Nigerian Journal of Environmental Sciences and Technology
(NIJEST)
www.nijest.com
ISSN (Print): 2734-259X | ISSN (electronic): 2734-2603
Vol 9, No. 2 October 2025, pp 149 - 163

Principal Component Analysis of Hydrochemical Parameters in Hydrocarbon-Affected Groundwater at Otorogu Delta State, Nigeria

Omena, A.L[1]., Ihimekpen, N.I[1]., and Okonofua, S.E[1]

[1]Department of Civil Engineering, Faculty of Engineering, University of Benin, Benin City, Nigeria

*Corresponding Author ome1602@yahoo.com

https://doi.org/10.36263/nijest.2025.02.63

ABSTRACT

Groundwater contamination arising from hydrocarbon exploitation and related human activities has become a major environmental concern in Otorogu, Delta State. Understanding the key hydrochemical factors that influence groundwater quality is essential for sustainable water resource management. This study therefore aimed to identify the dominant hydrochemical and trace-metal factors influencing groundwater quality and to assess the extent of hydrocarbon-related pollution in the area using Principal Component Analysis (PCA). A total of 100 groundwater samples were collected from domestic boreholes; 50 during the wet season and 50 during the dry season. Thirty-one (31) physico-chemical and microbiological parameters were analyzed following standard laboratory protocols, and all measurements were performed in triplicate to ensure data reliability. Prior to PCA, data were standardized (z-score transformation) to minimize scale bias, and sample adequacy was confirmed using the Kaiser-Meyer-Olkin (KMO) test (0.849) and Bartlett's test of sphericity ($p < 0.001$). PCA extraction was based on eigenvalues greater than one and varimax rotation to enhance component interpretability. The PCA results showed that the first principal component (PC1), explaining 36.3% of the total variance, was dominated by ionic and nutrient variables (NO_3^-, Cl^-, TDS, EC), indicating influences from salinization and agricultural runoff. The second and third components were characterized by heavy metals (Pb, Cu, Ni, Zn) and microbial indicators, respectively, reflecting anthropogenic and localized hydrocarbon impacts. Overall, the study revealed that groundwater chemistry in Otorogu is primarily controlled by anthropogenic inputs, while hydrocarbon influence is moderate but spatially localized. Continuous seasonal monitoring and effective pollution control measures are therefore recommended to safeguard groundwater quality and ensure aquifer sustainability in hydrocarbon-impacted environments.

Keywords: Groundwater quality, Principal Component Analysis, Hydrocarbon pollution, Heavy metals, Niger Delta, Multivariate analysis.

1.0. Introduction

Groundwater remains an indispensable source of potable, agricultural, and industrial water; however, in hydrocarbon-producing regions such as the Niger Delta, its quality is increasingly jeopardized by petroleum-related contamination (Oyebamiji et al., 2024; Ikenna and Chukwudozie, 2024). Hydrocarbon leakage from exploration, refining, and distribution activities introduces complex organic compounds and trace metals into aquifers, resulting in significant alteration of physico-chemical characteristics such as pH, total dissolved solids (TDS), and electrical conductivity (EC) (Eze and Chinemelu, 2025; Osisanya et al., 2024). These transformations compromise the potability and ecological safety of groundwater and have been linked to chronic health issues and ecosystem degradation (Sahoo et al., 2015; Singh et al., 2005; Adebayo et al., 2023).

Traditional assessment techniques such as the Water Quality Index (WQI) provide a useful overview of groundwater suitability for use (Sutadian et al., 2017; Sunardi et al., 2020). However, WQI has limited diagnostic capacity, it does not distinguish which parameters or contaminants drive deterioration, nor does it identify the underlying hydrochemical processes responsible for observed changes (Thurston et al., 2011; Simeonov et al., 2003). This shortcoming restricts effective management interventions in hydrocarbon-impacted regions, where remediation must target specific pollutants and sources.

To address this analytical gap, multivariate statistical techniques, particularly Principal Component Analysis (PCA), have gained prominence in hydrogeochemical research. PCA simplifies large, inter-correlated datasets into a smaller number of uncorrelated components that explain the dominant variance in water quality data (Shrestha and Kazama, 2007; Shrestha, 2021). It is particularly useful in identifying pollutant linkages, hydrochemical controls, and spatial-temporal variations in contaminated aquifers (Sousa et al., 2007; Simeonov et al., 2004; Arıman et al., 2024). Compared to clustering or discriminant methods, PCA is data-driven, assumption-free, and offers a robust framework for uncovering key hydrochemical processes (Tabachnick and Fidell, 2001; Rashid et al., 2012). Recent applications in hydrocarbon-bearing terrains have shown that PCA effectively isolates dominant factors such as salinity, heavy metal enrichment, and organic load as indicators of petroleum-related groundwater contamination (Osisanya et al., 2024; Eze and Chinemelu, 2025).

Despite these advances, few studies in the Niger Delta have integrated PCA with hydrochemical analysis to quantify and interpret the dominant factors controlling groundwater quality under sustained hydrocarbon pressure. Existing works are often limited to descriptive or index-based evaluations without statistically determining the interrelationships among key contaminants. This gap underscores the need for a multivariate analytical approach capable of identifying and ranking the principal parameters influencing groundwater degradation in the region.

Therefore, this study applies Principal Component Analysis (PCA) to determine the dominant hydrochemical processes and key contaminants influencing groundwater quality in hydrocarbon-impacted areas of the Niger Delta. Specifically, the study aims to: characterize the spatial variability of major physico-chemical parameters, identify the principal components explaining groundwater quality variation; and interpret the hydrochemical processes and pollution sources controlling groundwater composition. This integration of PCA with hydrochemical datasets provides a robust framework for groundwater quality diagnosis, pollution source identification, and the development of sustainable management strategies for petroleum-impacted aquifers.

2.0 Materials and Methods

2.1 Description of Study Area

The study was conducted in Otorogu, located within Udu Local Government Area, Delta State, Nigeria. Geographically, the area lies approximately between latitude 5°31′N and 5°37′N and longitude 5°48′E and 5°54′E, about 12 km from Warri city. The area falls within the Niger Delta sedimentary basin, which is characterized by extensive hydrocarbon exploration and production activities.

Geologically, Otorogu is underlain by the Benin Formation, consisting predominantly of unconsolidated sands and gravels with minor clay lenses, which form the major aquifer. The underlying Agbada Formation (interbedded sand and shale) and the deeper, more impermeable Akata Formation act as confining layers. The sediments are generally highly permeable, favoring groundwater recharge and flow, although clay or shale interbeds introduce heterogeneity (DFC, 2018).

Hydrogeologically, the aquifer is largely unconfined to semi-confined, shallow, and highly susceptible to contamination from surface activities, including petroleum-related pollution. Groundwater flow generally follows a north-to-south gradient, but local topography, human activities, and subsurface heterogeneities can affect flow directions (SCIRP, 2020). The region receives an annual rainfall of approximately 1,900 mm, which supports aquifer recharge but also facilitates contaminant leaching.

The study area hosts significant oil and gas infrastructure, including the Otorogu Gas Plant and associated pipelines, increasing the risk of hydrocarbon releases. The terrain is low-lying, interspersed with swampy areas, creeks, and farmlands, with sandy soils that promote pollutant infiltration. Vegetation ranges from wet forest to mangrove transition zones, with swampy areas near creeks. The 3D study area map is presented in Figure 1.

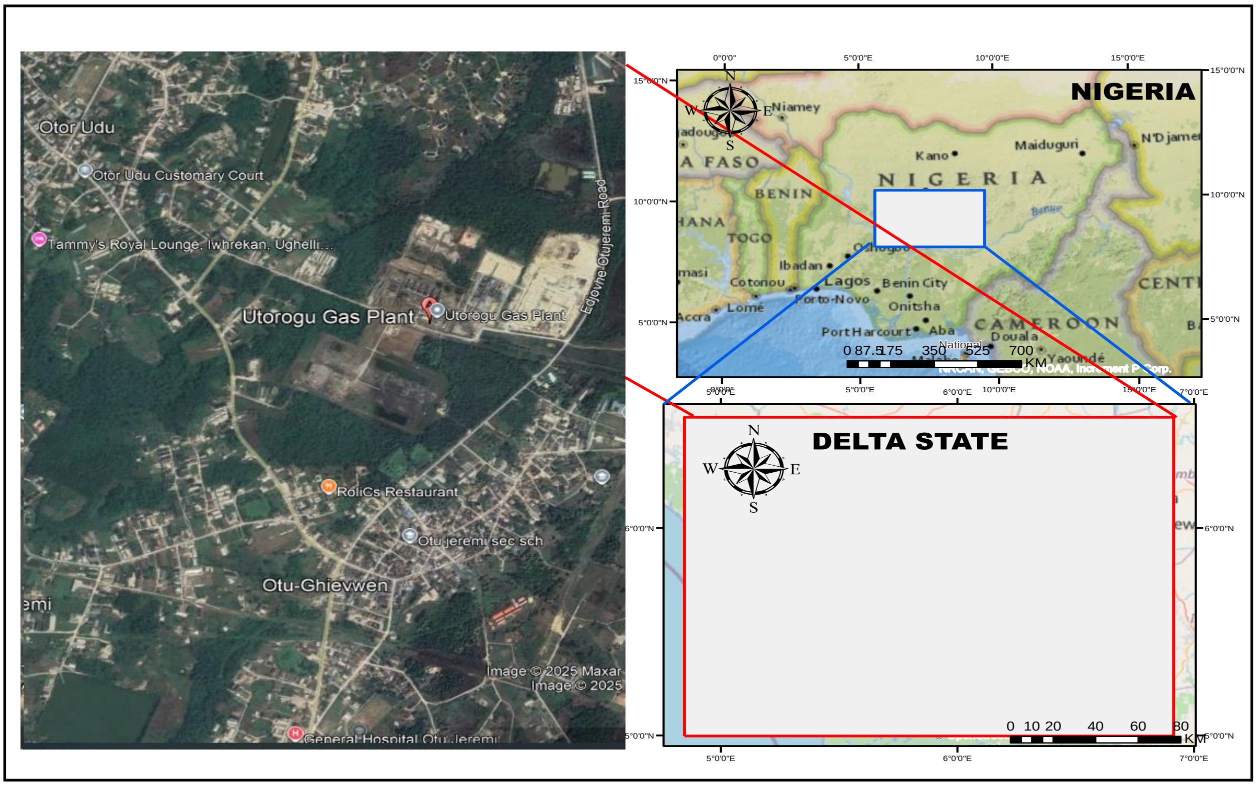

Figure 1: 3D Study Area Map

2.2 Groundwater Sampling Design and Data Acquisition

A total of 100 groundwater samples were collected from 100 domestic boreholes, with 50 samples collected during the dry season and 50 during the wet season, ensuring seasonal variation in groundwater quality was captured. Each borehole was sampled once. Borehole locations were georeferenced using a handheld GPS (accuracy ±3 m), and coordinates were recorded for spatial analysis. Sampling points were selected to cover areas with varying proximity to industrial and residential zones.

Thirty-one (31) physico-chemical and microbiological parameters were analyzed for each sample. Physico-chemical parameters included temperature, odor, color/clarity, pH, electrical conductivity (EC), turbidity, total suspended solids (TSS), salinity, alkalinity, total dissolved solids (TDS), dissolved oxygen (DO), total hydrocarbon content (THC), biochemical oxygen demand (BOD), chemical oxygen demand (COD), bicarbonate (HCO_3^-), sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), magnesium (Mg^{2+}), chloride (Cl^-), phosphorus (P), ammonium (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-), sulfate (SO_4^{2-}), and heavy metals including iron (Fe), manganese (Mn), zinc (Zn), copper (Cu), chromium (Cr), cadmium (Cd), nickel (Ni), and lead (Pb). Microbiological parameter included total coliform counts (TCC).

2.2.1 Laboratory Analysis and QA/QC

All samples were collected in pre-cleaned polyethylene bottles and preserved according to APHA (2017) guidelines. Physico-chemical analyses were conducted using calibrated instruments: pH and EC meters, spectrophotometers for trace metals, and standard titration procedures for alkalinity and hardness. Total hydrocarbon content was measured using UV-visible spectrophotometry. Microbiological analyses were performed using membrane filtration and culture techniques.
Quality assurance and control (QA/QC) procedures included:

1. Calibration of all instruments before analysis using standard solutions.
2. Analysis of method blanks and laboratory duplicates.
3. Use of certified reference materials for trace metal determination.
4. Detection limits and uncertainty ranges were maintained according to standard laboratory protocols.

2.3 Data Pre-Processing

Prior to statistical analysis, all continuous variables were checked for missing values and outliers. Missing values (<2% of dataset) were replaced using the variable mean, and extreme outliers (>3 standard deviations from the mean) were examined and winsorized based on contextual justification. To eliminate scale effects among variables, data were standardized using z-score transformation presented in equation 1:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \tag{1}$$

Where; x_{ij} is the observed value, $\bar{x}_j$ the mean, and s_j the standard deviation of variable j.

2.4 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) was employed to reduce the dimensionality of the water quality dataset and to identify the dominant factors controlling groundwater chemistry (Jolliffe and Cadima, 2016). Prior to PCA, data suitability was assessed using the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity. A KMO value ≥0.7 indicates that the sample is adequate for factor analysis, while a significant Bartlett's test ($p < 0.05$) confirms that sufficient correlations exist among variables to justify PCA.

PCA was conducted on the standardized correlation matrix because the variables were measured in different units. Eigenvalues (λ_i) and eigenvectors (v_i) were extracted from the matrix according to equation (2):

$$Rv_i = \lambda_i v_i \tag{2}$$

Components with eigenvalues greater than one (Kaiser Criterion) and those appearing before the inflection point on the screen plot were retained. The total variance explained by each component was computed using Equation (3):

$$Var_i = \frac{\lambda_i}{\sum \lambda_i} \times 100 \tag{3}$$

To further validate component retention, parallel analysis was performed. The retained components were subjected to Varimax (orthogonal) rotation to enhance interpretability by maximizing the variance of squared loadings and minimizing cross-loadings among factors. Varimax rotation was selected because the underlying hydrochemical processes were assumed to act largely independently, consistent with groundwater systems dominated by distinct geochemical controls. Variables with loadings ≥0.70 were regarded as significant contributors to the principal components (Jolliffe, 2002; Hair et al., 2010). Communalities were examined to determine the proportion of variance in each variable explained by the retained components. The derived factor loadings were then interpreted to infer hydrochemical processes influencing groundwater quality (Field, 2013).

All statistical analyses were conducted using IBM SPSS Statistics 25. Although the sample-to-variable ratio (N/p = 100/31 = 3.2) was slightly below the ideal range recommended for highly stable PCA solutions, the high KMO value (0.849) and acceptable communalities confirmed that the dataset was adequate and the extracted components were robust (Field, 2013; Jolliffe and Cadima, 2016).

3.0 Results and Discussion

3.1 Suitability of PCA for Groundwater Data

Preliminary diagnostics presented in Table 1 confirmed the suitability of the dataset for Principal Component Analysis (PCA).

Table 1: KMO and Bartlett's Test for Dry and Wet Season

Test		Dry Season	Wet Season
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.849	0.769
Bartlett's Test of Sphericity	Approx. Chi-Square	3.645E3	3.334E3
	df	465	465
	Sig.	.000	.000

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy yielded values of 0.849 for the dry season and 0.769 for the wet season, both exceeding the recommended threshold of 0.6, indicating that correlations among variables were sufficiently compact for PCA (Kaiser, 1974; Yong and Pearce, 2013). Bartlett's Test of Sphericity was significant in both seasons ($p < 0.001$), confirming that the correlation matrices were not identity matrices and that the variables were adequately interrelated for factor extraction (Hair et al., 2019). Communalities were examined to determine how well each variable was represented by the extracted components as observed in Table 2.

Table 2: Estimated Communalities for Dry and Wet Season

S/N	Variables	Dry Season		Wet Season	
		Initial	Extraction	Initial	Extraction
1	pH	1.000	.509	1.000	.647
2	NO_3	1.000	.989	1.000	.972
3	EC	1.000	.974	1.000	.943
4	Turb.	1.000	.942	1.000	.443
5	DO	1.000	.793	1.000	.907
6	TDS	1.000	.974	1.000	.898
7	Na	1.000	.941	1.000	.924
8	Pb	1.000	.901	1.000	.939
9	SO_4	1.000	.986	1.000	.971
10	Zn	1.000	.917	1.000	.991
11	Cu	1.000	.947	1.000	.976
12	Cl	1.000	.987	1.000	.986
13	Fe	1.000	.914	1.000	.991
14	HCO_3	1.000	.978	1.000	.972
15	TSS	1.000	.712	1.000	.616
16	NO_2	1.000	.988	1.000	.968
17	Cd	1.000	.943	1.000	.903
18	Mg	1.000	.895	1.000	.946
19	P	1.000	.962	1.000	.928
20	Alkalinity	1.000	.884	1.000	.804
21	Ca	1.000	.943	1.000	.982
22	Sal.	1.000	.974	1.000	.976
23	Col.	1.000	.962	1.000	.714
24	BOD	1.000	.562	1.000	.814
25	COD	1.000	.962	1.000	.976
26	K	1.000	.951	1.000	.956
27	NH_4N	1.000	.988	1.000	.972
28	Mn	1.000	.918	1.000	.989
29	Cr	1.000	.970	1.000	.974
30	THC	1.000	.821	1.000	.906
31	Temp.	1.000	.762	1.000	.575

The analysis of communalities for the dry season indicated that the majority of water quality parameters were well represented in the extracted components. Parameters such as nitrate (0.989), sulphate (0.986), chloride (0.987), ammonium (0.988), and chromium (0.970) exhibited particularly high communalities, suggesting that they contributed substantially to the total variance in groundwater quality. These variables can therefore be regarded as key indicators of contamination during the dry season. In contrast, parameters including pH (0.509), biological oxygen demand (BOD, 0.562), and temperature (0.762) showed lower communalities, implying a relatively minor role in explaining the observed variability.

This pattern suggests that groundwater quality deterioration in the dry season is predominantly influenced by nutrient enrichment and toxic heavy metals, likely arising from leachates associated with waste disposal, industrial activities, and residual petroleum hydrocarbons that facilitate the mobilization of such contaminants.

During the wet season, the communalities reflected a somewhat different distribution of variance. Trace metals and ionic species including zinc (0.991), iron (0.991), chloride (0.986), sulphate (0.971), calcium (0.982), and manganese (0.989) as well as chemical oxygen demand (COD, 0.976) and THC (0.906) exhibited high communalities, highlighting their significant contributions to groundwater variability under conditions of enhanced recharge.

Conversely, parameters such as turbidity (0.443) and temperature (0.575) showed weak communalities, indicating that these variables were less consistent in explaining variance across samples. The prominence of trace metals and organic-related parameters during the wet season suggests that hydrological processes, including leaching, dissolution, and redox-mediated mobilization, play major roles in transporting hydrocarbon contaminants.

These observations are consistent with established evidence that wet-season recharge promotes the infiltration of dissolved metals and organic pollutants from hydrocarbon-impacted soils into groundwater systems, thereby altering the distribution and bioavailability of contaminants.

3.2 Component Extraction and Variance Explained

The Total Variance Explained for both the dry and wet seasons provided insight into the proportion of total variability in groundwater quality accounted for by the extracted components. Tables 3 and 4 present the eigenvalues, percentage of variance, and cumulative variance associated with each principal component before and after Varimax rotation.

As shown in Table 3, the first principal component (PC1) in the dry season accounted for 68.75% of the total variance, indicating that a single dominant factor controls most of the groundwater chemistry during this period. This suggests strong interrelationships among major ions and potential influence from concentrated sources such as evaporative enrichment, ion exchange, or anthropogenic inputs intensified by low recharge conditions.

The second, third, and fourth components contributed 11.19%, 5.84%, and 4.91% of the variance, respectively, bringing the cumulative variance to 90.69%. The high total variance captured by the first few components reflects that the chemical characteristics of the groundwater are largely driven by a few key hydrochemical and contamination processes, notably nutrient loading and heavy metal mobilization associated with industrial and domestic activities during the dry season.

In contrast, during the wet season (Table 4), the first component (PC1) explained 63.17% of the total variance, while the second, third, and fourth components accounted for 11.84%, 6.09%, and 5.55%, respectively, with a cumulative variance of 86.66%. The slightly lower variance explained by PC1 compared to the dry season suggests a more complex hydrochemical regime under recharge conditions.

This could be attributed to the dilution effect of rainfall and increased leaching of both inorganic and organic constituents from surface and subsurface materials. The redistribution of variance among multiple components in the wet season also implies the involvement of mixed geogenic and anthropogenic processes, including metal dissolution, redox reactions, and hydrocarbon-related influences that become more pronounced when the water table rises.

Table 3: Total Variance Explained for Dry Season

Component	**Initial Eigenvalues**			**Extraction Sums of Squared Loading**			**Rotation Sums of Squared Loading**		
S/N	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	21.313	68.751	68.751	21.31	68.751	68.751	18.69	60.282	60.282
2	3.467	11.185	79.936	3.467	11.185	79.936	4.729	15.256	75.538
3	1.812	5.844	85.780	1.812	5.844	85.780	2.466	7.954	83.492
4	1.522	4.911	90.691	1.522	4.911	90.691	2.232	7.200	90.691
5	.983	3.172	93.864						
6	.620	2.001	95.864						
7	.484	1.562	97.427						
8	.365	1.176	98.603						
9	.130	.418	99.021						
10	.114	.369	99.390						
11	.047	.151	99.541						
12	.045	.145	99.685						
13	.028	.090	99.775						
14	.019	.062	99.837						
15	.013	.041	99.879						
16	.010	.033	99.912						
17	.008	.026	99.938						
18	.008	.025	99.963						
19	.004	.012	99.975						
20	.003	.010	99.985						
21	.002	.008	99.993						
22	.001	.003	99.996						
23	.000	.002	99.998						
24	.000	.001	99.999						
25	.000	.000	100.000						
26	7.32E-5	.000	100.000						
27	4.44E-5	.000	100.000						
28	3.35E-5	.000	100.000						
29	9.87E-7	3.19E-6	100.000						
30	1.24E-7	4.01E-7	100.000						
31	1.0E-16	3.4E-16	100.000						

Table 4: Total Variance Explained for Wet Season

Component	**Initial Eigenvalues**			**Extraction Sums of Squared Loadings**			**Rotation Sums of Squared Loadings**		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	20.848	63.174	63.174	20.85	63.174	63.174	14.84	44.981	44.981
2	3.909	11.844	75.019	3.909	11.844	75.019	7.397	22.416	67.396
3	2.009	6.089	81.108	2.009	6.089	81.108	3.989	12.089	79.485
4	1.833	5.553	86.661	1.833	5.553	86.661	2.368	7.176	86.661
5	.962	2.915	89.576						
6	.876	2.654	92.230						
7	.554	1.678	93.908						
8	.471	1.427	95.336						
9	.409	1.239	96.575						
10	.341	1.032	97.607						
11	.243	.737	98.344						
12	.171	.518	98.862						
13	.103	.314	99.176						
14	.088	.266	99.442						
15	.057	.173	99.615						
16	.042	.129	99.743						
17	.023	.069	99.812						
18	.019	.056	99.869						
19	.012	.036	99.904						
20	.009	.028	99.932						
21	.006	.019	99.951						
22	.005	.014	99.966						
23	.003	.010	99.976						
24	.003	.009	99.985						
25	.002	.006	99.991						
26	.001	.004	99.995						
27	.001	.002	99.997						
28	.000	.001	99.998						
29	.000	.001	100.000						
30	9.03E-5	.000	100.000						
31	1.56E-5	4.74E-5	100.000						

Overall, the pattern observed across both seasons indicates that a few principal factors dominate the hydrochemical variability, but their relative strength shifts with seasonal dynamics reflecting the interplay between dry-season concentration effects and wet-season recharge and mobilization mechanisms. The results of the total variance explained were further supported by the scree plots for both seasons, as presented in Figures 2 and 3.

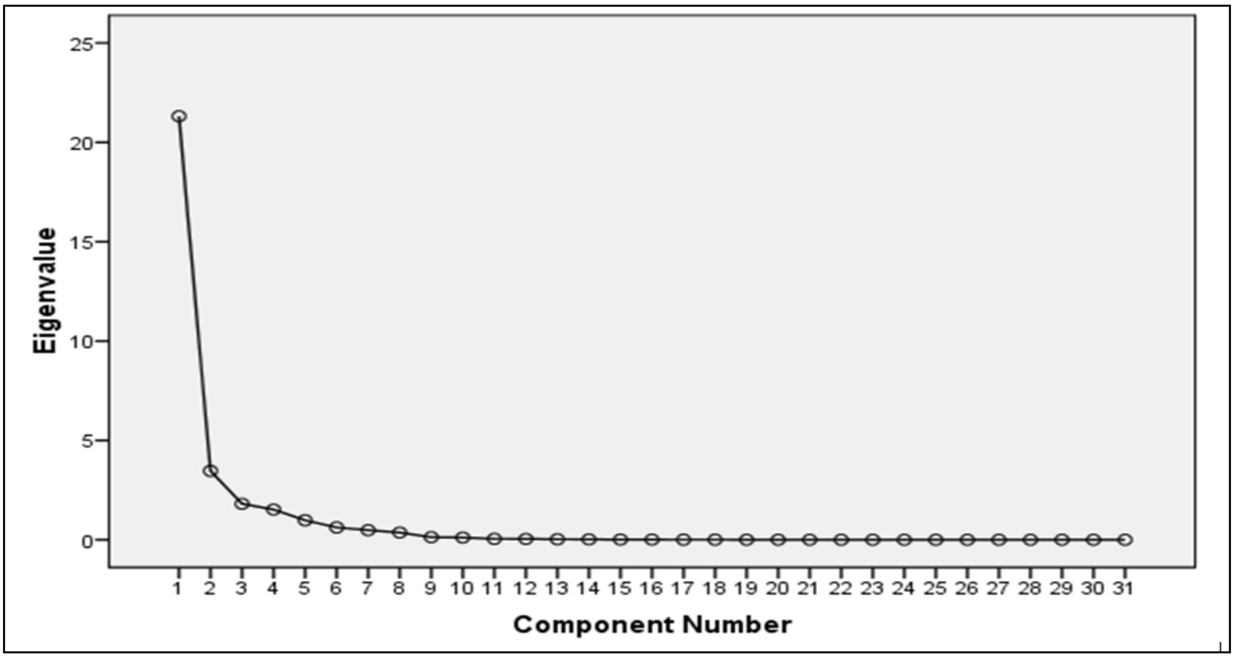

Figure 2: Scree Plot for Dry Season

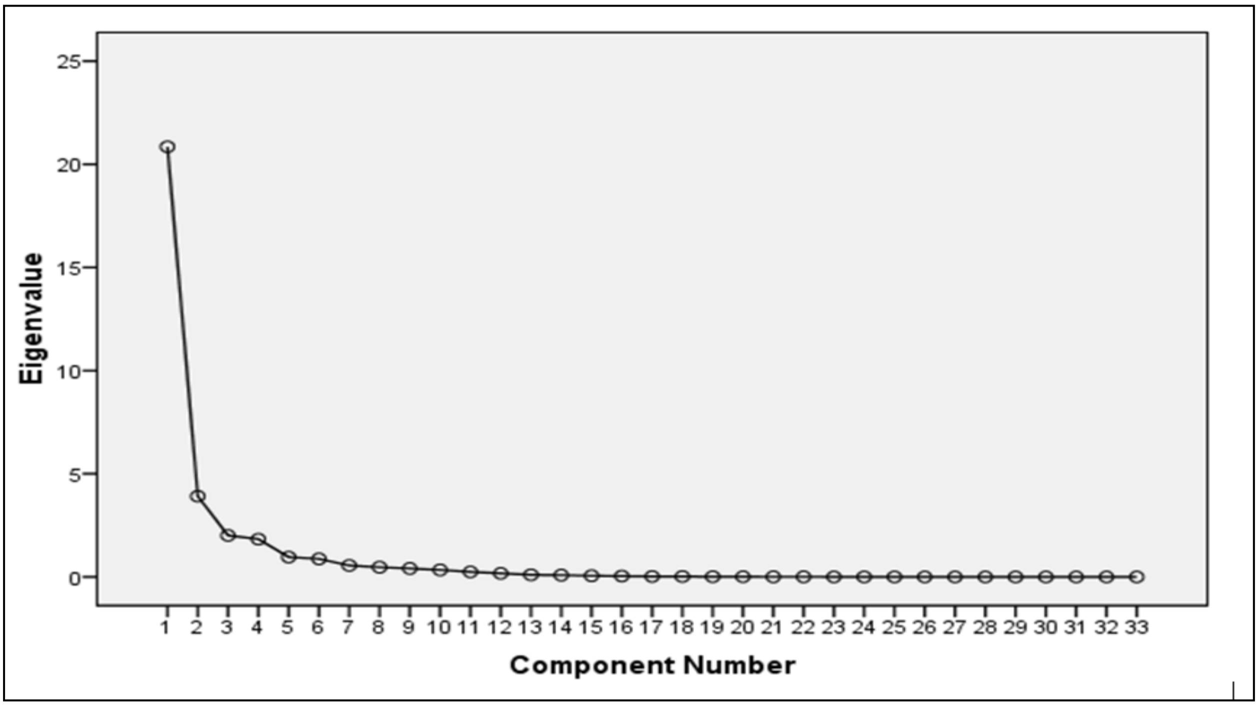

Figure 3: Scree Plot for Wet Season

The scree plots illustrate the distribution of eigenvalues across all principal components, providing a visual confirmation of the number of components to retain. In the dry season (Figure 2), a sharp decline in eigenvalues is observed after the first component, followed by smaller decreases for the second and third components, with the curve leveling off after the fourth component indicating that four components adequately capture the majority of the variance. Similarly, in the wet season (Figure 3), the eigenvalues drop steeply after the first component, with the “elbow” occurring at the fourth component, consistent with the results from the total variance explained table. These scree plots validate the selection of four principal components for both dry and wet seasons, confirming that they effectively summarize the dominant factors influencing groundwater quality.

3.3 Component Interpretation

The rotated component matrices highlighted the dominant factors influencing groundwater quality in both seasons (Table 5). In the dry season, four components were extracted. The first, an Ionic/Salinity–Nutrient Factor (60% variance), was dominated by nitrate, EC, TDS, sulphate, chloride, bicarbonate, calcium, phosphate, ammonium, and salinity, reflecting ionic enrichment and nutrient loading from agricultural runoff, sewage, and waste leachates. The second component, a Heavy-Metal Factor (15%), included Pb, Cd, Cr, Cu, Fe, and Zn, indicating contamination from industrial effluents and petroleum residues. The third, a Turbidity/Microbial Factor (7.95%), represented microbial contamination and suspended solids, while the fourth, a Physical–Organic/Thermal Factor (7.2%), captured influences of TSS, temperature, and alkalinity from organic matter decomposition and thermal variations this findings is similar to (Ali et al. 2024).

For the wet season, four components were similarly identified. The first, an Ionic–Nutrient/Salinity Factor (58–60%), included major ions and nutrients, reflecting leaching and dilution effects under recharge. The second, a Heavy-Metal Factor (14–16%), comprised Pb, Cd, Cr, Cu, Zn, Fe, and Mn, showing mobilization of metals during infiltration. The third, a Microbial/Turbidity Factor (8–9%), indicated surface-derived microbial contamination and suspended solids, while the fourth, an Organic–Thermal/Hydrocarbon Factor (6–7%), highlighted the influence of TSS, temperature, and total hydrocarbons under rainfall-driven mobilization this is further comfirmed by the study of (Aduojo et al. 2024).

Overall, groundwater quality is primarily controlled by ionic enrichment, nutrient loading, heavy metals, microbial activity, and organic/thermal processes, with the wet season showing a more distributed influence due to recharge and contaminant mobilization, while the dry season is dominated by concentrated nutrients and metals.

Table 5: Summary Results of Critical Variables Classification

Season	Component	Dominant Factor (Interpretation)	Variance Explained (%)	Key Variables with High Loadings (≥0.7)
Dry Season	1	Ionic / Salinity–Nutrient Factor	60%	NO_3 (.969), EC (.959), TDS (.959), SO_4 (.961), Cl (.960), HCO_3 (.917), Ca (.912), P (.926), NH_4N (.960), Salinity (.958), COD (.912), Na (.869), K (.865), Mg (.823), Mn (.857)
	2	Heavy-Metal / Trace Contaminant Factor	15%	Pb (.892), Cd (.899), Cr (.740), Cu (.586), Fe (.840), Zn (.859)
	3	Turbidity / Microbial Contamination Factor	7.95%	Turbidity (.943), Coliform (.960), DO (.668), BOD (.549)
	4	Physical–Organic / Thermal Factor	7.2%	TSS (.816), Temperature (.803), Alkalinity (.760), BOD (.401)
Wet Season	1	Ionic–Nutrient / Salinity Factor	58–60%	NO_3 (.865), EC (.870), TDS (.817), Salinity (.892), SO_4 (.899), Cl (.686), HCO_3 (.932), Ca (.922), Mg (.926), P (.940), K (.912), NH_4N (.827), COD (.923)
	2	Heavy-Metal Contamination Factor	14–16%	Pb (.900), Cd (.885), Cr (.902), Cu (.786), Zn (.736), Fe (.743), Mn (.707), Cl (.501)
	3	Microbial / Turbidity Factor	8–9%	Turbidity (–.609), Coliform (–.811), BOD (.771), TSS (.707)
	4	Organic–Thermal / Hydrocarbon Factor	6–7%	TSS (.707), Temperature (–.579), Total Hydrocarbon Content (–.356), Coliform (–.811)

Results of the rotated component matrix was further cemented with the component plot in rotated space presented in Figures 4 and 5 respectively

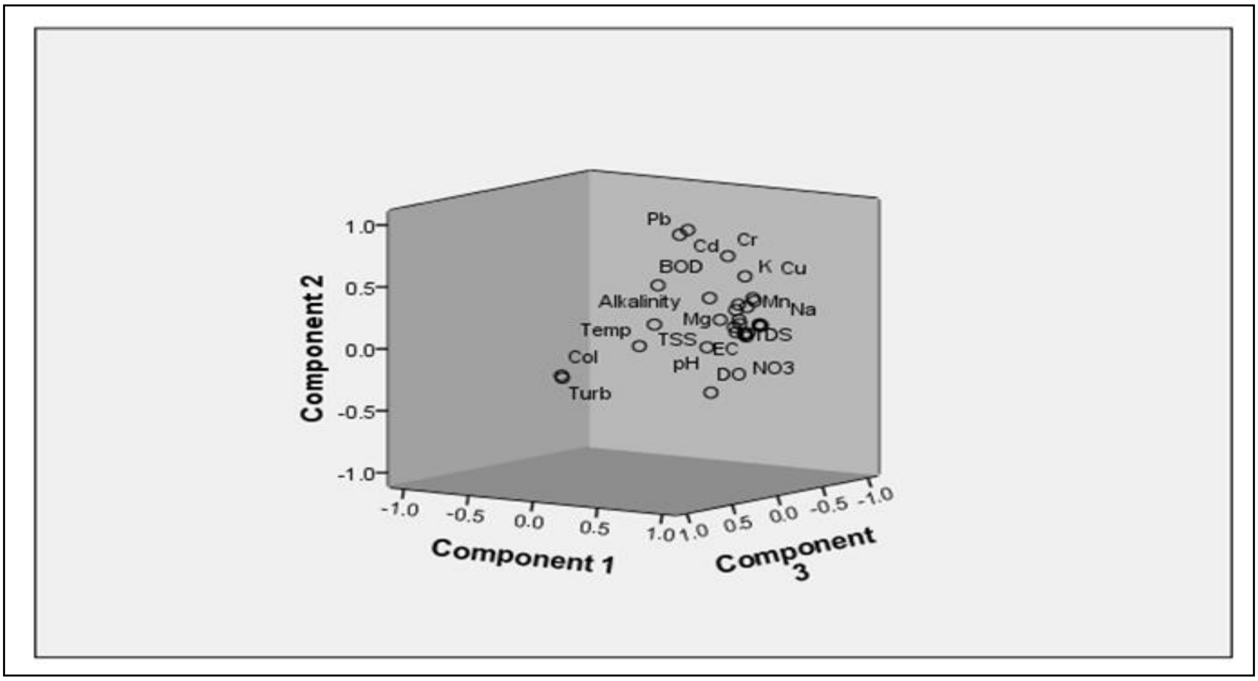

Figure 4: Component Plot in Rotated Space for Dry Season

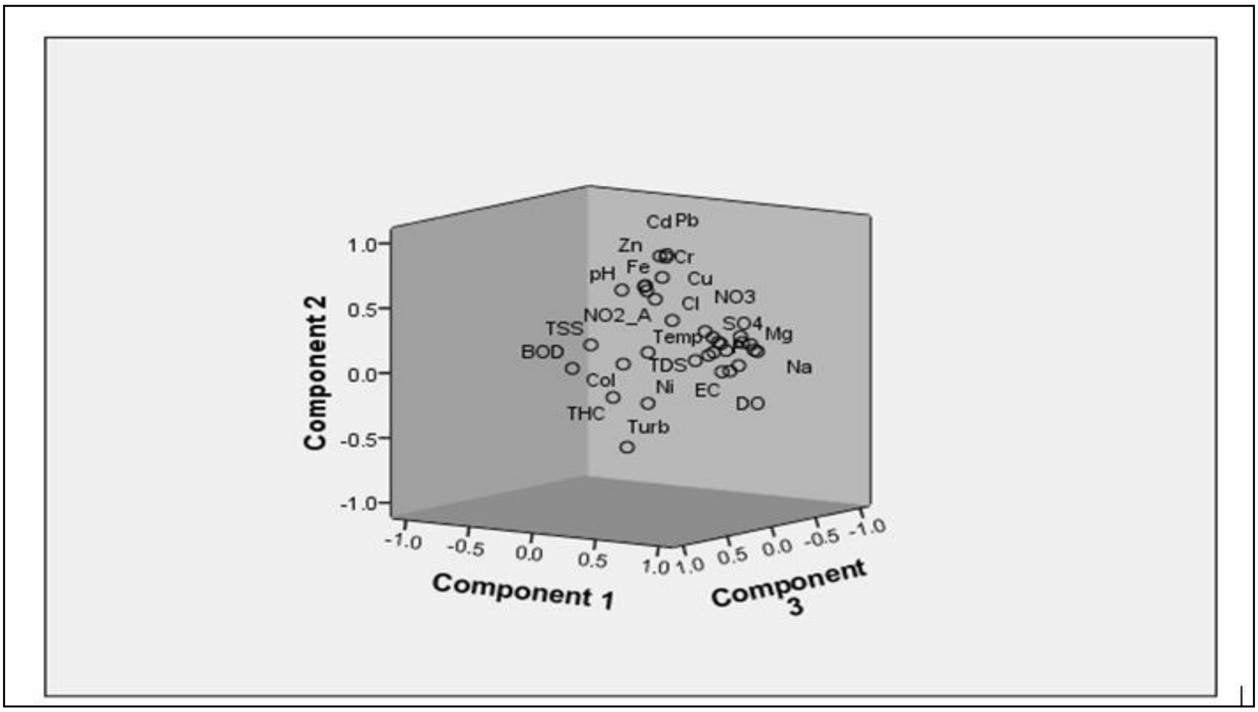

Figure 5: Component Plot in Rotated Space for Wet Season

Figures 4 and 5 present the component plots in rotated space for the dry and wet seasons, respectively. These three-dimensional visualizations illustrate how the measured groundwater quality parameters are grouped under the extracted principal components, confirming the factor classification obtained in the rotated component matrices.

For the dry season, the plot shows a clear clustering of ionic and nutrient variables (NO_3, EC, TDS, SO_4, Cl, HCO_3, Ca, NH_4N, and salinity) along the first component axis, indicating their dominant role in explaining the variance in groundwater quality. Heavy metals such as Pb, Cd, and Cr, together with Cu and Fe, cluster distinctly along the second component axis, supporting the identification of a heavy-metal contamination factor. Turbidity and coliforms plot away from the ionic and metal clusters, aligning with the third component and confirming their microbial and suspended solids-related influence. Finally, TSS, temperature, and alkalinity separate into the fourth component, reflecting organic–thermal processes. The spatial segregation of these clusters demonstrates that different processes ionic enrichment, trace metal mobilization, microbial contamination, and organic/thermal effects independently shape groundwater chemistry during the dry season.

In contrast, the wet season component plot shows a more compact but interconnected clustering of variables. Ionic and nutrient species again dominate the first component, but the grouping is broader, with phosphate, potassium, and ammonium tightly aligned with EC, TDS, and major anions. Heavy metals (Pb, Cd, Cr, Cu, Fe, Zn, and Mn) form a distinct but stronger cluster compared to the dry season, reflecting enhanced mobilization during recharge. Microbial indicators (coliforms, BOD, turbidity) plot closer to the center of the component space, suggesting that their influence overlaps with both ionic and organic factors in the wet season. Notably, total hydrocarbon content (THC) aligns with TSS and temperature under the fourth component, illustrating the hydrocarbon mobilization effect during rainfall and runoff infiltration. This configuration indicates that recharge processes intensify interactions among chemical, microbial, and hydrocarbon contaminants, producing a more distributed pattern of influence compared to the dry season.This align with the study of (Longe and Edoka, 2025)

The rotated component matrices and component plots in rotated space revealed that the underlying factors were dominated by groups of chemical, microbial, and hydrocarbon-related parameters, but their relative importance varied with season. By isolating the dominant loadings, it was possible to identify critical parameters that acted as indicators of groundwater impairment. The summary table of critical parameters for both seasons is presented in Table 6

Table 6: Summary Table of Critical Parameters Influencing Groundwater Quality

Contamination Factor	Critical Parameters (Dry Season)	Critical Parameters (Wet Season)
Ionic–Nutrient Factor	NO_3, Cl^-, SO_4^{2-}, HCO_3^-, Salinity, EC, TDS, Ca, NH_4-N	NO_3, Cl^-, SO_4^{2-}, HCO_3^-, PO_4^{3-}, K^+, NH_4-N, EC, TDS, Salinity
Heavy Metal Factor	Pb, Cd, Cr, Cu, Fe	Pb, Cd, Cr, Cu, Fe, Zn, Mn
Microbial–Turbidity Factor	Coliforms, Turbidity	Coliforms, Turbidity, BOD
Organic–Thermal / Hydrocarbon Factor	TSS, Temperature, Alkalinity	TSS, Temperature, Alkalinity, THC

4.0 Conclusion

This study applied Principal Component Analysis (PCA) to identify and interpret the dominant factors influencing groundwater quality in a hydrocarbon-impacted environment. Four principal components explained over 80% of the total variance in both dry and wet seasons, reflecting the major controlling processes, including ionic–nutrient enrichment, heavy-metal contamination, microbial activity, and organic–thermal dynamics. Hydrocarbons were specifically detected during the wet season, highlighting their mobilization under recharge conditions and the potential for pollutant transport into aquifers. Seasonal variations indicated that while dilution occurs during rainfall, trace metals and nutrients remain persistent, emphasizing their long-term impact on groundwater quality. The study provides essential baseline information for groundwater monitoring and management, identifying critical indicators such as nitrate, chloride, sulphate, lead, and cadmium. Limitations include the lack of direct correlation mapping for hydrocarbons and the absence of long-term temporal monitoring, suggesting that future work should integrate spatial hydrogeochemical analysis and continuous sampling to better guide remediation strategies in hydrocarbon-producing regions.

\

References

Aduojo, A. A., Mosobalaje, O. O., Uchegbulam, O., Johnson, A. O., & Ifeanyi, O. (2024). Multivariate analysis of seasonal changes of chemical elements in groundwater around Solous III dumpsite, Lagos, South-West Nigeria. *Scientific African*, *23*, e02084. https://doi.org/10.1016/j.sciaf.2024.e02084

Ali, S., Verma, S., Agarwal, M. B., Islam, R., Mehrotra, M., Deolia, R. K., Kumar, J., Singh, S., Mohammadi, A. A., Raj, D., Gupta, M. K., Dang, P., & Fattahi, M. (2024). Groundwater quality assessment using water quality index and principal component analysis in the Achnera block, Agra district, Uttar Pradesh, Northern India. *Sci Rep* **14**, 5381 (2024). https://doi.org/10.1038/s41598-024-56056-8

Arıman, S., Göksu Soydan-Oksal, N., Beden, N. and 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), p. 1570.

Eze, S. U., and Chinemelu, E. S. (2025) 'Evaluation of hydrocarbon contamination in groundwater through electrical resistivity imaging and geochemical techniques', Earth Science Malaysia,

Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics (4th ed.). London: Sage Publications Ltd.

Hair, J.F., Black, W.C., Babin, B.J., and Anderson, R.E. (2010). Multivariate Data Analysis (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Ikenna, S. J., and Chukwudozie, O. S. (2024) 'Groundwater Quality Assessment in Crude Oil-Impacted Communities in Ogoniland in the Niger Delta Using the Water Quality Index (WQI)', International Journal of Research in Engineering, Science and Management, Vol. 7, No. 9.

Jolliffe, I.T. (2002). Principal Component Analysis (2nd ed.). New York: Springer-Verlag.

Jolliffe, I.T. and Cadima, J., (2016). Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), p.20150202.

Kaiser, H.F., (1974). An index of factorial simplicity. Psychometrika, 39(1), pp.31–36.

Longe, E., and Edoka, J. (2025). Principal Component Analysis of Seasonal Trends in Groundwater Contamination Around the Abule-Egba Dumpsite, Lagos, Nigeria. *Technical Journal*, *30*(03), 22-32. Retrieved from https://tj.uettaxila.edu.pk/index.php/technical-journal/article/view/2340

Oyebamiji, A., Hoque, M.A. & Whitworth, M., 2024. Regional Assessment of Groundwater Contamination Risk from Crude Oil Spillages in the Niger Delta: A Novel Application of the Source-Pathway-Receptor Model. *Earth Systems and Environment*

Osisanya, W. O., Akpeji, B. H., Agho, I. F., Saleh, S. A., and Oyanameh, O. E. (2024) 'Integration of Heavy Metal Indexes and Health Risk Assessment in Groundwater Studies in Urban Area of Port

Harcourt, Niger Delta Region of Nigeria', Anchor University Journal of Science and Technology, 5(2).

Rashid, B., Destrade, M. and Gilchrist, M.D., (2012). Determination of friction coefficient in unconfined compression of brain tissue. Journal of the Mechanical Behavior of Biomedical Materials, 14, pp.163–171.

Sahmoun, A.E., Case, L.D., Jackson, S.A. and Schwartz, G.G., (2005). Cadmium and prostate cancer: a critical epidemiologic analysis. Cancer Investigation, 23(3), pp.256–263.

Sahoo, M.M., Patra, K.C. and Khatua, K.K., (2015). Inference of water quality index using ANFIA and PCA. Aquatic Procedia, 4, pp.1099–1106.

Shrestha, N., (2021). Factor analysis as a tool for survey analysis. American Journal of Applied Mathematics and Statistics, 9(1), pp.4–11.

Shrestha, S. and Kazama, F., (2007). Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji River Basin, Japan. Environmental Modelling & Software, 22(4), pp.464–475

Simeonov, V., Simeonova, P. and Tzimou-Tsitouridou, R., (2004). Chemometric quality assessment of surface waters: two case studies. Chemia Inżynieria i Ekologia, 11(6), pp.449–469.

Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A. and Kouimtzis, T., (2003). Assessment of the surface water quality in northern Greece. Water Research, 37(17), pp.4119–4124.

Singh, K.P., Malik, A. and Sinha, S., (2005). Water quality assessment and apportionment of pollution sources of Gomti River (India) using multivariate statistical techniques—A case study. Analytica Chimica Acta, 538(1–2), pp.355–374

Sousa, S., Martins, F.G., Alvim-Ferraz, M.C. and Pereira, M.C., (2007). Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environmental Modelling & Software, 22(1), pp.97–103

Su, S., Zhi, J., Lou, L., Huang, F., Chen, X. and Wu, J., (2011). Spatio-temporal patterns and source apportionment of pollution in Qiantang River (China) using neural-based modeling and multivariate statistical techniques. Physics and Chemistry of the Earth, Parts A/B/C, 36(9–11), pp.379–386

Sunardi, S., Ariyani, M., Agustian, M., Withaningsih, S., Parikesit, P., Juahir, H. and Abdoellah, O.S., (2020). Water corrosivity of polluted reservoir and hydropower sustainability. Scientific Reports, 10(1), pp.1–8

Sutadian, A.D., Muttil, N., Yilmaz, A.G. and Perera, B.J.C., (2017). Using the analytic hierarchy process to identify parameter weights for developing a water quality index. Ecological Indicators, 75, pp.220–233.

Tabachnick, B.G. and Fidell, L.S., (2001). Using Multivariate Statistics. 4th ed. New York: Harper & Row.

Taş, B., Tepe, Y., Ustaoğlu, F. and Alptekin, S., (2019). Benthic algal diversity and water quality evaluation by biological approach of Turnasuyu Creek, NE Turkey. Desalination and Water Treatment, 155, pp.402–415.

Thurston, G.D., Ito, K. and Lall, R., (2011). A source apportionment of US fine particulate matter air pollution. Atmospheric Environment, 45(24), pp.3924–3936.

Yong, A.G. and Pearce, S., (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutorials in Quantitative Methods for Psychology, 9(2), pp.79–94.

Cite this article as:

Omena, A.L., Ihimekpen, N.I., and Okonofua, S.E. (2025). Principal Component Analysis of Hydrochemical Parameters in Hydrocarbon-Affected Groundwater at Otorogu Delta State, Nigeria. *Nigerian Journal of Environmental Sciences and Technology*, 9(2), pp. 149-163. https://doi.org/10.36263/nijest.2025.02.63