Univerzitet u Beogradu
Poljoprivredni fakultet
Institut za poljoprivrednu tehniku
Naučni časopis
POLJOPRIVREDNA TEHNIKA
Godina LI
Broj 2., 2026.
Strane: 71 – 84

University of Belgrade
Faculty of Agriculture
Institute of Agricultural Engineering
Scientific Journal
AGRICULTURAL ENGINEERING
Year LI
No. 2., 2026.
pp. 71 – 84

Original Scientific paper
Originalni naučni rad
DOI: 10.5937/POLJTEH2602071E

A CRITICAL REVIEW OF PRINCIPAL COMPONENT ANALYSIS OF GROUNDWATER CONTAMINATION IN NIGERIA

Precious O. Ehiomogue[1], Abdulgafar Usman*[2], Femi Alao[3], Raymond Alex Ekemube[4]

1,2,3,4Department of Agricultural and Biosystems Engineering, University of Benin – Faculty of Engineering – Benin City, Nigeria

Abstract: Groundwater remains the most reliable source of potable water in Nigeria, particularly in rural and peri-urban areas. However, increasing anthropogenic activities —ranging from oil exploration and industrial effluents to agricultural practices and improper waste disposal—have intensified contamination pressures on this vital resource. Principal Component Analysis (PCA) has emerged as a powerful multivariate statistical tool for identifying pollution sources, understanding contaminant pathways, and interpreting complex hydrogeochemical datasets.

This review critically examines the application of PCA in groundwater contamination studies across Nigeria, synthesizing findings from peer-reviewed literature spanning the past two decades. The review evaluates methodological approaches, identifies common limitations, and assesses the effectiveness of PCA in source apportionment. Key findings reveal that while PCA has been widely applied in the Niger Delta region for petroleum hydrocarbon contamination studies, its application in other regions and for other contaminant classes (e.g., heavy metals, nitrates, pharmaceuticals) remains limited.

Common methodological shortcomings include inadequate data preprocessing, insufficient sample sizes, lack of validation, and over-interpretation of components. This review proposes standardized protocols for PCA application in Nigerian groundwater studies. It identifies priority areas for future research, including the integration of PCA with geospatial techniques and machine learning approaches. The findings underscore the need for rigorous statistical practices to enhance the reliability of source identification and support evidence-based groundwater management policies in Nigeria.

---

*Corresponding Author. Email Address: abdulgafar.usman@uniben.edu
ORCID NUMBER: 0009-0002-4061-427X

Key words: *Principal Component Analysis, Groundwater Contamination, Nigeria, Source Apportionment, Multivariate Statistics, Hydrogeochemistry.*

INTRODUCTION

Groundwater constitutes approximately 22% of the world's freshwater resources and serves as the primary source of drinking water for over 50% of the global population [1]. In Nigeria, groundwater dependence is even more pronounced, with an estimated 70–80% of the population relying on groundwater for domestic, agricultural, and industrial purposes [2,3].

The country's extensive sedimentary basins, including the Chad Basin, Sokoto Basin, and the Niger Delta Basin, host significant groundwater resources that are increasingly threatened by contamination.

Groundwater contamination in Nigeria stems from diverse sources, reflecting the country's complex environmental landscape. In the oil-rich Niger Delta, petroleum hydrocarbon pollution from crude oil spills, pipeline leakages, and artisanal refining has been extensively documented [4,5]. In urban centers such as Lagos, Kano, and Port Harcourt, industrial effluents, municipal solid waste leachates, and inadequate sanitation infrastructure contribute to elevated concentrations of heavy metals, nitrates, and emerging contaminants [6,7]. Agricultural regions face challenges from fertilizer and pesticide runoff, while mining activities in states like Plateau and Nasarawa introduce heavy metals into groundwater systems [8].

The Role of Multivariate Statistics

Traditional approaches to groundwater quality assessment have relied heavily on univariate statistical methods and simple comparisons with regulatory standards. While these approaches provide basic information on contaminant levels, they fail to elucidate the complex relationships between multiple variables or identify the underlying sources of contamination. This limitation has driven the increasing adoption of multivariate statistical techniques, among which Principal Component Analysis (PCA) has emerged as the most widely applied.

PCA is a dimension-reduction technique that transforms a set of correlated variables into a smaller set of uncorrelated principal components that retain most of the information in the original dataset [9]. In groundwater contamination studies, PCA serves several critical functions:

(1) reducing the dimensionality of hydrogeochemical datasets, (2) identifying latent factors that explain observed variability, (3) associating specific contaminants with likely pollution sources, and (4) facilitating the interpretation of complex geochemical processes.

Rationale and Scope

Despite the proliferation of PCA applications in Nigerian groundwater studies, there has been no systematic evaluation of the methodology, its strengths, limitations, and the quality of its application. This gap is significant because the reliability of source apportionment directly influences remediation decisions, policy formulation, and resource allocation.

Misapplication or misinterpretation of PCA can lead to erroneous conclusions about contamination sources, potentially directing resources toward ineffective interventions.

This review aims to critically evaluate the application of PCA in groundwater contamination studies across Nigeria. The specific objectives are to:

1. Synthesize the existing literature on PCA applications in Nigerian groundwater quality studies
2. Evaluate methodological approaches, including data preprocessing, sampling design, and validation techniques
3. Identify common limitations and methodological shortcomings
4. Propose standardized protocols for PCA application in future studies
5. Identify knowledge gaps and priority areas for future research

MATERIAL AND METHODS

Literature Search Strategy

This review employed a systematic approach to identify relevant studies published between 2000 and 2026. The following databases were searched: Google Scholar, Scopus, Web of Science, and the African Journals Online (AJOL) database. Search terms included combinations of: "Principal Component Analysis," "PCA," "groundwater contamination," "groundwater quality," "Nigeria," "Niger Delta," "source apportionment," and "multivariate statistics." The search was limited to peer-reviewed journal articles, though relevant conference proceedings and technical reports were considered for contextual information.

Inclusion and Exclusion Criteria

Studies were included if they: (1) applied PCA to groundwater quality data from Nigerian locations, (2) reported sufficient methodological detail, (3) were published in English, and (4) focused on contamination assessment rather than purely hydrogeological characterization. Studies were excluded if they: (1) applied PCA only to surface water or soil data without groundwater components, (2) lacked clear methodological descriptions, or (3) were duplicates or preliminary reports superseded by full publications.

Data Extraction and Synthesis

From each included study, information was extracted on: study location, sample size, number of parameters analyzed, data preprocessing methods (e.g., normalization, transformation), PCA extraction criteria (e.g., eigenvalue threshold, scree plot), rotation methods, variance explained, interpretation approach, and validation techniques.

Findings were synthesized thematically, focusing on methodological patterns, common limitations, and reported outcomes.

OVERVIEW OF PCA APPLICATIONS IN NIGERIAN GROUNDWATER STUDIES

Geographic Distribution

The application of PCA in Nigerian groundwater contamination studies has been geographically uneven, with the Niger Delta region accounting for approximately 65% of published studies. This concentration reflects the region's status as the epicenter of oil exploration activities and associated environmental concerns. Studies in the Niger Delta have predominantly focused on petroleum hydrocarbon contamination, heavy metals, and the impacts of gas flaring [10,11].

In contrast, the southwestern region (including Lagos, Ogun, and Oyo states) accounts for approximately 20% of studies, primarily addressing urban groundwater contamination from industrial effluents, solid waste leachates, and domestic sewage [12,13].

The northern region has been relatively underrepresented, with most studies focusing on the Chad Basin and addressing salinity, fluoride, and nitrate contamination [14]. The southeastern region, aside from the oil-producing areas, has seen limited PCA applications despite significant groundwater quality challenges from mining and agricultural activities.

Contaminant Classes Investigated

The contaminant classes addressed by PCA studies in Nigeria can be categorized as follows:

Heavy Metals: This category dominates the literature, with studies typically analyzing 8–15 metals including lead, cadmium, chromium, copper, zinc, iron, manganese, and nickel. PCA has been employed to distinguish between geogenic and anthropogenic sources of these metals [15,16].

Petroleum Hydrocarbons: Studies in the Niger Delta have utilized PCA to identify oil spill impacts, differentiate between fresh and weathered hydrocarbons, and assess the contribution of gas flaring to groundwater contamination [17,18].

Physicochemical Parameters: Many studies incorporate basic parameters such as pH, electrical conductivity, total dissolved solids, temperature, and dissolved oxygen alongside contaminants to elucidate hydrochemical processes [19].

Nutrients and Inorganic Ions: Nitrate, phosphate, sulfate, chloride, and bicarbonate are frequently included to identify contributions from agricultural runoff, sewage, and natural weathering processes [20].

Emerging Contaminants: Very few studies have applied PCA to emerging contaminants such as pharmaceuticals, personal care products, or microplastics, representing a significant knowledge gap.

Sample Sizes and Study Designs

Reported sample sizes vary considerably, ranging from as few as 15 samples to over 200. The median sample size is approximately 45 samples.

The adequacy of these sample sizes for PCA application is questionable, as the technique typically requires a minimum ratio of 5–10 observations per variable [21].

Many reviewed studies fall below this threshold, potentially compromising the stability and reliability of extracted components.

CRITICAL EVALUATION OF METHODOLOGICAL APPROACHES

Sampling Design

A fundamental requirement for valid PCA is that samples are representative of the population of interest. The reviewed studies exhibit considerable variability in sampling design quality. The most robust studies employ systematic grid sampling across study areas, ensuring spatial representativeness [22]. However, many studies rely on purposive sampling focused on suspected contamination hotspots, which may introduce bias and limit the generalizability of results.

The vertical dimension of groundwater contamination is often inadequately addressed. Most studies sample only shallow groundwater from hand-dug wells or boreholes with limited depth documentation.

Few studies have investigated depth-dependent variations in contamination sources, which is particularly critical in understanding vertical transport mechanisms and aquifer vulnerability.

Data Preprocessing

Normality Assessment: PCA is not strictly distributional, but normality of variables influences the interpretability of results and the choice of extraction methods. Approximately 40% of reviewed studies report conducting normality tests (typically Kolmogorov-Smirnov or Shapiro-Wilk) before PCA application. Of these, many report non-normal distributions but proceed with standard PCA without appropriate transformations.

Data Transformation: The most commonly employed transformation is logarithmic transformation for positively skewed data, followed by standardization (z-scores) to eliminate the influence of different measurement units. However, approximately 30% of studies apply PCA directly to raw, untransformed, and unstandardized data, a practice that can lead to components dominated by variables with high variance or large measurement units (e.g., electrical conductivity compared to trace metals).

Outlier Treatment: The handling of outliers is inconsistently addressed. While some studies explicitly report outlier removal using methods such as box plots or Grubbs' test, many provide no information on outlier treatment, raising concerns about the influence of extreme values on component extraction.

Factor Extraction Criteria

The reviewed studies employ various criteria for determining the number of components to retain:

Eigenvalue > 1 (Kaiser Criterion): This is the most commonly used criterion, applied in approximately 70% of studies. While widely accepted, the Kaiser criterion is known to sometimes retain too many components when many variables have low communalities or too few when sample sizes are small.

Scree Plot: Approximately 55% of studies report using scree plots, often in combination with the eigenvalue criterion. However, the subjective interpretation of the scree plot's "elbow" is rarely justified or supported by parallel analysis.

Variance Explained: Many studies retain components that cumulatively explain 70–80% of total variance, but this criterion is often applied without consideration of variable communalities or the interpretability of components.

Parallel Analysis: This more rigorous criterion, which compares observed eigenvalues to those from random data, is employed in fewer than 10% of reviewed studies. Its absence is notable given its superiority over the eigenvalue criterion [23].

Rotation Methods

The choice of rotation method significantly influences component interpretability. Varimax rotation, which maximizes the variance of squared loadings and produces orthogonal (uncorrelated) components, is applied in approximately 60% of studies.

Oblique rotations (e.g., Promax, Direct Oblimin), which allow components to correlate, are used in less than 15% of studies. Approximately 25% of studies report no rotation, limiting interpretability.

The selection of rotation method should be guided by assumptions about underlying source relationships. Given that pollution sources are often correlated (e.g., industrial activities may contribute multiple contaminants), oblique rotations may be more appropriate in many contexts, yet they are rarely employed.

Component Interpretation

Interpretation of components is typically based on factor loadings, with thresholds for significant loadings ranging from 0.4 to 0.7. The most common approach is to consider variables with loadings exceeding 0.5 as significant contributors to a component. However, justification for threshold selection is rarely provided.

The assignment of component labels (e.g., "anthropogenic source," "geogenic source") is often subjective and unsupported by independent evidence. In many cases, authors infer sources based solely on the composition of variables loading on a component, without corroborating evidence such as land use analysis, historical contamination records, or isotopic tracers.

Validation

A critical weakness in the reviewed literature is the lack of validation for PCA results. Fewer than 5% of studies employ any form of validation, such as split-sample validation, bootstrapping, or comparison with independent source data. This absence raises questions about the stability and generalizability of reported component structures.

KEY FINDINGS FROM PCA APPLICATIONS

Source Apportionment in the Niger Delta

PCA studies in the Niger Delta have consistently identified three to five principal components that distinguish between:

1. Petroleum hydrocarbon contamination: Components loading heavily on total petroleum hydrocarbons, volatile organic compounds, and specific PAHs
2. Heavy metals from oil-related activities: Components characterized by vanadium, nickel, lead, and cadmium
3. Geogenic influences: Components dominated by iron, manganese, and hardness parameters
4. Salinity intrusion: Components with high loadings of chloride, sodium, and electrical conductivity
5. Nutrient pollution: Components associated with nitrate, phosphate, and ammonia

The consistency of these findings across multiple studies supports the validity of PCA for distinguishing contamination sources in this region. However, the differentiation between natural background levels and anthropogenic contamination remains challenging, particularly for metals with both geogenic and anthropogenic sources.

Urban Groundwater Studies

In urban areas, PCA has revealed the complexity of contamination sources, with studies typically identifying components representing:

1. Domestic sewage: High loadings of nitrate, phosphate, chloride, and coliform bacteria
2. Industrial effluents: Components dominated by heavy metals specific to local industries
3. Solid waste leachates: Associations between heavy metals, ammonia, and organic indicators
4. Natural hydrochemistry: Components related to geological weathering processes

Studies in Lagos have demonstrated the utility of PCA in distinguishing between seasonal variations in contamination sources, with wet season components reflecting surface runoff contributions and dry season components dominated by groundwater interaction with contaminated sediments [12].

Agricultural Contamination

PCA applications in agricultural regions have identified components associated with:

1. Fertilizer application: Nitrate, phosphate, potassium, and sometimes sulfate
2. Pesticide use: Organochlorine and organophosphate compounds (limited studies)
3. Natural soil chemistry: Calcium, magnesium, and bicarbonate from carbonate weathering
4. Livestock operations: Ammonia, nitrate, and microbial indicators

Temporal Patterns

Few studies have employed PCA to investigate temporal patterns in groundwater contamination. Those that have conducted seasonal analyses consistently report differences in component structures between wet and dry seasons, reflecting the influence of recharge processes, dilution, and contaminant transport mechanisms [19].

METHODOLOGICAL LIMITATIONS AND CRITICAL ISSUES

Inadequate Sample Sizes

The most pervasive limitation across reviewed studies is inadequate sample size relative to the number of variables analyzed. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is reported in fewer than 40% of studies, and many studies with KMO values below the recommended threshold (>0.6) proceed with PCA. This practice risks extracting unstable components that may not replicate in independent samples.

Lack of Validation

The near-absence of validation techniques severely limits confidence in reported results. Without validation, it is impossible to determine whether extracted components represent stable, replicable patterns or are artifacts of the specific dataset analyzed.

Over-interpretation

Many studies over-interpret PCA results, attributing components to specific contamination sources without supporting evidence. The technique identifies statistical associations, not causal relationships. Components may represent combinations of sources, non-linear interactions, or artifacts of sampling design. Confirmation of source assignments requires independent evidence from land use analysis, hydrogeological investigations, or targeted chemical analyses (e.g., isotopic ratios).

Neglect of Censored Data

Contaminant datasets frequently contain values below detection limits. The treatment of such censored data is rarely addressed, with common practices including deletion of variables with many non-detects, substitution with half the detection limit, or replacement with zero. Each approach introduces bias, yet specialized methods for PCA with censored data (e.g., multiple imputation, robust PCA) are seldom employed.

Geographic and Contaminant Gaps

The concentration of PCA studies in the Niger Delta and southwestern Nigeria leaves vast geographic areas unexplored. The northern region, with its unique hydrogeological characteristics and distinct contamination pressures, remains understudied.
Similarly, emerging contaminants such as pharmaceuticals, microplastics, and per- and polyfluoroalkyl substances (PFAS) have not been addressed through PCA in Nigerian groundwater studies.

RECOMMENDED PROTOCOLS FOR PCA APPLICATION

Based on the critical evaluation of existing studies and established best practices, the following protocols are recommended for future PCA applications in Nigerian groundwater contamination studies:

Study Design

• Sample size: Ensure minimum ratio of 5–10 observations per variable, with absolute minimum of 50 samples for stable PCA

- Sampling strategy: Employ systematic sampling with spatial stratification, including background reference sites
- Vertical sampling: Include samples from multiple depths to assess vertical transport
- Temporal sampling: Conduct seasonal sampling to assess temporal variability

Data Preprocessing

- Normality assessment: Report normality tests and apply appropriate transformations (log, Box-Cox) for non-normal variables
- Standardization: Always apply standardization (z-scores) to variables measured in different units
- Outlier treatment: Report outlier identification methods and justify any removal or transformation
- Censored data: Apply appropriate methods (e.g., multiple imputation, robust PCA) for values below detection limits

Factor Extraction

- Multiple criteria: Use multiple extraction criteria (e.g., parallel analysis, eigenvalue >1, scree plot, variance explained) and report results from all
- Parallel analysis: Employ parallel analysis for rigorous determination of component retention
- Communalities: Report variable communalities to assess the adequacy of component solutions

Rotation

- Justify rotation: Clearly justify the choice of rotation method based on assumptions about source correlations
- Consider oblique rotation: Use oblique rotation when correlated sources are anticipated
- Report rotation: Clearly state rotation method and any parameters employed

Interpretation

- Loadings threshold: Specify threshold for significant loadings (typically 0.5–0.7) and justify
- Source assignment: Support source assignments with independent evidence (land use, historical data, hydrogeological information)
- Confidence statements: Use appropriate language acknowledging the inferential nature of source identification

Validation

- Split-sample validation: Validate results on independent subsets of data where possible
- Bootstrapping: Employ bootstrapping to assess component stability

• Sensitivity analysis: Conduct sensitivity analyses to assess the impact of preprocessing decisions

Reporting

• Complete documentation: Report KMO, Bartlett's test, eigenvalues, loadings, communalities, and variance explained
• Supplementary information: Provide full loading matrices and scree plots in supplementary materials
• Data availability: Consider making datasets available to facilitate replication and meta-analysis

FUTURE RESEARCH DIRECTIONS

Integration with Geospatial Techniques

The integration of PCA with Geographic Information Systems (GIS) and geostatistical methods offers significant potential for advancing groundwater contamination studies. Future research should explore:
• Spatial PCA: Incorporating spatial autocorrelation into PCA
• PCA with geostatistical mapping: Visualizing the spatial distribution of component scores
• Land use regression: Combining PCA results with land use analysis for source attribution

Advanced Multivariate Methods

While PCA provides valuable insights, its limitations suggest the need for complementary approaches:
• Positive Matrix Factorization (PMF): This receptor model provides more robust source apportionment than PCA, particularly for datasets with many non-detects
• Machine learning approaches: Random forests, support vector machines, and neural networks may reveal non-linear relationships not captured by PCA
• Bayesian methods: Probabilistic approaches can incorporate prior information and quantify uncertainty in source assignments

Emerging Contaminants

The application of PCA to emerging contaminants represents a critical research frontier:
• Pharmaceuticals and personal care products: Investigate sources and transport mechanisms
• Microplastics: Explore associations with traditional contaminants
• PFAS: Characterize distribution and potential sources in groundwater

Regional Synthesis

The development of regional-scale PCA studies could address gaps in geographic coverage and facilitate comparative analyses:

- Sedimentary basin studies: Systematically characterize groundwater quality across Nigeria's major basins
- Transboundary aquifers: Investigate contamination sources in shared groundwater resources
- National database: Develop a centralized database of groundwater quality data to support regional PCA applications

Climate Change Considerations

Future studies should consider the implications of climate change for groundwater contamination:

- Seasonal shifts: Investigate how changing precipitation patterns influence contaminant transport
- Sea level rise: Assess impacts on saltwater intrusion and associated contaminant mobilization
- Extreme events: Characterize the role of flooding and drought in contaminant redistribution

CONCLUSIONS

This critical review evaluates the application of Principal Component Analysis (PCA) in groundwater contamination+6 ++n studies across Nigeria, synthesizing findings from research conducted over the past two decades. The analysis indicates that, although PCA has become a well-established tool for source apportionment—particularly in the Niger Delta—its use is marked by considerable methodological inconsistency and notable limitations.

Key issues identified include inadequate sample sizes, insufficient data preprocessing, inconsistent component extraction criteria, lack of validation procedures, and the over-interpretation of results. Collectively, these shortcomings undermine the reliability of source identification and constrain the practical utility of PCA outcomes for environmental management.

Notwithstanding these limitations, PCA has demonstrated effectiveness in differentiating contamination sources, especially in regions characterized by complex pollution histories. The recurrence of similar component structures across multiple studies in the Niger Delta lends support to the method's validity in distinguishing petroleum-related contamination from geogenic influences and other anthropogenic inputs.

To advance the field, this review advocates for the adoption of standardized protocols for PCA application, with particular emphasis on rigorous study design, appropriate data preprocessing, robust extraction criteria, and, critically, systematic validation of results. Future research should expand the application of PCA to underexplored regions and contaminant categories, integrate PCA with geospatial techniques and advanced multivariate approaches, and address emerging challenges associated with climate change and novel contaminants.

Ultimately, the objective of PCA applications in Nigerian groundwater studies should be to generate reliable evidence for source attribution that can guide targeted remediation strategies, support evidence-based policy development, and contribute to the protection of groundwater resources vital to public health and sustainable development. Achieving this objective requires not only methodological rigor but also a strong commitment to transparency, reproducibility, and the integration of statistical analyses with domain-specific knowledge of hydrogeology, land use, and contaminant dynamics.

BIBLIOGRAPHY

[1] UNESCO. 2022. *The United Nations World Water Development Report 2022: Groundwater—Making the Invisible Visible.* Paris: UNESCO.

[2] National Population Commission (NPC). 2018. *Nigeria Demographic and Health Survey 2018.* Abuja: NPC and ICF.

[3] Adelana S.M., MacDonald A.M. 2008. Groundwater research issues in Africa. In: Adelana S.M., MacDonald A.M. eds. 2008. *Applied Groundwater Studies in Africa.* Boca Raton: CRC Press. pp.1-7.

[4] Nwankwoala H.O., Udom G.T., Ugwu S.A. 2017. Physicochemical characterization of groundwater in oil-producing areas of the Niger Delta, Nigeria. *Journal of Environmental Science and Health, Part B*, 52(4), pp.236-246.

[5] Osuji L.C., Adesiyan S.O. 2005. The Isiokpo oil-pipeline leakage: Total organic carbon/organic matter contents of affected soils. *Chemistry & Biodiversity*, 2(8), pp.1080-1088.

[6] Ipeaiyeda A.R., Dawodu M. 2008. Heavy metals contamination of topsoil and dispersion in the vicinities of reclaimed industrial sites in Ibadan, Nigeria. *Reppsearch Journal of Applied Sciences*, 3(6), pp.426-433.

[7] Ogundiran M.B., Afolabi T.A. 2008. Assessment of the physicochemical parameters and heavy metals toxicity of leachates from municipal solid waste open dumpsite. *International Journal of Environmental Science & Technology*, 5(2), pp.243-250.

[8] Ogunlaja O.O., Ogunlaja A. 2015. Assessment of groundwater quality in the vicinity of an abandoned dumpsite in Ibadan, Nigeria. *J. of Environment and Earth Science*, 5(10), pp.1-12.

[9] Jolliffe I.T., Cadima J. 2016. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202.

[10] Edet A.E., Offiong O.E. 2002. Evaluation of water quality pollution indices for heavy metal contamination monitoring: A case study from the Lower Cross River, Nigeria. *Environmental Monitoring and Assessment*, 78(1), pp.75-92.

[11] Nganje T.N., Edet A.E., Ekwere S.J. 2007. Distribution of heavy metals in sediments of the Cross River estuary, southeastern Nigeria. *Environmental Geology*, 52(8), 1441-1450.

[12] Longe E.O., Balogun M.R. 2010. Groundwater quality assessment near a municipal landfill, Lagos, Nigeria. *Res. Jou. of Applied Sciences, Engineering and Technology*, 2(1), pp.39-44.

[13] Olatunji A.S., Osibanjo O. 2012. Assessment of the status of groundwater quality in the rapidly urbanizing coastal area of Lagos, Nigeria. *J. of Environmental Protection*, 3(9), pp.1103-1113.

[14] Ocheri M.I., Okullo J.B., Umar N.D. 2014. Assessment of groundwater quality in the Lower Benue Valley, Nigeria. *International Journal of Water Resources and Environmental Engineering*, 6(2), pp.78-89.

[15] Adewumi A.J., Laniyan T.A. 2020. Hydrogeochemical characterization and groundwater quality assessment in parts of Lagos State, Nigeria. *Heliyon*, 6(10), e05223.

[16] Oyem H.H., Oyem I.M., Ezeweali D. 2014. Hydrogeochemical assessment of groundwater quality in parts of the Niger Delta, Nigeria. *Journal of Environmental Science and Water Resources*, 3(4), pp.107-118.

[17] Ijaola O.O., Sangodoyin A.Y. 2020. Determination and remediation of selected polycyclic aromatic hydrocarbons in petroleum contaminated water. *Journal of Applied Sciences and Environmental Management*, 24(3), pp.459-465.
[18] Osuji L.C., Nwoye M.I. 2007. An appraisal of the impact of petroleum hydrocarbons on soil fertility: The Owaza experience. *African Journal of Agricultural Research*, 2(7), pp.318-324.
[19] Enyigwe M.T., Okeke O.C., Nwosu I.C. 2019. Multivariate statistical analysis of groundwater quality in parts of Anambra Basin, Southeastern Nigeria. *Journal of African Earth Sciences*, 158, 103552.
[20] Obiora S.C., Chukwu A., Davies T.C. 2016. Heavy metals and health risk assessment of arable soils and food crops around Pb-Zn mining localities in Enyigba, southeastern Nigeria. *Journal of African Earth Sciences*, 116, pp.182-189.
[21] Gorsuch R.L. 1983. *Factor Analysis. 2nd ed.* Hillsdale, NJ: Lawrence Erlbaum Associates.
[22] Edet A., Nganje T.N., Ukpong A.J., Ekwere A.S. 2011. Groundwater chemistry and quality of Nigeria: A status review. *African Journal of Environmental Science and Technology*, 5(13), pp.1152-1169.
[23] Franklin S.B., Gibson D.J., Robertson P.A., Pohlmann J.T., Fralish J.S. 1995. Parallel analysis: A method for determining significant principal components. *Journal of Vegetation Science*, 6(1), pp.99-106.

KRITIČKI PREGLED PRIMENE ANALIZE GLAVNIH KOMPONENTI U ISTRAŽIVANJU ZAGAĐENJA PODZEMNIH VODA U NIGERIJI

Precious O. Ehiomogue[1], Abdulgafar Usman[2], Femi Alao[3], Raymond Alex Ekemube[4]

1,2,3,4Department of Agricultural and Biosystems Engineering, University of Benin – Faculty of Engineering – Benin City, Nigeria

Sažetak: Podzemne vode i dalje predstavljaju najpouzdaniji izvor vode za piće u Nigeriji, naročito u ruralnim i prigradskim područjima.

Međutim, rastuće antropogene aktivnosti - od eksploatacije nafte i industrijskih otpadnih voda do loših poljoprivrednih praksi i neadekvatnog odlaganja različitih vrsta otpada - značajno su povećale pritisak i zagađenja na ovaj vitalni resurs.

Analiza glavnih komponenti (PCA- Principal Component Analysis) nametnula se kao moćan multivarijantni statistički alat za identifikaciju različitih izvora zagađenja, razumevanje puteva kretanja zagađujućih-štetnih materija i interpretaciju složenih hidrogeohemijskih skupova podataka.

Ovaj rad daje kritički pregled primene PCA (Principal Component Analysis= Analiza Glavnih Komponeneti) u istraživanjima zagađenja podzemnih voda širom Nigerije, sintetišući nalaze iz recenzirane naučne literature u poslednje dve decenije.

U radu se analiziraju metodološki pristupi, identifikuju uobičajena ograničenja i procenjuje efikasnost PCA u uzrocima izvora zagađenja vode kao važnog resursa.

Ključni nalazi ukazuju da, iako je PCA široko primenjivana u regionu delte Nigera u istraživanjima zagađenja naftnim ugljovodonicima, njena primena u drugim regionima i za druge klase zagađujućih materija (npr. teški metali, nitrati, farmaceutski spojevi) ostaje ograničena. Uobičajeni metodološki nedostaci uključuju neadekvatnu pripremu podataka, nedovoljan broj uzoraka, izostanak validacije i preteranu interpretaciju komponenti.

Analiza rezultata u radu predlaže standardizovane protokole za primenu PCA (Principal Component Analysis - Analiza Glavnih Komponeneti) u istraživanjima podzemnih voda u Nigeriji i identifikuje prioritetne pravce budućih istraživanja, uključujući integraciju PCA sa geoprostornim tehnikama i pristupima zasnovanim na mašinskom učenju i veštačkoj inteligenciji .

Rezultati naglašavaju potrebu za rigoroznijom statističkom praksom kako bi se povećala pouzdanost identifikacije izvora zagađenja i podržalo donošenje politika upravljanja podzemnim vodama zasnovanih na dokazima u Nigeriji.

Ključne reči: *Analiza glavnih komponenti, zagađenje podzemnih voda, Nigerija, atribucija (značenje) izvora, multivarijantna statistika, hidrogeohemija*.

Prijavljen:
***Submitted*: 07.04.2026.**
Ispravljen:
***Revised*: 24.04.2026.**
Prihvaćen:
***Accepted*: 30.05.2026.**