behaviour interpretation in which the ensemble of dissolved species at each site is not a set of independently varying quantities but a jointly determined outcome of a single controlling physical process: the duration and intensity of water–rock contact experienced by that parcel of water before sampling. Framed this way, the groundwater/surface-water contrast in this basement complex setting is analogous to an order parameter distinguishing two macroscopically different states of the same underlying system short-residence-time, dilute, mildly acidic, alkalinity-buffered surface flow versus long-residence-time, mineralised, near-neutral, non-carbonate-hardness-bearing groundwater with the five sampled sites falling cleanly either side of that boundary rather than distributing continuously between the two end-members. This pattern is broadly consistent with hydrogeochemical contrasts reported for other basement complex settings in Nigeria (Musa et al., 2024; Adetona *et al.,* 2024; Omeka et al., 2022), lending external plausibility to the internal statistical structure recovered here, even though the present dataset alone (n = 5) cannot establish this as a general result for the wider study area.

Comparison with Recent Literature and Points of Disagreement

The groundwater-mineralisation-exceeds-surface-water pattern reported here is consistent with hydrogeochemical assessments elsewhere in Nigeria's

basement complex provinces, including the Pago area of north-central Nigeria (Musa *et al.,* 2024) and agrarian basement terrains studied using combined graphical, indexical, and statistical modelling approaches (Omeka *et al.,* 2022). Where this study diverges from the broader multivariate groundwater literature is in the strength and completeness of the multivariate signal that a wider parameter panel typically resolves: PCA and HCA studies conducted with ten or more ionic parameters and dozens of sampling wells (e.g., Arıman *et al.,* 2024; Taşan *et al.,* 2022) typically resolve two to three interpretable components tied to distinguishable geochemical processes (e.g., silicate weathering versus evaporite dissolution versus anthropogenic input), whereas the present six-parameter, five-site dataset collapses almost the entire variance structure onto a single component. This is not necessarily evidence that only one process is operating in the Federal University Lokoja study area; it more likely reflects the resolving power available from six parameters and five sites, and should not be over-interpreted as evidence of geochemical simplicity relative to better-instrumented comparison sites.

Environmental and Public Health Significance

From a public health standpoint, the physicochemical results indicate that both boreholes, while requiring hardness-related treatment (particularly BH), are of substantially better raw-water physicochemical quality than the three surface-water sources, which additionally carry the generic microbiological and turbidity risks characteristic of open stream channels in a mixed land-use catchment risks that fall outside the physicochemical parameter set analysed here but are well documented in the regional literature (Agbasi *et al.,* 2023). The detection of trace lead specifically at BH, even though compliant with current standards, is consistent with international guidance that lead exposure should be minimised toward the lowest achievable level given the absence of a safe threshold (WHO, 2025), and supports the borehole-specific infrastructure inspection (pump casing, fittings) recommended in the underlying field study.

CONCLUSION

This study re-examined a compact physicochemical dataset from a basement complex terrain in north-central Nigeria through a combined computational and condensed matter physics framework. Multivariate analysis generated a single dominant latent axis (69.1% of variance) separating mineralised groundwater from dilute surface water, which was also reproduced by a hierarchical cluster analysis, and a statistically significant (-0.949, $p = 0.014$), geochemically informative anti-correlation between the hardness and alkalinity components, which points towards a non-carbonate hardness signature at the most mineralised borehole. Mechanistic interpretations, and not just descriptions, were offered for each of the measured parameters via concepts from condensed matter physics and statistical physics: the hydrogen-bond network proton transfer, Debye–Hückel electrostatics, ion hydration thermodynamics, diffusion-limited weathering kinetics, and trace-ion substitution chemistry. A methodological result of significant interest came out of the calculations of the water quality index: the classic WQI based on weighted arithmetic indexes is structurally dominated by the parameter with the lowest regulatory standard, which is likely to be the case in any future set of data that combines trace heavy metals with the majority ionic ones, and which is the reason for the use of entropy-weighted alternatives in water quality index calculations in the future.

The conclusions presented are made with the clear acknowledgement that they are based on five sampling locations and six parameters sampled from one sampling campaign in the sense that the resulting pattern recognitions are exploratory, formulating hypotheses and providing mechanistic insights, but cannot be regarded as complete in lieu of the added parameters of the expanded major-ion panel, the added temporal dimension of seasonal replication, and the added microbiological detail needed in order to comprehensively certify the water sources characterised here as potable for domestic use. This is primarily a methodological and interpretive contribution, since it shows that a physically-based, computational explicit re-analysis can generate mechanistic hypotheses and reveal true statistical artifacts in common water quality indices that would not have been uncovered by a standards-based approach to the data.

REFERENCES

Adetona, A. A., Aliyu, S. B., Kwaghhua, F. I., & Damidami, L. M. (2024). Groundwater investigation within the basement complex, North Central Nigeria, using magnetic and resistivity method. Arabian Journal of Geosciences, 17, Article 6. https://doi.org/10.1007/s12517-023-11784-5

Agaku, R. M., Ikyumbur, J. T., Amanyi, M. I., & Onwoke, E. E. (2025). Assessment of heavy metals concentration in ground water in Gboko, Nigeria. Nigerian Journal of Physics, 34(1), 122–133. https://doi.org/10.62292/njp.v34i1.2025.390

Agbasi, J. C., Egbueri, J. C., Ayejoto, D. A., Unigwe, C. O., Omeka, M. E., Nwazelibe, V. E., Ighalo, J. O., Pande, C. B., & Fakoya, A. A. (2023). The impact of seasonal changes on the trends of physicochemical, heavy metal and microbial loads in water resources of Southeastern Nigeria: A critical review. In J. C. Egbueri et al. (Eds.),

Climate change impacts on Nigeria. Springer. https://doi.org/10.1007/978-3-031-21007-5_25

Agmon, N. (2012). Liquid water: From symmetry distortions to diffusive motion. Accounts of Chemical Research, 45(1), 63–73. https://doi.org/10.1021/ar2002186

American Public Health Association (APHA). (2017). Standard methods for the examination of water and wastewater (23rd ed.). APHA, AWWA, WEF.

Arıman, S., Soydan-Oksal, N. G., Beden, N., & Ahmadzai, H. (2024). Assessment of groundwater quality through hydrochemistry using principal components analysis (PCA) and water quality index (WQI) in Kızılırmak Delta, Turkey. Water, 16(11), Article 1570. https://doi.org/10.3390/w16111570

Debenedetti, P. G., & Stanley, H. E. (2003). Supercooled and glassy water. Physics Today, 56(6), 40–46. https://doi.org/10.1063/1.1595053

Freeze, R. A., & Cherry, J. A. (1979). Groundwater. Prentice-Hall.

Hem, J. D. (1985). Study and interpretation of the chemical characteristics of natural water (3rd ed.). U.S. Geological Survey Water-Supply Paper 2254.

Musa, A., Tsado, F., Amadi, A. N., Aweda, A. K., Habib, I. A., & Abdulkadir, H. (2024). Hydrogeochemical investigation of surface and groundwater quality of Pago, North Central Nigeria. FUDMA Journal of Sciences, 8(6), 100–106. https://doi.org/10.33003/fjs-2024-0806-2812

Omeka, M. E., Egbueri, J. C., & Unigwe, C. O. (2022). Investigating the hydrogeochemistry, corrosivity and scaling tendencies of groundwater in an agrarian area (Nigeria) using graphical, indexical and statistical modelling. Arabian Journal of Geosciences. https://doi.org/10.1007/s12517-022-10514-7

Standards Organisation of Nigeria. (2015). Nigerian Standard for Drinking Water Quality (NSDWQ 2015). Standards Organisation of Nigeria.

Taşan, M., Demir, Y., & Taşan, S. (2022). Groundwater quality assessment using principal component analysis and hierarchical cluster analysis in Alaçam, Turkey. Water Supply, 22(3), 3431–3447. https://doi.org/10.2166/ws.2021.390

World Health Organization. (2022). Guidelines for drinking-water quality (4th ed., incorporating the first addendum). World Health Organization.

World Health Organization. (2025). Lead poisoning and health [Fact sheet]. Retrieved July 2, 2026, from https://www.who.int/news-room/fact-sheets/detail/lead-poisoning-and-health