



## EVALUATIONS OF GROUNDWATER QUALITY USING PRINCIPAL COMPONENT ANALYSIS AND ASSOCIATED MULTIVARIATE TECHNIQUES: A CASE HISTORY IN EWEKORO COMMUNITIES, SOUTH-WEST NIGERIA

Ishola S. A.

Department of Earth Sciences, Olabisi Onabanjo University, Ago-Iwoye, P.M.B 2002, Ago-Iwoye, Ogun State, Nigeria.

Email: [ishola.sakirudeen@oouagoiwoye.edu.ng](mailto:ishola.sakirudeen@oouagoiwoye.edu.ng)

### Cite this article:

Ishola, S. A. (2024), Evaluations of Groundwater Quality using Principal Component Analysis and Associated Multivariate Techniques: A Case History in Ewekoro Communities, South-West Nigeria. African Journal of Mathematics and Statistics Studies 7(4), 57-81. DOI: 10.52589/AJMSS-TH71XIP9

### Manuscript History

Received: 12 Aug 2024

Accepted: 6 Oct 2024

Published: 10 Oct 2024

### Copyright © 2024 The Author(s).

This is an Open Access article distributed under the terms of Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), which permits anyone to share, use, reproduce and redistribute in any medium, provided the original author and source are credited.

**ABSTRACT:** *The quest for quality water consumption is comprised of integration of numerous techniques from different professionals whose contributions go a long way in determining overall health status and wellness of any given society. The water quality status of boreholes in Ewekoro communities of Ewekoro Local Government Area, South-West Nigeria was investigated in this work. Water samples were collected from 25 boreholes at various sampling stations across the study area. Standard analytical water quality methods using Inductively Coupled Mass Spectrometry (ICP-MS)/Optical Emission Spectrometry (ICP-OES) and Pour Plate Techniques were respectively employed in the laboratory for the samples' geochemical and bacteriological analyses. The raw data were processed and analyzed using Principal Component Analysis and other multivariate techniques with Excel and SPSS 20.0 statistical software packages with the corresponding outcomes systematically compared with water quality standards. For the PCA, the varimax rotation converged in 23 iterations with 6 components extracted which accounted for 92.76% of the total cumulative variance respectively observed for geochemical parameters.  $\text{Na}^{2+}$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^+$ ,  $\text{PO}_4^{3-}$ , and  $\text{HCO}_3^-$ ;  $\text{Mg}^{3+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Cd}^{2+}$ ,  $\text{Zn}^{2+}$ ,  $\text{Ni}$ ,  $\text{S}$ ,  $\text{Si}$ ,  $\text{Fe}^{3+}$ ,  $\text{NO}_3^-$  and  $\text{Al}^{3+}$  played a significant contribution with higher factor loading supported by the dominance of  $\text{HCO}_3^-$  among the water quality determinants. The rotation converged in 8 iterations with 3 extracted components that accounted for a total cumulative variance of 90.99% with pH, TEMP, EC, BOD, TS, TURB, ALK, TH, TDS, THC, and DO in the Physico-chemical Parameters series. Most quality determinants in the sampled water are within allowable limits of the approved water quality standards including WHO except Alkalinity,  $\text{HCO}_3^-$ ,  $\text{MgCO}_3^-$ ,  $\text{Cl}^-$ , BOD,  $\text{Cd}^{2+}$ ,  $\text{Fe}^{3+}$  and  $\text{Zn}^{2+}$ . The overall ionic dominance pattern follows the same trend  $\text{K}^+ > \text{Na}^{2+} > \text{Fe}^{3+} > \text{Mn}^{2+}$  and  $\text{Cl}^- > \text{NO}_3^- > \text{NO}_2^-$ . The results of the microbial analyses revealed the highest Total Coliform Count of  $17 \times 10^2$  MPN/100 ml while that of THBC and FCC were reported as  $7.8 \times 10^2$  cfu/ml and  $1.2 \times 10^2$  cfu/ml respectively. The hygienic status of the water sources was found to be poor. High BOD, Coliform Count and  $\text{BOD}_5/\text{NO}_3^-$  ratios of 248.23 recorded for the groundwater samples were suggestive of possible organic pollution due to faecal invasion with the lowest measured depth of 35m. Multivariate statistical approaches namely Correlation, Principal Component Analysis and Descriptive statistics simplified and identified the interrelationship among the analyzed physicochemical parameters and the pollution sources. This study has therefore revealed that boreholes in Ewekoro were polluted and posed potential risk to biomedical safety and overall human health. Intervention measures are therefore necessary to safeguard the inhabitants from water-related diseases and their consequences.*

**KEYWORDS:** Borehole; Geochemical; Groundwater Quality; Descriptive Statistics; PCA; Nigeria.



## INTRODUCTION

Water is highly essential for the sustenance and maintenance of all forms of life, aside from daily domestic needs that it meets making its relevance to be ranked as one of the unavoidable natural endowments on earth. Water occupies approximately 71% of the earth surface naturally distributed in surface as seas, rivers, streams, lakes, and ponds and subsurface as groundwater (Pasquini and Alexander, 2004; Verplanck *et al.*, 2006; Ishola *et al.*, 2023). Despite its high state of abundance in different continents of the world, the quality as well as the availability of potable water still remains a daily challenge in the society; most commonly, in rural and semi-rural communities in the developing countries of the world (Faremi and Oloyede, 2010; Lashkaripour, 2003; UNEP/ERCE/UNESCO, 2008). Today, it is shocking and overwhelming that contaminated water has been reported to endanger more lives than AIDS, cancer, war or even accident (WHO, 2011a). Diarrhea diseases alone in the circle water borne diseases account for an estimated 4.1% of the total daily global challenge of diseases and are typically responsible for the deaths of over 1.8 million people every year (WHO, 2011a; AGWT, 2013); of which 88% of this burden is attributable to unsafe water consumption, poor sanitary conditions and hygienic environment among others (WHO, 2004). Microbial invasion of drinking water emanating from human faeces has been a major contributor to diarrhea diseases that claim millions of children and infants on a yearly basis (Foppen, 2002; Horward *et al.*, 2006; Verplanck *et al.*, 2006; UNEP *et al.*, 2008). Thus, it is highly recommended that for a given water to be safe for consumption, it must be completely free from all disease causing microbes and harmful chemicals that threaten public health and wellness. In Nigeria, it has been reported that only 58% of urban cum semi-urban residents and 39% of the rural residents have access to potable water supply while the rest of the population relies solely on groundwater and surface water for their domestic needs (FGN, 2012; Ishola *et al.*, 2023). As long as urbanization, industrial and domestic pollution, atmospheric input from fossil fuel burning and environmental degradation, daily human and other biological population increase, chemical and biological contaminants are expected to also increase alongsides in water which may lead to series of threats to available water in the affected environment. Earlier before now, several environmental investigations have revealed surface and subsurface groundwater quality invasions in Nigeria, Uganda, and India because of the chemical and biological inputs from external sources (Foppen, 2002; APHA, 1998; Kumar and Pal, 2010; Oluseyi *et al.*, 2011; Sha'Ato *et al.*, 2010; Ishola *et al.*, 2023). As these daily experiences continue to linger, the challenges of water quality status equally becomes much more arduous task. Therefore, there is an urgent need for regular water quality inspection, monitoring and overall management as it is enshrined in the health promotion strategy network in underdeveloped and developing countries of the world (Ishola *et al.*, 2021). Over two decades, multivariate statistical analyses (PCA and CA) have been applied for solving diverse hydraulic, hydro-geochemical and biological problems relating to water availability and consequent qualities for domestic and industrial use (Amadi *et al.*, 2012). With this technique, large datasets involving geological, hydrological and biological variables are simplified, organised and classified to produce beneficial and interpretable results (Wu *et al.*, 2005). Moreover, most subsurface water data once acquired from rural communities in Nigeria were quite inadequate and doubtful in their presentations; on these notes, this study is principally carried out for the purpose of evaluating the quality of subsurface conditions of groundwater using geochemical and microbiological variables exemplifying it on borehole water sources in Ewekoro communities within Ewekoro Local Government Area of Ogun State.



## STUDY AREA

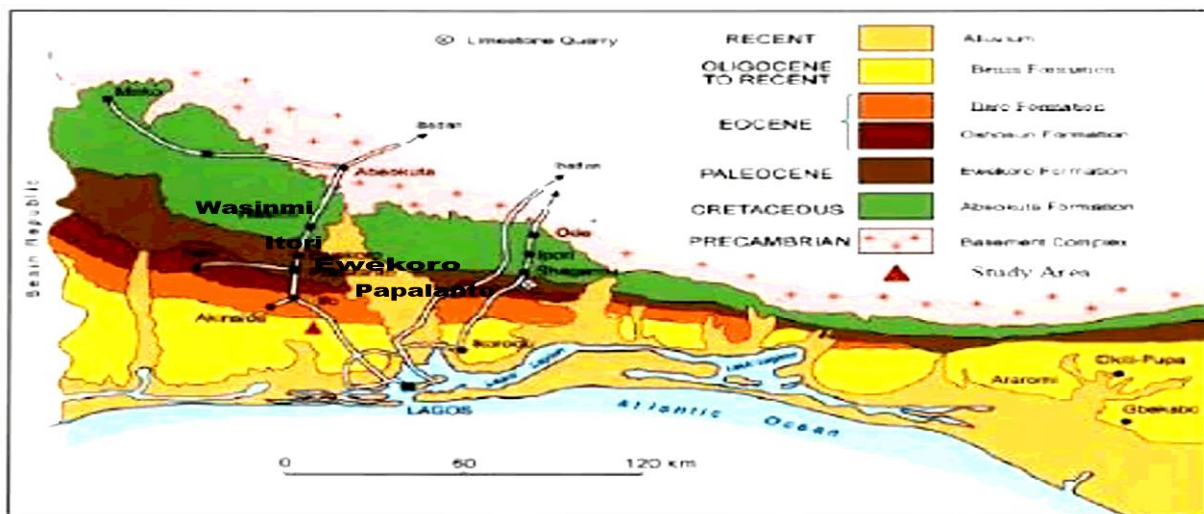
### Location, accessibility and Geologic Setting

Ewekoro community in Ogun State is one of the mills of West African Portland Cement Company (WAPCO) and Dangote group Cement Company. It is a sleepy neighbouring town to Papalanto, a name known for sugarcane plantation. It lies between a latitude of  $6^{\circ}53'N$  and a longitude of  $3^{\circ}14'E$  (Ishola, 2019). The sedimentary rocks of Ogun State consist of Ewekoro formation and Abeokuta formation. The Ewekoro formation is fossiliferous and consists of economic deposits of limestones that were quarried by WAPCO (Ishola *et al.*, 2021). Ewekoro cement production facility is located at 5 kilometres north of Ewekoro town ( $6^{\circ}55'N$  and  $3^{\circ}12'E$ ). Also, it is approximately 64 kilometers north of Lagos and 42 kilometers south of Abeokuta and within the tropical rainforest belt of Nigeria. Farming settlements such as Olapeleke (West), Itori (North), Elebute and Alaguntan (East) which predate the factory are located within 10 km radius of the production facility (Ishola, 2019). The settlements are perennially drained by Itori, Ewekoro, Eshe, Elebute and Alaguntan Rivers. Only Alaguntan River receives waste water directly from the cement plant. However, the water qualities of other catchment rivers are primarily influenced by non-point pollutants from run-offs and atmospheric depositions. Cement dust and other particulate depositions are facilitated by an average wind speed of 1.0 and 0.72 msG, respectively 10 m above the ground during the dry (January-March) and wet periods (May-November) (Ishola, 2019). Other significant and prevailing climatic conditions over the cement plant catchment area include an average relative humidity of  $65\pm 10\%$  and an average annual rainfall of  $1500\pm 120$  mm (Ishola, 2019). Ewekoro formation at the type locality is composed of 11m to 12m of limestones. It is sandy at the base grading downward into Abeokuta Formation. The Ewekoro formation is overlain by a phosphatic glauconitic grey shale (Obaje, 2009; Ishola, 2024). The Ewekoro formation is the local geology in the study area which is generally consistent with the regional geology of eastern part of the Dahomey Basin; predominantly comprises of the non-crystalline and highly non-fossiliferous limestone and thinly laminated fissile and probably non-fossiliferous shale (Ushie *et al.*, 2014). It is the sedimentary terrain of southwestern Nigeria. Ewekoro formation consists of intercalations of argillaceous sediment. The rock is soft and friable but in some places cemented by ferruginous and siliceous materials. The lithological units in Ewekoro formation are clayey sand, clay, shale, marl, limestone and sandstone (Obaje, 2009; Ishola *et al.*, 2021).

The study area is generally a low lying to gentle undulating terrain that falls within the humid tropical climate characterized by two distinct seasons predominant in the tropics in the southern part of Nigeria namely, the wet and dry seasons. The wet season usually occurs from March to October, the climate is dominated by the tropical maritime airmass or moisture laden Southwest winds from the Atlantic Ocean that produces heavy rainfall; most of the rainfall comes in torrential showers resulting in high run-off while the dry season occurs from November to late February or early March under the influence of the dry continental airmass or North-Easterly winds from Sahara desert. The major water bodies in the region are Yewa and Ogun rivers which flow into Lagos lagoon while their tributaries are found in Ewekoro Local Government Area as Alaguntan River, Akinbo River and Eshe River. There are however streams running parallel in the area. Also ponds are not left out. Due to the alternation of wet and dry seasons, the water table fluctuates in response to the seasonality of rainfall. During the wet season, groundwater level rises towards the surface and drops as the dry season sets in (Ishola, 2019).

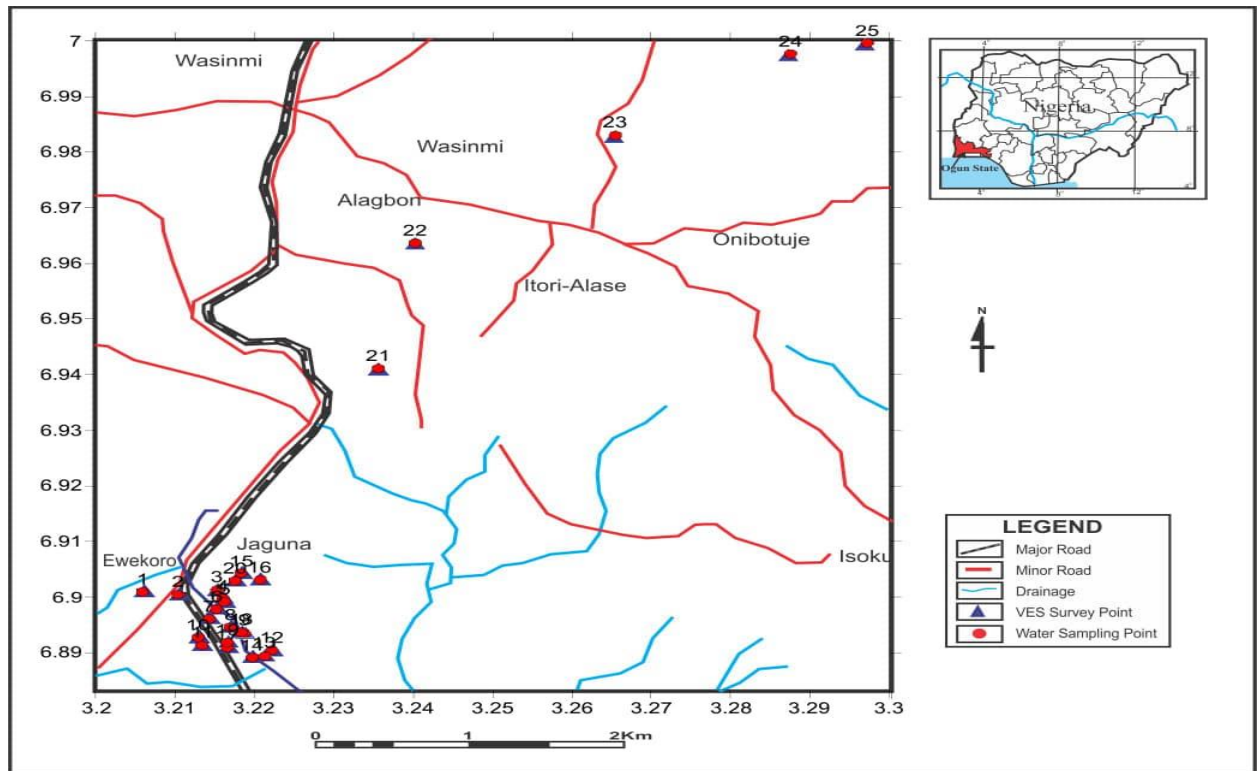
Fig. 1 shows the Geological Map of the Selected Locations of the Study Area within Dahomey Embayment, the inset map showing political divisions of the study area within Nigerian continental environment is shown in Fig. 2, the map of the investigated locations in the study area are shown in Fig. 3 and Fig. 4 is a base map showing the location and accessibility of the study area in Ewekoro LGA.

The entire study area is generally accessible by major roads and several footpaths, although the road from Abeokuta town to the investigated area is tarred. In addition to Ewekoro-Papalanto road, the survey locations can equally be accessed through a major road from Lagos State through Sango-Arigrbajo Ifo express road (Ishola, 2019).

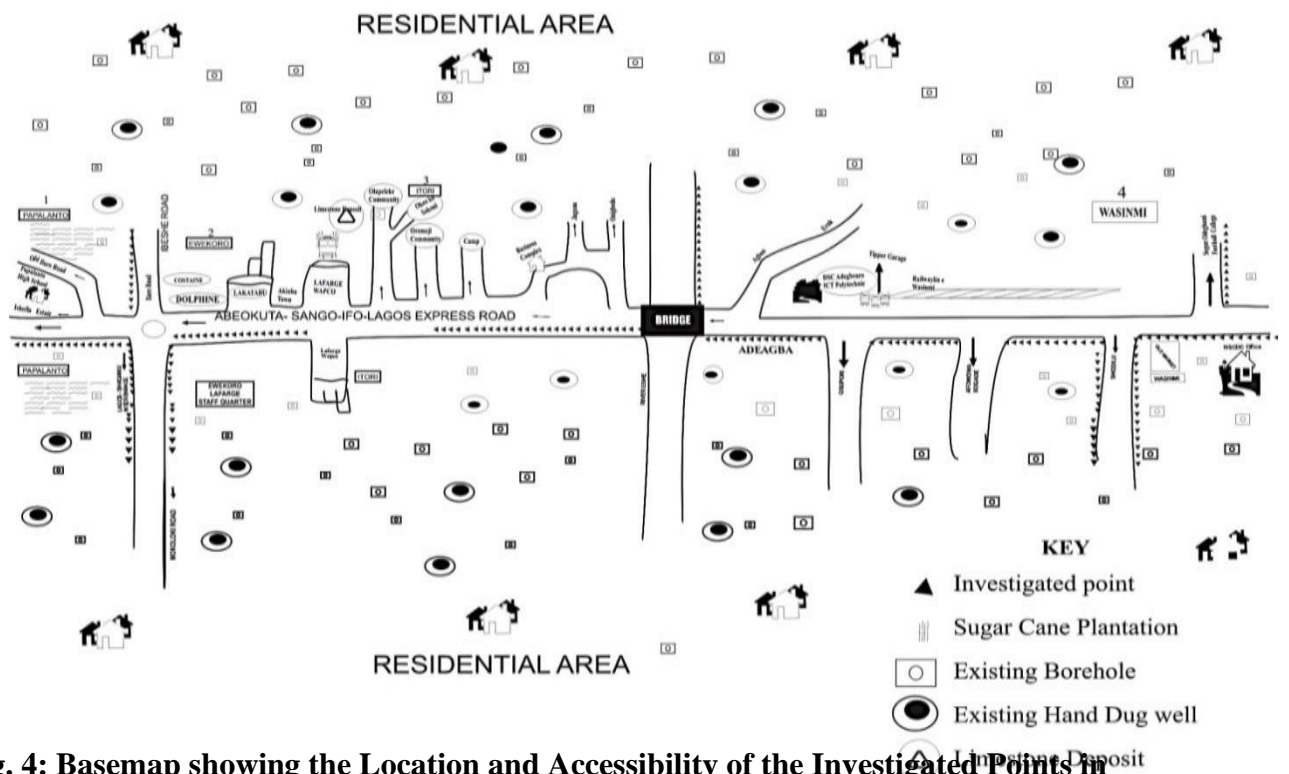


**Fig. 1: Geological Map Showing the Study Area within the Nigerian Part of Dahomey Embayment (after Billman, 1992; modified by Ishola, 2019).**





**Fig. 3: Data Acquisition Map showing the Investigated Locations in Ewekoro Study Area in Ewekoro LGA, Southwest Nigeria (Ishola, 2019).**



**Fig. 4: Basemap showing the Location and Accessibility of the Investigated Points in Ewekoro LGA, Southwest Nigeria (Ishola, 2019).**



## MATERIALS AND METHODS

### Collection and Analyses of Subsurface Water Samples

Different water samples were collected from 25 existing and functional borehole sites at separate and strategic sampling points within Ewekoro communities. Samples collected were immediately stored after collection in clean air-tight-proof plastic bottles and labeled appropriately while 1 ml concentrated  $\text{HNO}_3$  per liter of the sample was used for the preservation of the samples for heavy metals. The water samples were analyzed for the physico-chemical parameter to quantify the physical properties and chemical constituents present in the water namely pH, cations, anions, and trace elements among others. Electrical Conductivity (EC), Total Dissolved Solids (TDS), Temperature, Dissolved Oxygen (DO) and pH were determined in-situ using Hannah Combo TDS/pH/EC/Temperature meter series multi-parameters (model HI991300), whereas Hannah (model HI9147) equipment was used for daily measurement of DO to ensure that they were not subjected to physical alteration such as temperature as JYD-IA DO meter was used for the BOD5 calculation after the expiration of five days incubation. Aside from other analyzed physicochemical parameters, bacteriological assessment and heavy-metal levels were equally measured following the standard laboratory procedures (APHA, 1992; APHA, 1998; Ishola *et al.*, 2021). The geographical coordinates of sample points were also taken with a GPS meter, and their location is indicated on the data acquisition map (Fig. 4). Samples specifically meant for anion determination were collected in 500 ml bottles, unfiltered and unpreserved, and later stored below  $8^\circ\text{C}$  before analysis while the third sampling bottles were used for the determination of microbial loads. Ion Chromatography (IC) was used to determine the concentration levels of the anions, while nitrates, phosphates, bicarbonates, chlorides and sulphates were measured after chromatography separation utilizing conductivity detectors. Inductively Coupled Mass Spectrometer (ICP-MS) and Inductively Coupled Optical Emission Spectrometry (ICP-OES) were used for detection of heavy metals and trace metals. Water samples were filtered to less than  $0.45\ \mu\text{m}$  using a Pall Corporation GN-6 metricell sterilized membrane to improve the accuracy and to prevent cloudiness of the water while ensuring that the minute particles of clay sizes were removed before analysis. When lower levels of contamination were identified, ICP-MS provided lower detection limits for measurement. At the same time, ICP-OES was useful for higher concentrations, such as cases of high contamination levels (Udousoro and Umoren, 2014). Furthermore, a cell-based ICP-MS served as a very veritable integration tool for the removing possible interferences that might prevent the detection of a contamination at its emergence. The samples were further examined for microbial analyses under the microscope. The Heterotrophic total bacterial counts were determined by pour plate technique using standard analytical methods (APHA, 1998; Pepper *et al.*, 2004). Total Coliform Counts were equally determined by the Most probable number (MPN) index technique using 3-3-3 regimen for coliform and total coliform (Pepper *et al.*, 2004). Faecal coliform counts were determined using Eosin Methylene Blue medium employing the pour plate technique (APHA, 1992; Burnett and Beuchat, 2001). These were consequently compared to water quality standards as provided by the World Health Organization (WHO, 2011) and the National Agency for Food and Drug Administration and Control (NAFDAC, 2004). The accuracy of the methods utilized for all determinants was better than  $\pm 10\%$ , the bias is within  $\pm 3\%$ , and the repeatability at the 95% confidence interval was better than at 5%. The analyses were deliberately repeated to the correctness of the outcomes in relation to laboratory precision and accuracy. These analyses were ultimately



carried out to evaluate the concentration status of the constituting elements with respect to the possible activities in the study area. The depth measurements of the investigated boreholes were determined using the Heroin Dipper-T Water level meter. All the laboratory sample analyses were conducted in the Central Laboratory, Institute of Agricultural Research and Training (IART), Obafemi Awolowo University, Moor Plantation Ibadan Campus.

### Analyses of Water Sample Data

Water quality datasets were subjected to Multivariate analyses comprising basically Principal Component Analysis (PCA), Correlation Analysis, (CA) and other descriptive statistical techniques. CA was determined using the statistical software- Statgraphic® Centurion XV while SPSS Statistics 20.00 was used for Pearson's correlation coefficients analysis and PCA synthesis. Principal component analyses were performed on a matrix of 25 experimental data in groundwater (boreholes) standardized through a set of Varimax rotations with Kaiser Normalization to eliminate the interference of drifting units of measurement and to render the acquired data dimensionless (Ishola et al., 2023). Adherence to the analytical quality control was ensured through procedural blank measurements, duplicate analyses of water samples and standardization of analytical laboratory devices.

### Mathematical Principles of Principal Component Analysis

An  $n \times m$  matrix forms of the dataset when only we have  $n$  objects in relation to measured  $m$  characteristics. Every pair of available objects is subjected to the computations of certain measurement of similarity between them (Ishola *et al.*, 2023). Wide varieties of coefficients of resemblance have been used which includes the correlation coefficient  $r_{ij}$  and a standardized  $m$ -space Euclidean distance,  $d_{ij}$  while the distance coefficient is computed by

$$d_{ij} = \sqrt{\frac{\sum_{k=1}^m (X_{ik} - X_{jk})^2}{m}} \quad (\text{Ishola } et al., 2023) \quad 1.0$$

where  $X_{ik}$  is followed by the corresponding  $K_{th}$  variable measured on object  $i$  as  $X_{jk}$  is the  $K_{th}$  variable measured on object  $j$  with  $m$  variables as measured on each object, and  $d_{ij}$  is the distance between object  $i$  and object  $j$  (Ishola, 2019). A low distance depicts similarity in the two objects while a high distance indicates dissimilarity as often expected. As commonly observed, the  $n \times m$  raw data matrix is a standardized data prior to computing distance measurements; ensuring that each identified variable is weighted equally (Davis *et al.*, 1970; Gangopadhyay *et al.*, 2001).

Groundwater sample data were collected from  $b$  boreholes which were often subjected subsequent correlation. This correlation is a possible reflection of existing complexity of the aquifer hydrogeology as an indication of its constituents (elemental/geochemical concentrations and physicochemical compositions in this case study) collected from one borehole is also contained in the remaining  $b-1$  boreholes. Therefore, the objective of principal component analysis is to assess its concentration level based on the number of independent variables that characterize the elemental concentration cum physicochemical variations in the aquifer system. This goes a long way in reducing the number of variables measured or observed at a given instance. The reviewed aspects of this particular principal component technique are significant for the present application in this work as outlined in this section.





Let  $\mathbf{X}$  be assigned as the observation matrix of deviations from the mean of the order  $n \times p$ , where  $n$  serves as the number of observations on  $b$  variables. Here  $n$  is the number for which elemental concentration measurement of water samples have been considered for  $b$  boreholes. The mean concentration of the borehole is assumed to be the average of the  $n$  number of observation values for the boreholes. So in  $\mathbf{X}$  elemental concentrations for a given borehole represent the deviation from their respective mean concentration level. It is also considered that this original  $b$  variable set of observations in  $\mathbf{X}$  contains an identified correlation. This correlation can be analyzed using principal components to identify the relative relevance of any borehole in representing variations of elemental concentrations among the  $b$  wells. This is achieved based on the characterizations of the variance of  $\mathbf{X}$  with  $q$  ( $q < b$ ) principal or orthogonal components and subsequent evaluation of the correlation between the  $i$ th original variable and  $j$ th principal component. The  $j$ th principal component,  $\mathbf{z}_j$  ( $n \times 1$ ) column vector, is the linear function such that

$$\mathbf{z}_j = \mathbf{X}\mathbf{a}_j \quad (\text{Ishola et al., 2023}) \quad 1.1$$

where  $\mathbf{a}_j$  is a  $p \times 1$  linear transformation coefficient vector corresponding to  $\mathbf{z}_j$ .

From the basic theory of Principal Component Analysis (Gangopadhyah et al., 2001; Ishola et al., 2023), It can be revealed that  $\mathbf{a}_j$  is the characteristic vector associated with the characteristic roots  $\lambda_j$  (roots of the scalar  $\lambda$ ) which are obtained by solving the equation defined by

$$|\mathbf{S} - \lambda\mathbf{I}| = 0 \quad 1.2$$

Where  $\mathbf{I}$  and  $\mathbf{0}$  are the  $p \times 1$  unit and zero vectors respectively;

$\mathbf{S}$  (order  $b$ ) is the estimate of the variance-covariance matrix of  $\mathbf{X}$ ; it is given by

$$\mathbf{S} = \frac{\mathbf{X}^T\mathbf{X}}{(n-1)} \quad (\text{Ishola et al., 2023}) \quad 1.3$$

where the superscript  $T$  denotes the matrix-transpose operator.

The solution of equation (2.7) is subject to

$$\text{Normalizing Constraints:} \quad \mathbf{a}_j^T \mathbf{a}_j = 1; \text{ and}$$

$$\text{Orthogonality constraints:} \quad \mathbf{a}_i^T \mathbf{a}_j = \mathbf{a}_j^T \mathbf{a}_i = 0$$

These constraints ensure a unique solution and uncorrelated principal components (Gangopadhyah et al., 2001; Ishola, 2019).

Thus, corresponding to  $\lambda_1, \lambda_2, \lambda_3$  to  $\lambda_b$ , we have the  $n \times 1$  column vectors  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ , to  $\mathbf{a}_b$  respectively. Then, the above equations can be extended to

$$\mathbf{a}_i^T. \text{ Where } \mathbf{Z} = \mathbf{X} \quad 1.4$$

$\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$  to  $\mathbf{z}_b)$  is the transformed  $n \times p$  matrix of  $n$  values for each of the  $b$  components, and  $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ , to  $\mathbf{a}_b)$  is the  $b \times b$  linear transformation coefficient matrix (Gangopadhyah et al., 2001; Ishola, 2019; Ishola et al., 2023). In  $\mathbf{Z}$ , the first  $q$  components



explain most of the variation contained in  $\mathbf{X}$ , and the remaining  $p-q$  provides only a small contribution to the total variation and which may be neglected. The advantage at this stage is working with  $\mathbf{Z}$  in that these new variables (principal components) are mutually uncorrelated and the dimensionality of the original problem is consequently reduced (Davis *et al.*, 1970; Gangopadhyay *et al.*, 2001; Ishola, 2019; Ishola *et al.*, 2023). However, the objective is to utilize the correlation between  $\mathbf{X}_i$  and  $\mathbf{Z}_j$  to determining the significance of any sampled borehole  $\mathbf{b}_i$  in representing the elemental concentration variation among the group of  $\mathbf{b}$  boreholes in the same hydrogeological environment in this analysis. Therefore, the correlation between the  $i_{th}$  standardized observed variable and the  $j_{th}$  computed component can be expressed as

$$\cos \theta = (x_i, z_j) = \lambda_j^{1/2} a_{ij} \quad (\text{Ishola } et al., 2023) \quad 1.5$$

where  $a_{ij}$  ( $i$  = rows,  $j$  = columns) are the elements of the coefficients matrix  $\mathbf{A}$ .

Equation 1.0 can therefore be utilized in computing the elements of the correlation matrix that indicate the extent of correlation between the original observed variable and the derived principal component. This correlation matrix is referred to as the factor loading matrix as commonly observed while the elements given by equation 1.0 are called factor loadings.

The concept of varimax criterion involves the maximization of the variance of the aforementioned loadings on the factors. We may define the variance  $S^2_k$  of the loadings on the  $K_{th}$  factors as expressed in equation 1.6

$$S^2_k = \frac{p \sum_{j=1}^m (l^2 j p / h^2 j)^2 - (\sum_{j=1}^m (l^2 j p / h^2 j))^2}{p^2} \quad (\text{Ishola } et al., 2023) \quad 1.6$$

In order to enhance the interpretation of the factor loading matrix, an orthogonal transformation such as varimax rotation is often carried out (Gangopadhyay *et al.*, 2001; Ishola *et al.*, 2023) as diagrammatically exemplified and displayed in Fig. 5. Setting a cut-off value for factor loading which serves as a measure of significant correlation; boreholes that have loadings below the significance levels for all the extracted components may not be observed. In other words, only  $q$  ( $q < b$ ) boreholes that have a significant correlation with extracted component can adequately explain the significant contribution of the elemental concentration variation in the group of  $\mathbf{b}$  wells or boreholes as the cosine theta coefficient is defined as:

$$\cos \theta_{ij} = \frac{\sum_{k=1}^n X_{ik} Z_{jk}}{\sqrt{\sum_{k=1}^n X_{ik}^2 \sum_{k=1}^n Z_{jk}^2}} \quad (\text{Ishola } et al., 2023) \quad 1.7$$

where  $X_{ik}$  is the  $k^{th}$  variable of the  $i^{th}$  observation and  $n$  in equation 1.7 serves as the number of variables. If we regard the identified two samples,  $i$  and  $j$  as defining points in  $n$ -dimensional variable space, this measurement ascribes the cosine of the angle between the vectors to those two points. In a sense, cosine theta is a measure of the ‘‘great circle’’ distance between samples  $i$  and  $j$  lying on the surface of the hypersphere (Davis, 1986; Ishola, 2019; Ishola *et al.*, 2023).

Comparison of equation 1.6 with that derived for the correlation will display a close similarity between the two; if  $n$  variables used to compute cosine theta were standardized to



have zero mean and variance of 1.00, the two measures would be identical, because the cosine of zero degrees is 1.00. Decreasing similarity is indicated by lower values of coefficient, and vice-versa just as in the correlation (Davis *et al.*, 1970; Yeung, 1999; Ishola *et al.*, 2023).

Figure 5 shows the geometrical expression of the process in two dimensions. The mean values of the variables ( $\mu_{x1}, \mu_{x2}$ ) and the covariance matrix which is an  $n \times p$  ( $2 \times 2$ ) matrix were found in this case using all the available data points. If we compute the eigenvectors of the covariance matrix we get the direction vectors indicated by  $\phi_1$  and  $\phi_2$  while placing appropriately the two eigenvectors as columns in the matrix  $\Phi = [\phi_1, \phi_2]$ ; we can therefore create a transformation matrix which consequently transfers the data points from the previous  $[x_1, x_2]$  axis system to the axis  $[\phi_1, \phi_2]$  system with the equation:

$$p_\phi = (p_x - \mu_x) \cdot \Phi \quad 1.8$$

where  $p_x$  is any point in the  $[x_1, x_2]$  axis system,  $\mu_x = (\mu_{x1}, \mu_{x2})$  is the data mean, and  $p_\phi$  is the coordinate of the point in the  $[\phi_1, \phi_2]$  axis system (Ishola, 2019).

### Theoretical Background of PCA in the Hydrogeochemical Characterization of Aquifers

In consideration of the simplicity of its algebra and its straightforward approach cum interpretation, PCA stands as the most widely used method of multivariate data analysis (Ishola *et al.*, 2023). A linear transformation can be defined as simply transforming a set of variables from the state of correlated factors into uncorrelated factors. These transformed orthogonal factors can extract successively a maximal part of the local variance of the variables. The basic problem solved by the principal component analysis is to transform a set of correlated variables into uncorrelated quantities which could be interpreted in an ideal, multi-Gaussian context as independent factors underlying the phenomenon (Wackernagel, 1995; Ishola, 2019; Ishola *et al.*, 2023).

If  $Z$  be the  $n \times N$  matrix of the data from which the means of the variables have already been earlier extracted. Then the corresponding  $N \times N$  variance-covariance matrix  $V$  is therefore given as:

$$V = [\sigma_{ij}] = \frac{1}{n} Z^T Z \quad (\text{Yeung, 1999; Ishola } et al., 2023) \quad 1.9$$

Let  $Y$  be an  $n \times N$  matrix containing in its rows the  $n$  samples of the factors  $Y_p$  ( $p = 1, \dots, N$ ) which are uncorrelated and of zero mean (Yeung, 1999; Ishola, 2019; Ishola *et al.*, 2023).

The variance-covariance matrix of the factors is diagonal because the covariances between the factors are nil by definition, hence

$$D = \frac{1}{n} Y^T Y = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_{NN} \end{bmatrix} \quad (\text{Yeung, 1999; Ishola, 2019; Ishola } et al., 2023). \quad 1.10$$

and the diagonal elements  $d_{pp}$  are the variances of the factors.



An  $N \times N$  orthogonal matrix  $A_2$  which linearly transforms the measured variables into synthetic factors is given by

$$Y = ZA \text{ with } A^T A = \mathbf{1} \quad (\text{Yeung, 1999; Ishola, 2019; Ishola } et al., 2023) \quad 1.11$$

Finding the product of the equation from the left with  $\frac{1}{n}$  and  $Y^T$  with the replacement of  $Y$  by  $ZA$ , the following is obtained

$$\frac{1}{n} Y^T Y = \frac{1}{n} Y^T Z A = \frac{1}{n} (Z A)^T (Z A) = \frac{1}{n} A^T Z^T Z A = A^T \frac{1}{n} (Z^T Z) A \quad 1.12$$

Finally

$$D = A^T V A \Rightarrow V A = A D \quad (\text{Yeung, 1999; Ishola, 2019; Ishola } et al., 2023) \quad 1.13$$

Rendering the Principal component as nothing other than a statistical interpretation of the eigenvalue problem as shown in equation 2.13 and 2.15:

$$V Q = Q A \text{ with } Q^T Q = \mathbf{1} \quad (\text{Wackernagel, 1995; Ishola } et al., 2023)$$

This can be inferred immediately that the matrix  $Q$  orthonormal of eigenvectors of  $V$  offers a solution to the problem and the eigenvalues  $\lambda_p$  are then simply the variances of the factors  $Y_p$ .

Defining the factors as:

$$Y = Z Q \quad 1.14$$

Principal component analysis is also very useful for defining a sequence of orthogonal factors which successively adopt a maximum amount of variance of the data (Gangopadhyah *et al.*, 2001; Ishola, 2019; Ishola *et al.*, 2023). For instance, take a vector  $y_1$  corresponding to the first factor obtained by transforming the centred data matrix  $Z$  with a vector  $a_1$  calculated to a unit length:

$$y_1 = Z a_1 \text{ with } a_1^T a_1 = 1 \quad (\text{Gangopadhyah } et al., 2001). \quad 1.15$$

The variance of  $y_1$  is

$$\text{Var}(y_1) = \frac{1}{n} y_1^T y_1 = \frac{1}{n} a_1^T Z^T Z a_1 = a_1^T V a_1 \quad (\text{Gangopadhyah } et al., 2001). \quad 1.16$$

If we are to appropriate a maximal part of the acquired variance of the data to  $y_1$ , an objective function  $\Phi_1$  with a defined Lagrange parameter  $\lambda_1$ , which multiplies the constraints that the transformation vector  $a_1$  should be of unit norm as shown in equation 1.17.

$$\Phi_1 = a_1^T V a_1 - \lambda_1 (a_1^T a_1 - 1) \quad (\text{Wackernagel, 1995}) \quad 1.17$$

Also, when the derivative with respect to  $a_1$  is set at zero:

$$\frac{\delta \Phi_1}{\delta a_1} = 0 \Rightarrow 2 V a_1 - 2 \lambda_1 a_1 = 0 \quad (\text{Wackernagel, 1995}) \quad 1.18$$

$\lambda_1$  is an eigenvalue of the variance-covariance matrix and the  $\mathbf{a}_1$  is equal to the eigenvector  $\mathbf{q}_1$  associated with these eigenvalue:

$$V\mathbf{q}_1 = \lambda_1\mathbf{q}_1 \quad (\text{Gangopadhyah } et al., 2001; \text{Ishola } et al., 2023). \quad 1.19$$

Equation 2.20 shows a second vector  $\mathbf{y}_2$  orthogonal to the first  $\mathbf{y}_1$ :

$$Cov(\mathbf{y}_2, \mathbf{y}_1) = Cov(\mathbf{Za}_2, \mathbf{Za}_1) = \mathbf{a}_2^T V \mathbf{a}_1 = \mathbf{a}_2^T \lambda_1 \mathbf{a}_1 = 0 \quad 1.20$$

The function  $\Phi_2$  maximize and incorporates two constraints: the fact that  $\mathbf{a}_2$  should be unit norm and the orthogonally between  $\mathbf{a}_2$  and  $\mathbf{a}_1$ . These constraints produce two new langrange multipliers  $\lambda_2$  and  $\mu$ :

$$\Phi_2 = \mathbf{a}_2^T V \mathbf{a}_2 - \lambda_2 (\mathbf{a}_2^T \mathbf{a}_2 - 1) + \mu \mathbf{a}_2^T \mathbf{a}_1 = 0 \quad 1.21$$

Setting the derivative with respect to  $\mathbf{a}_2$ :

$$\frac{\delta \Phi_2}{\delta \mathbf{a}_2} = 0 \Rightarrow 2V\mathbf{a}_2 - 2\lambda_2 \mathbf{a}_2 + \mu \mathbf{a}_1 = 0 \quad 1.22$$

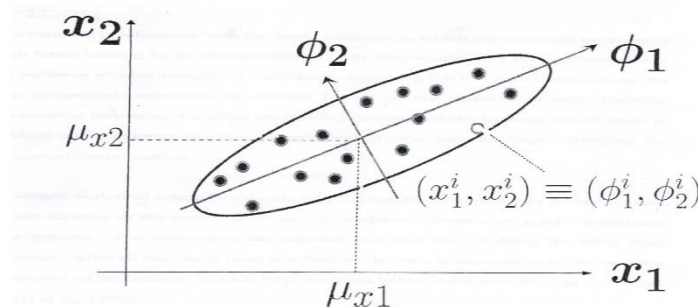
$$\underbrace{2\mathbf{a}_1^T V \mathbf{a}_2}_{0} = \underbrace{2\lambda_2 \mathbf{a}_1^T \mathbf{a}_2}_{0} + \underbrace{\mu \mathbf{a}_1^T \mathbf{a}_1}_{1} = 0 \quad (\text{Ishola, 2019; Ishola } et al., 2023). \quad 1.23$$

$$0 \qquad 0 \qquad 1$$

It can be seen that  $\mu$  is null (the constraint is not active). Therefore,

$$V\mathbf{a}_2 = \lambda_2 \mathbf{a}_2 \quad (\text{Gangopadhyah } et al., 2001; \text{Ishola, 2019; Ishola } et al., 2023). \quad 1.24$$

Again  $\lambda_2$  is later turned out to be an eigenvalue of the variance-covariance matrix and  $\mathbf{a}_2$  in the corresponding eigenvectors  $\mathbf{q}_2$  continuing similarly. One can find the rest of the  $N$  eigenvalues and eigenvectors of  $V$  as a solution to the maximization problem (Francisco *et al.*, 2001). The components are subsequently rotated orthogonally (Fig. 5) using the varimax method as an analytical tool to obtain a more significant distribution of the weights of the different variables on the components (Davis, 1986; Ishola *et al.*, 2023). These multivariate techniques are very useful tools for the hydrogeochemical characterization of aquifers (Francisco *et al.*, 2001) and for consequent studies of contamination of groundwater (Davis *et al.*, 1970; Wackernagel, 1995; Francisco *et al.*, 2001; Mahannada *et al.*, 2010; Gangopadhyah *et al.*, 2001; Ishola, 2019; Ishola *et al.*, 2023).



**Fig. 5: PCA Orthogonal Transformation (Ishola, 2019)**



## RESULTS AND DISCUSSION

### Physico-chemical Quality Characteristics

Thirty-eight (38) geochemical parameters (25 Elemental and 13 Physico-chemical parameters) and microbiological variables of the different groundwater (boreholes) samples collected in the study area were investigated. The results obtained are hereby presented jointly in Table 1 and 2.

The pH levels of the water sources range from 6.60 to 7.67 with a mean of  $6.97 \pm 0.32$ . The temperatures of the water bodies range from 24 to 28°C with a mean value of  $26.00 \pm 1.44$ . The colour values (in Hazen unit) of water samples range from 5 to 10 while the water turbidity ranges from 0.11 to 0.77 NTU. Electrical conductivity is a good measure of dissolved solid content in water; it is an important parameter in assessing the suitability of any given water source for irrigation purposes (Mahananda et al., 2001; Kumar and Pal, 2012). The values of Electrical Conductivity for the investigated water sources range from 558 to 941  $\mu\text{s}/\text{cm}$  with a recorded mean of  $720.36 \pm 93.68$  across the study area. Similarly, the total dissolved solids (TDS) range from 0.28 to 9.70 mg/L with a mean of  $4.99 \pm 3.89$  while the Total Suspended Solids (TSS) range from 0.23 to 4.95 mg/L with a mean of  $0.85 \pm 1.04$ . Dissolved Oxygen (DO) range of 6.70 to 21.48 mg/L with a mean of  $10.30 \pm 5.69$  while the BOD levels in the boreholes range from 16.92 to 32.27 mg/L with a mean of  $22.26 \pm 4.82$  while Total hardness (TH) for water samples ranges from 14.00 to 32.63 mg/L with a mean of  $25.74 \pm 4.81$ . The alkalinity levels of boreholes in the study area are very high and range from 28.42 to 2210.10 mg/L with a mean value of  $299.12 \pm 407.89$ ; the Standard deviation was higher than the mean value as observed because there was a wider distribution of Alkalinity contribution than the average contribution of other associated physico-chemical parameters in the groundwater system of the study area (Ishola *et al.*, 2016). The levels of COD obtained from borehole water sources range from 8.96 to 34.92 mg/L with a mean value of  $28.06 \pm 7.58$ . Ammonium concentration levels in samples from the study area range from 0.58 to 1.79 mg/L in Ewekoro boreholes (Table 1 and 2).

### Anion and Heavy Metal Concentrations

The anions levels in the different borehole sources are hereby presented as follows: Chloride concentrations ranged from 57.20 to 379.32 mg/L with a mean of  $234.57 \pm 126.15$  while Nitrate concentration levels ranged from 0.11 to 0.13 mg/L with a mean of  $0.13 \pm 0.006$  and Nitrite was in the range 0.01 to 0.04 mg/L with a mean of  $0.02 \pm 0.008$ . Manganese concentration levels ranged from 0.01 mg/L to 0.07 mg/L with a mean of  $0.04 \pm 0.002$  while Potassium, the range is 31.75 to 61.46 mg/L with a mean of  $55.30 \pm 7.66$ . Sodium concentration values are in the range of 27.99 to 49.98 mg/L with a mean of  $43.45 \pm 5.80$ . Zinc concentrations ranged from 0.67 to 1.68 mg/L with an average value of  $1.41 \pm 0.32$  while the range for Iron was 0.02 to 1.35 mg/L with a mean value of  $0.65 \pm 0.46$ . The range of the bicarbonates varied from 71.70 to 1108.75, with a mean value of  $408.22 \pm 445.40$ . It was observed from the results that the contribution and concentration of BOD, Alkalinity,  $\text{Cl}^-$ ,  $\text{Fe}^{3+}$ ,  $\text{Cd}^{2+}$ ,  $\text{Zn}^{2+}$ , Cd, and  $\text{HCO}_3^-$  and  $\text{MgCO}_3$  were greater than the approved set standards as revealed in Table 1.



## Total Bacterial Count

Borehole water sources from different sampling stations have considerably elevated levels of coliform counts. The THBC ranged from 0.00 to  $7.80 \times 10^2$  cfu/ml while the TCC ranges from 0 to 17 MPN/100mL and the FCC ranges from 0 to  $1.2 \times 10^2$  cfu/ml (Table 2). Considering the highest concentration level, the mean BOD: NO<sub>3</sub> ratio exhibited a value of 248.23; this served as a very significant index for estimating organic pollution levels in a given water source (Ishola *et al.*, 2016). the investigated water sources has the lowest measured depth of 35m in EWEBH4 and highest measured depth of 100m in EWEBH18 (Table 1 and 2).

**Table 1: Descriptive Statistics showing the Concentration Values of Physico-Chemical and Elemental Parameters of Ewekoro Boreholes (N=25)**

Parameters	Min	Max	Range	Mean±SD	WHO (mg/L)	NESRE A (mg/L)	NSDW Q (mg/L)	USEPA (mg/L)	NAFDA C (mg/L)
pH	6.60	7.67	1.07	6.97±0.32	6.5 – 9.5	7.00-8.50	6.50-8.50	6.50-8.50	6.50-8.50
TEMP(°C)	24.00	28.00	4.00	26.00±1.44	27	NA	NA	27	27
EC( $\mu S cm^{-1}$ )	558.0	941.00	383.00	720.36±93.68	1200	NA	900	1200	1000
DO (mg/L)	6.70	21.48	14.78	10.30±5.69	7.5	NA	7.5	NA	NA
BOD (mg/L)	16.92	32.27	15.35	<b>22.26±4.82</b>	10	NA	10	NA	NA
COD (mg/L)	8.96	34.92	25.96	28.06±7.58	NA	NA	NA	NA	NA
TDS (mg/L)	0.28	9.70	9.42	4.99±3.89	100	1500	500	500	500
TSS (mg/L)	0.23	4.95	4.72	0.85±1.04	> 10	>10	NA	NA	NA
TS (mg/L)	0.35	1.35	1.00	0.90±0.27	1500	NA	NA	NA	NA
TURB (NTU)	0.11	0.77	0.65	0.32±0.25	< 4	5.0	5.0	5.0	5.0
ALK (mg/L)	28.42	2210.10	2181.68	<b>299.12±407.89</b>	200	500	100	100	100
TH (mg/L)	14.00	32.63	18.63	25.74±4.81	< 200	100 – 300	500	NA	100
THC (mg/L)	0.01	0.28	0.28	0.21±0.07	NA	NA	NA	NA	NA
Na <sup>2+</sup> (mg/L)	27.99	48.98	20.99	43.45±5.80	< 200	NA	200	NA	200
K <sup>+</sup> (mg/L)	31.75	61.46	29.71	55.30±7.66	250	200	NA	200	10
Ca <sup>2+</sup> (mg/L)	10.60	29.84	19.24	24.58±6.42	100	75	NA	75	75
Mg <sup>3+</sup> (mg/L)	2.80	4.96	2.16	3.77±0.84	20	15	NA	20	20
Cl <sup>-</sup> (mg/L)	57.20	379.32	322.12	<b>234.57±126.15</b>	250	200	250	100	100
NO <sub>3</sub> <sup>-</sup> (mg/L)	0.11	0.13	0.02	0.13±0.006	50	45	NA	10	10
NO <sub>2</sub> <sup>-</sup> (mg/L)	0.01	0.04	0.02	0.02±0.008	< 3.0	NA	NA	NA	NA
SO <sub>4</sub> <sup>2-</sup> (mg/L)	7.70	19.20	11.50	11.70±2.84	400	500	200	250	100
NH <sub>4</sub> <sup>+</sup> (mg/L)	0.58	1.79	1.21	1.43±0.36	1.50	NA	NA	NA	NA
PO <sub>4</sub> <sup>3-</sup> (mg/L)	7.84	11.70	3.86	10.13±1.12	NA	NA	NA	NA	NA
HCO <sub>3</sub> <sup>-</sup>	71.70	1108.7	1037.0	<b>408.22±445.2</b>	100	NA	NA	NA	NA



(mg/L)		5	5	<b>.40</b>					
MgCO <sub>3</sub> (mg/L)	8.94	13.92	4.98	<b>12.10±1.50</b>	10	NA	NA	NA	NA
Cu <sup>2+</sup> (mg/L)	0.01	0.05	0.04	0.03±0.01	2.0	NA	1.0	1.3	1.0
Pb <sup>2+</sup> (mg/L)	0.00	0.01	0.01	0.001±0.002	0.01	0.01	0.01	0.01	0.01
Cd <sup>2+</sup> (mg/L)	0.00	0.04	0.04	<b>0.004±0.01</b>	0.003	0.003	0.001	0.005	0.005
Mn <sup>2+</sup> (mg/L)	0.01	0.07	0.06	0.04±0.02	0.1	0.2	0.5	0.4	2.0
Zn <sup>2+</sup> (mg/L)	0.67	1.68	1.01	<b>1.41±0.32</b>	0.01	NA	NA	NA	NA
Fe <sup>3+</sup> (mg/L)	0.02	1.35	1.33	<b>0.65±0.46</b>	0.3	0.3	0.3	0.3	0.3
Ni (mg/L)	0.00	0.03	0.03	0.007±0.01	0.02	0.05	NA	NA	0.05
S (mg/L)	0.16	5.50	5.34	3.16±2.05	250	NA	NA	NA	NA
Al <sup>3+</sup> (mg/L)	0.00	0.01	0.01	0.003±0.005	0.2	NA	NA	0.2	0.5
I (mg/L)	0.02	0.06	0.04	0.04±0.01	NA	NA	NA	NA	NA
Si (mg/L)	0.00	0.01	0.01	0.002±0.004	NA	NA	NA	NA	NA

## KEY

NA = Not Available in the National Guidelines and Standards for Drinking Water Quality

WHO = World Health Organization

NESREA = National Environmental Standard Regulation Agency

USEPA = United State Environmental Protection Agency

NSDWQ = National Standards for Drinking Water Quality

NAFDAC = National Food and Drug Administration and Control

**Table 2: Bacteriological Counts of Borehole Water Sample in Ewekoro, Southwest Nigeria**

SAMPLE CODE	BHD (m)	DO (mg/L)	BOD (mg/L)	COLOUR (TCU)	THBC (×10 <sup>2</sup> cfu/ml)	TCC (MPN/100ml)	FCC (×10 <sup>2</sup> cfu/ml)
EWEBH1	52	8.10	20.68	5	5.5	18	0
EWEBH2	38	8.21	20.74	5	7.2	18	0
EWEBH3	40	6.70	21.07	10	6.3	18	0
EWEBH4	83	6.80	21.10	5	7.8	17	0
EWEBH5	40	6.80	21.17	5	7.4	18	1.2
EWEBH6	45	6.82	21.20	5	6.3	18	1.2
EWEBH7	75	7.90	26.90	5	4.3	17	0
EWEBH8	80	7.96	27.10	5	3.8	15	0
EWEBH9	95	7.40	18.60	5	4.3	15	0.2
EWEBH10	60	7.42	18.62	5	3.9	14	0
EWEBH11	40	7.52	20.95	10	6.5	14	0
EWEBH12	45	7.48	20.79	10	6.3	14	0
EWEBH13	40	21.40	32.20	5	6.3	13	0
EWEBH14	35	21.46	32.24	5	0	0	0
EWEBH15	40	21.48	32.27	5	5.9	2	0
EWEBH16	72	7.84	16.96	5	2.8	2	0
EWEBH17	70	7.82	16.92	6	4.8	14	0
EWEBH18	100	8.32	18.25	5	3.7	14	0
EWEBH19	65	8.25	18.21	6	6.3	15	0





EWEBH20	65	7.24	18.70	5	4.8	13	0
EWEBH21	76	7.20	18.50	6	6.8	13	1.2
EWEBH22	62	7.32	19.52	6	3.3	15	0.9
EWEBH23	65	7.28	19.75	5	0	0	0
EWEBH24	45	21.43	26.58	5	0	0	0
EWEBH25	80	21.28	27.33	5	0.4	9	0
				<b>WHO</b>	<b>1.0 x 10<sup>2</sup></b>	<b>0</b>	<b>0</b>
				<b>(1996)</b>			
				<b>USEPA</b>			
				<b>(2002)</b>			
				<b>Guidelines</b>			

Key:

EWEBH= Ewekoro Borehole Water

DO = Dissolved Oxygen

BOD= Biological Oxygen Demand

THBC = Total Heterotrophic Bacteria Counts

TCC = Total Coliform Counts

FCC = Total Faecal Coliform Counts

BHD = Borehole Depth

### Principal Component Analysis (PCA) of Wasinmi Groundwater

The outcomes of PCA analysis as obtained from the processed principal components is sometimes not readily interpreted as varimax rotation with Kaiser normalization as it is executed with the primary aim of reducing the dimensionality of the dataset, extracting the most significant variables thereby prescribing the possible principal processes that control chemical concentrations in water sources (Saima *et al.*, 2009; Ishola *et al.*, 2023). The coefficient of Varimax factor loading with a correlation of  $> 0.75$  are interpreted as strong significant factor loading (FL);  $0.75-0.50$  are considered as moderate FL; and  $0.50-0.30$  are considered as weak FL. Only FL  $> 0.75$  which was adjudged as being positive or negative, was later used in this study to explain the contamination sources of the investigated boreholes (Rail, 2000; Liu *et al.*, 2003; Ishola *et al.*, 2023). The multivariate statistics of the rotated component matrix revealed the extracted components for both Physico-chemical and Elemental parameters with their proportionate eigenvalues  $>1$ . The Principal Component Analysis (PCA) results for Elemental Parameter and Physico-chemical variables using the principal component analytical method of extraction for rotation sums of squared loadings were presented in Table 5 to Table 6. To determine the number of components to retain in the results, the value of variance after rotation was examined where only components with Eigenvalue  $\geq 1$  after a complete rotational process were considered significant in the long run.

The rotation for the PCA analysis of Ewekoro geochemical parameters converged in 11 iterations with 6 components extracted which accounted for 92.76% of the total variance observed. The first components with eigenvalue of 8.86 accounted for 36.92% of the total variance observed which were significantly contributed by  $\text{Na}^{2+}$ ,  $\text{K}^{+}$ ,  $\text{Ca}^{2+}$ ,  $\text{SO}_4^{2-}$ ,  $\text{NH}_4^{+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Zn}^{2+}$ , Ni, and S. The second components revealed 18.80% as the total variance observed with eigenvalue of 4.51 displaying a high correlation for  $\text{Mg}^{3+}$ ,  $\text{SO}_4^{2-}$ ,  $\text{Cu}^{2+}$ ,  $\text{Cd}^{2+}$ , I and Si. The third components of eigenvalue of 3.176 represent 13.23% of the total variance that was



chiefly contributed by  $\text{Ca}^{2+}$ ,  $\text{NO}_3^-$ ,  $\text{Fe}^{3+}$ ,  $\text{MgCO}_3^-$  and  $\text{Cd}^{2+}$ . The fourth component possesses an eigenvalue of 2.661 and represents 11.09% of the total variance observed in the formation which was strongly and positively contributed by only  $\text{NO}_3^-$ . The fifth component accounted for 6.80% of the total variance observed with an eigenvalue of 1.63 and were significantly contributed by  $\text{Mg}^{2+}$ ,  $\text{NO}_2^-$ ,  $\text{NH}_4^+$ ,  $\text{Pb}^{2+}$  and  $\text{Fe}^{3+}$ . The sixth components accounted for 5.93% with a presumably lower eigenvalue of 1.42 when compared with earlier components and recorded a total variance observed as 5.93% with significant contributions from  $\text{Cl}^-$  and  $\text{PO}_4^{3-}$ . The result is well represented by the principal component scores of box data plotted on their respective positions on the components in rotated spaces in Fig. 6; supported by the dominance of  $\text{HCO}_3^-$  among the water quality determinants displayed in the bubble plot (Fig. 8).

The PCA analysis results for Physico-chemical Parameters of Ewekoro boreholes showed that the rotation converged in 8 iterations and 6 components were extracted which accounted for 90.99% of the total variance observed. The first component with eigenvalue of 3.46 accounted for 24.74% of the total variance observed with high factor loadings for DO, BOD and TH. The second component with eigenvalue of 2.78 accounted for 19.99% of the total variance observed which was significantly contributed by pH and TEMP. The third component with eigenvalue of 2.44 accounted for 17.43% and significantly contributed by high factor loadings for TDS and TS. The fourth component with eigenvalue of 1.76 accounted for 12.57% of the total variance and significantly contributed by high factor loadings for EC and TURB. The fifth component with eigenvalue of 1.22 accounted for 8.68% of the total variance with significant contributions from

TSS and THC. The sixth component possessed an eigenvalue of 1.06 which accounted for 7.58% of the exhibited total variance as shown in Table 4.

The degree of relationship of the tested Physico-chemical Parameters is displayed by the principal component scores of box data plotted on their respective positions on their components in rotated space in Fig. 7.

**Table 3: PCA for Geochemical Parameters of Ewekoro Borehole**

	Rotated Component Matrix <sup>a</sup> (Component)					
	C1	C2	C3	C4	C5	C6
$\text{Na}^{2+}$	<b>0.902</b>	0.168	0.334	0.107	-0.112	-0.080
$\text{K}^+$	<b>0.825</b>	0.200	0.231	-0.101	0.076	-0.276
$\text{Ca}^{2+}$	<b>0.756</b>	0.036	<b>0.467</b>	0.119	0.190	0.338
$\text{Mg}^{2+}$	-0.456	<b>0.572</b>	-0.170	0.141	<b>0.526</b>	0.173
$\text{Cl}^-$	-0.158	-0.577	-0.049	0.045	0.046	<b>0.786</b>
$\text{NO}_3^-$	0.368	-0.061	<b>0.859</b>	0.046	-0.013	-0.159
$\text{NO}_2^-$	0.057	0.277	0.098	0.226	<b>0.865</b>	-0.040
$\text{SO}_4^{2-}$	0.326	<b>0.571</b>	0.726	-0.012	-0.010	0.037
$\text{NH}_4^+$	<b>0.526</b>	0.086	<b>0.515</b>	-0.147	<b>0.556</b>	0.259
$\text{PO}_4^{3-}$	0.358	-0.044	0.403	0.038	0.098	<b>0.819</b>
$\text{HCO}_3^{2-}$	0.101	0.144	0.383	<b>0.713</b>	0.220	0.269

<b>MgCO<sub>3</sub><sup>-</sup></b>	0.373	-0.104	<b>0.831</b>	0.212	-0.086	0.266
<b>Cu<sup>2+</sup></b>	<b>0.512</b>	<b>0.725</b>	0.009	0.394	-0.105	0.010
<b>Pb<sup>2+</sup></b>	0.150	-0.273	0.329	-0.104	<b>0.529</b>	-0.608
<b>Cd<sup>2+</sup></b>	-0.055	<b>0.810</b>	<b>0.495</b>	0.201	-0.054	0.091
<b>Mn<sup>2+</sup></b>	0.437	0.279	0.052	<b>0.791</b>	-0.174	0.214
<b>Zn<sup>2+</sup></b>	<b>0.838</b>	0.256	0.123	0.277	0.082	0.285
<b>Fe<sup>3+</sup></b>	-0.062	-0.352	-0.138	-0.146	<b>0.863</b>	0.046
<b>Ni</b>	<b>0.596</b>	-0.357	0.225	0.104	-0.553	0.097
<b>S</b>	<b>0.683</b>	0.121	0.369	<b>0.462</b>	-0.388	0.074
<b>Al<sup>3+</sup></b>	0.270	0.373	-0.113	<b>0.836</b>	0.028	0.028
<b>I</b>	0.321	<b>0.819</b>	0.060	0.059	0.139	-0.190
<b>Si</b>	0.053	<b>0.904</b>	-0.229	0.317	-0.031	-0.097
<b>Eigenvalue</b>	<b>8.861</b>	<b>4.512</b>	<b>3.176</b>	<b>2.661</b>	<b>1.631</b>	<b>1.423</b>
<b>Variance (%)</b>	<b>36.920</b>	<b>18.798</b>	<b>13.232</b>	<b>11.086</b>	<b>6.797</b>	<b>5.929</b>
<b>Cumulative(%)</b>	<b>36.920</b>	<b>55.718</b>	<b>68.950</b>	<b>80.036</b>	<b>86.833</b>	<b>92.762</b>

Component Plot in Rotated Space

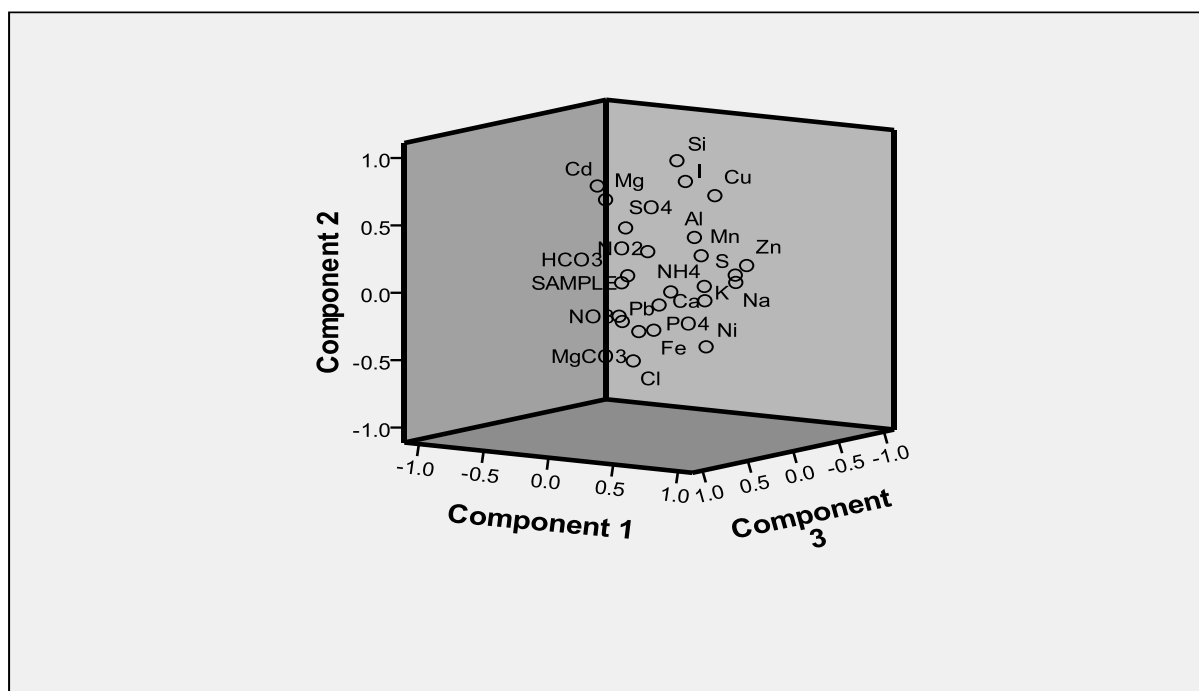
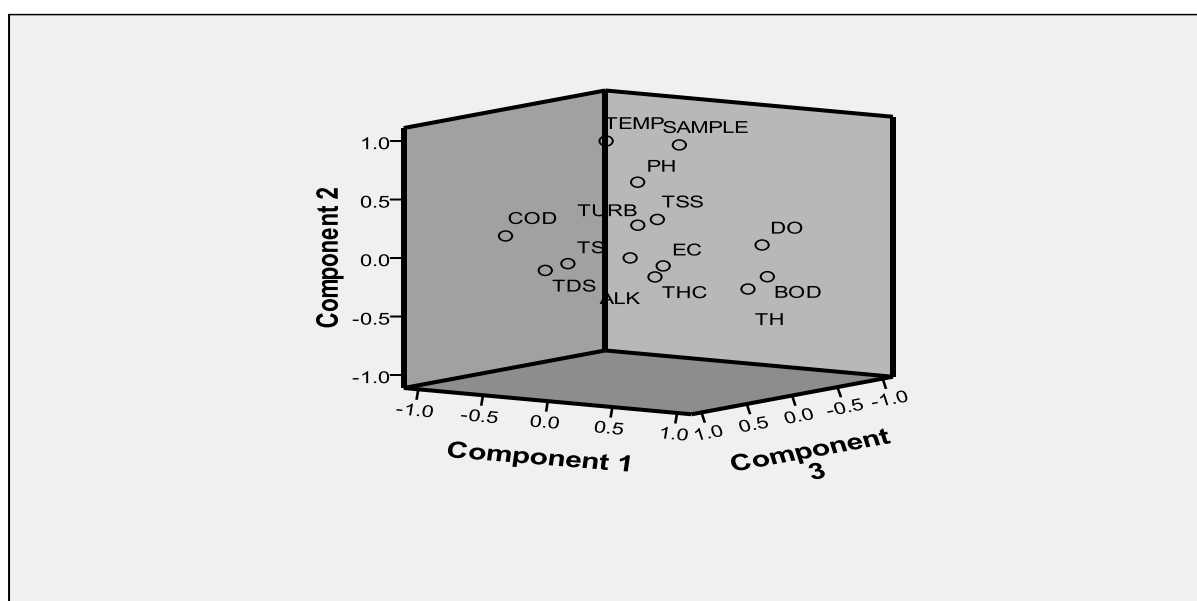


Fig. 6: Plot of Extracted Components 1, 2 and 3 of Elemental Parameters of Ewekoro Boreholes in Rotated Space

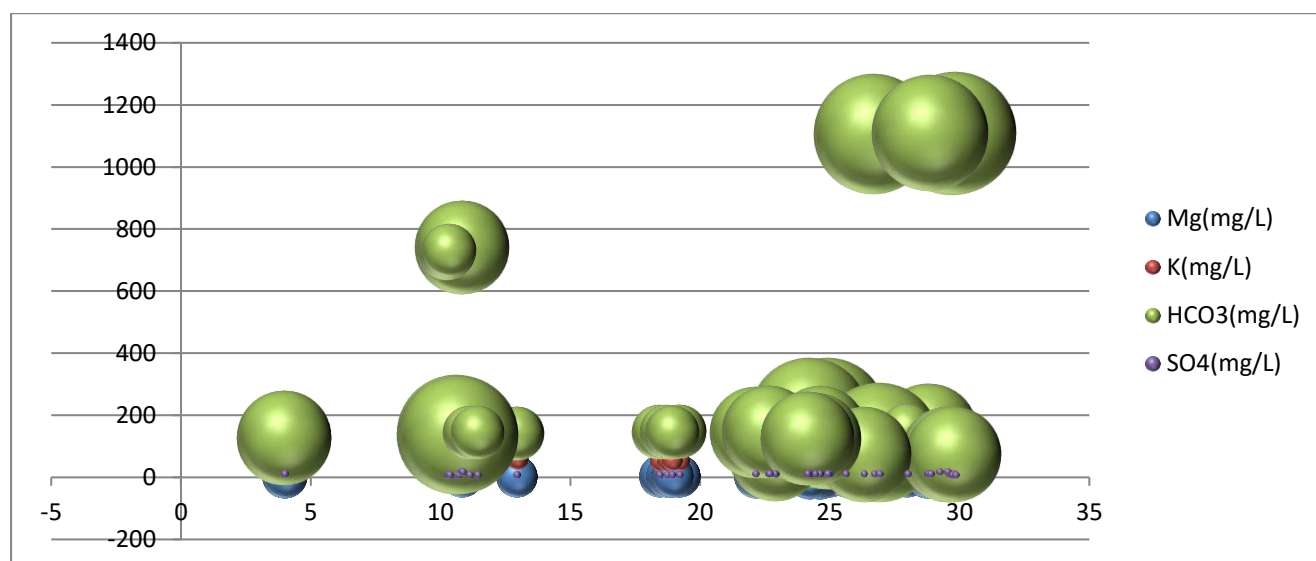
**Table 4: PCA for Physico-chemical Parameters of Ewekoro Borehole**

	Rotated Component Matrix <sup>a</sup> (Component)					
	C1	C2	C3	C4	C5	C6
PH	0.175	<b>0.671</b>	0.369	0.361	0.252	0.236
TEMP	-0.321	<b>0.920</b>	0.007	0.027	0.103	0.019
EC	0.155	-0.092	0.057	<b>0.945</b>	0.065	0.047
DO	<b>0.938</b>	0.170	0.084	0.091	0.055	-0.171
BOD	<b>0.935</b>	-0.110	0.022	0.114	0.088	0.262
COD	-0.795	0.124	0.441	0.129	-0.009	0.271
TDS	-0.193	-0.049	<b>0.862</b>	0.154	-0.193	0.263
TSS	-0.081	0.242	-0.216	-0.361	<b>0.769</b>	-0.153
TS	0.020	0.038	<b>0.917</b>	-0.028	0.213	-0.121
TURB	-0.047	0.235	0.050	<b>0.856</b>	-0.183	-0.222
ALK	-0.066	-0.039	0.107	-0.152	-0.134	<b>0.853</b>
TH	<b>0.562</b>	-0.299	-0.297	0.289	0.289	<b>0.502</b>
THC	0.313	-0.124	0.373	0.189	<b>0.790</b>	-0.016
Eigenvalue	<b>3.464</b>	<b>2.798</b>	<b>2.44</b>	<b>1.759</b>	<b>1.216</b>	<b>1.062</b>
Variance (%)	<b>24.742</b>	<b>19.985</b>	<b>17.434</b>	<b>12.565</b>	<b>8.682</b>	<b>7.583</b>
Cumulative (%)	<b>24.742</b>	<b>44.727</b>	<b>62.161</b>	<b>74.726</b>	<b>83.408</b>	<b>90.991</b>

**Component Plot in Rotated Space**



**Fig. 7: Plot of Extracted Components 1, 2 and 3 of Physico-Chemical Parameters of Ewekoro Boreholes in Rotated Space**



**Fig. 8: Bubble Plot of Ewekoro quality determinants showing the dominance of  $\text{HCO}_3^{2-}$  among other associated geochemical parameters in the study area.**

Furthermore, the output of the PCA and other multivariate techniques revealed the hydrogeochemical and other associated water quality status of the investigated area. The pH of boreholes from all the sampled locations in Ewekoro was well within the permissible limits of the approved standards alongside other physico-chemical parameters (NSDWQ, 2007; WHO, 2011; Udousoro and Umoren, 2014) (Table 1 and Table 2). The pattern of the overall ionic dominance for Wasinmi also follow the trend  $\text{Cl}^- > \text{K}^+ > \text{Na}^{2+} > \text{Fe}^{3+} > \text{NO}_3^- > \text{NO}_2^- > \text{Mn}^{2+}$  which is in consonance with the previous work done in the neighbouring town in Itori (Ishola *et al.*, 2023). The ionic concentration in the investigated borehole water samples contain more of  $\text{Fe}^{3+}$  than  $\text{Mn}^{2+}$  (Table 1) which is in agreement with report while the elevated values of  $\text{HCO}_3^-$  and  $\text{MgCO}_3^-$  are suggestive of possible subsurface dilution of the weathering of the carbonate rocks of the study area being a typical limestone formation (USEPA, 2002; Ishola *et al.*, 2023); this is supported by the obvious dominance of  $\text{HCO}_3^-$  among the water quality determinants displayed in the bubble plot (Fig. 8). In all the investigated hydrogeochemical parameters, Alkalinity,  $\text{HCO}_3^-$ ,  $\text{MgCO}_3^-$ ,  $\text{Cl}^-$ , BOD,  $\text{Cd}^{2+}$ ,  $\text{Fe}^{3+}$  and  $\text{Zn}^{2+}$  were found to be greater than the permissible standard of approved health organization, WHO inclusive. The results of the microbial analyses shown in Table 2 revealed the highest Total Coliform Count of  $17 \times 10^2$  MPN/100 ml while that of THBC and FCC are reported as  $7.8 \times 10^2$  cfu/ml and  $1.2 \times 10^2$  cfu/ml respectively. The coliform counts obtained from this study though lower in comparison to the previous work in Itori community (Ishola *et al.*, 2023); still this poses serious concerns and threatens the sanitary integrity of the groundwater system which may result in huge health implications in the study area (AGWT, 2013; APHA, 1992; Mahananda *et al.*, 2010; Sha'Ato *et al.*, 2010; WHO, 2011a).

## CONCLUSION

Diverse range of applications are utilizable with groundwater in the society ranging from domestic, agricultural, and industrial, among others, when it is within the allowable quality standards; however, when the physiochemical parameters and other quality determinants are



aversely altered rendering such supposedly fresh and quality water doubtful and consequently unfit depending on the required purposes, it becomes unsafe to continue with the use of such water irrespective of its abundance. It is therefore very important that any available water assigned for any specific usage must undergo different processes of qualitative analysis, treatment and monitoring processes to improve the quality status before applications, especially for consumption and other domestic usages. The PCA and other Multivariate statistics utilized in this work have characterized the hydrogeochemical determinants in terms of specific and principal weight contributions of each of the geochemical variables, ionic concentration levels and overall water quality status exemplifying them on the existing and still functional boreholes in the study area. Elemental, Physicochemical and Microbial analyses of potable water sources in Ewekoro communities, Ewekoro Local Government Area of Ogun State, South-West Nigeria revealed that the boreholes were Alkaline. Alkalinity,  $\text{HCO}_3^-$ ,  $\text{MgCO}_3^-$ ,  $\text{Cl}^-$ , BOD,  $\text{Cd}^{2+}$ ,  $\text{Fe}^{3+}$ ,  $\text{Zn}^{2+}$  and Total Coliform levels found to be greater than the recommended limits of the approved standards, including WHO and NESREA should be given special consideration for groundwater inspection, monitoring and management while the quality determinants revealed the dominance of  $\text{HCO}_3^{2-}$  in the groundwater system of the study area due to subsurface geology of the study area being of typical limestone formation in the study area. External pollution sources may include among others, wastes from domestic and agricultural activities, leachates from waste dumpsites and sewer tanks.

## REFERENCES

- AGWT (2013). Domestic water treatment for home owners. American Ground Water Trust's Consumer Awareness Information, Pamphlet #3. [Online] Available: <http://www.agwt.org/content/water-treatment>. (July 29, 2013).
- Amadi, A.N., Olasehinde, P.I., Yisa, J., Okosun, E.A., Nwankwoala, H.O., and Alkali, Y.B. (2012). Geostatistical Assessment of Ground Water Quality from Coastal aquifers of Eastern Niger Delta, Nigeria. *Goesciences* 2(3), 51-59.
- APHA (American Public Health Associated) (1998). Standard method for examination of water and waste water, 20th ed. American Public Health Associated Inc, New York. American Water Work Associated
- APHA (American Public Health Associated) (1992). Standard method for examination of water and waste water, 12th ed. American Public Health Associated Inc, New York. American Water Work Associated
- Burnett S.L. and Beuchat, L.R. (2001). Human Pathogens associated with raw produce and unpasteurized juices, and difficulties in contamination. *Journal of Industrial Microbiology and Biotechnology* 27(6), 104-110.
- Billman, H.G. (1992). Offshore stratigraphy and paleontology of the Dahomey (Benin) Embayment, West Africa, 1st. *NAPE Bull. No. 2, Vol. 7*, pp. 121 – 130.
- Davis, J.C., Manson, V., Ondrick C.W and Srivastava, B (1970). *Statistics and Data Analysis in Geology* Kansas Geological Survey Pp 500-505
- Davis, (1986). *Multivariate analyses of Geol. Data*, Kansas Geological Survey Pp 420.
- Faremi, A.Y. and Oloyede, O.B. (2010). Biochemical assessment of the effect of soap and detergent industrial effluents on some enzymes in the stomach of albino rats. *Research Journal of Environmental Toxicology* 4(3), 127-133. doi:10.2923/rjet.2010.127.133.



- FGN (Federal Government of Nigeria) (2012). A report of water resources in Nigeria in the world water day. The Nation Newspaper, Abuja. p.12.
- Foppen, J.W.A. (2002). Impact of high-strength wastewater infiltration on groundwater quality and drinking water supply: the case of Sama'a Yemen. *Journal of Hydrology* 263,198-216.
- Francisco S.M., Espinossa., R.J and Bosch., A.P (2001). Mapping Groundwater Quality Variables Using PCA and Geostatistics; A Case Study Of Banjo Andrax, Southwestern Spain. *Journal of Hydrological Sciences*, Vol 46, Pp 228-235.
- Gangopadhyay, S., Gupta, A.D., and Nachabe, M.H (2001). Evaluation of Groundwater Monitoring Network by Principal Component Analysis. Vol 39, No 2, *Groundwater Journal*, March-April, Pp 181-19.
- Howard, G., Pedley, S., Barrett, M., Nolubega, M., and Johal, K. (2003). Risk factors contributing to microbial contamination of shallow groundwater in Kampala, Uganda. *Water Research* 37, 3421-3429.
- Ishola S.A (2024). Groundwater Resource Potentials using Frequency Domain Electromagnetic Method and Associated Water Quality Techniques in Ewekoro Communities South-West Nigeria. *Federal University Wukari (FUW) Trends in Science and Technology Journal* [www.ftstjournal.com](http://www.ftstjournal.com). e-ISSN: 24085162; p-ISSN:20485170, April, 2024, Vol 9 (1). pp. 193-200.
- Ishola S.A, Emumejaye K and Ajetumobi A.E, (2023). Biogeochemical Assessment and Health Implications of Borehole Groundwater System using Principal Component Analysis in Itori Community South-West Nigeria. *Dutse Journal of Pure and Applied Sciences (DUJOPAS)*. ISSN Online 2635-34990. Vol 9(4b). DUJOPAS 9(4b). December, 2023, pp. 121-138. DOI: <https://dx.doi.org/0.43/4/dujopas.v9i4b.11> .
- Ishola S. A, Makinde V, Gbadebo A.M Mustapha A.O, Orebiyi E.A (2021): Quality Assessment of Groundwater System in Itori Community of Ewekoro Local Government Area, South-West Nigeria. *Science and Technology Publishing (SCI & TECH)* ISSN: 2632-1017, Multidisciplinary Engineering Science Studies Volume 5 Issue 12, December-2021; Southampton-UK).SCITECHP420180 <http://scitechpub.org/index.php/vol-5-issue-12-december-2021/>.
- Ishola S.A (2019). Characterization of Groundwater Resource Potentials using Integrated Techniques in Selected Communities within Ewekoro Local Government Area South-West Nigeria. Department of Physics, FUNAAB Ph.D Thesis.
- Ishola S.A., Makinde V., Aina J.O., Ayedun H., Akinboro F.G., Okeyode I.C., Coker J.O., and Alatise O.O (2016). Aquifer Protection Studies and Groundwater Vulnerability Assessment in Abeokuta South Local Government Area, South-West Nigeria, *Journal of Nigerian Association of Mathematical Physics*, © J.of NAMP Vol 33, (January, 2016), pp 347-362
- Kogbe C.A (1976). The Cretaceous and Paleogene sediments of southern Nigeria. In C. A. Kogbe (Ed.) *Geology of Nigeria*. pp. 325- 334.
- Kumar, J., and Pal, A. (2012). Water quality monitoring of Ken River of Banda District, Uttar Pradesh, India. *Elixir Pollution* 42, 6360-6364.
- Lashkaripour, G.R. (2003). Contamination of groundwater resources in Zaheden City due to rapid development. *Pakistan Journal of Applied Science* 3(5), 341-345.
- Liu, R.X., Kuang, J., Gong, Q., Hou, X.L (2003). Principal Component Regression with SPSS. *Computer Methods and Programs in Biomedicine* 71(2): 141 July 2003.



- Mahananda, M., Mohanty, B., and Behera, N. (2010). Physico-chemical analysis of surface and ground water of Bargarh District, Orissa, India. *International Journal of Research and Reviews in Applied Sciences* 2(3), 284-295.
- NAFDAC (2004). National Agency for Food and Drug Administration and Control, Guidelines for Production and Registration of Packaged Water, New guidelines and regulations. NAFDAC, Abuja Nigeria, pp. 22-30.
- NSDWQ (2007). Nigerian Standard for Drinking Water Quality. Nigerian Industrial Standard NIS 554, Standard Organisation of Nigeria, 30p.
- Obaje, N.G. 2009. Geology and mineral resources of Nigeria. In: Brooklyn SB, Bonn HJN, Gottingen JR, Graz KS, (ed), *Lecture Notes in Earth Sciences*, Springer, 120:22.
- Okosun, E.A. (1998). Review of the early Tertiary stratigraphy of southwestern Nigeria, *Jour. Min. and Geol.* Vol. 34, no 1, pp. 27–35.
- Oluseyi, T., Olayinka, K., and Adeleke, I. (2011). Assessment of ground water pollution in the residential areas of Ewekoro and Shagamu due to cement production. *African Journal of Environmental Science and Technology* 5(10), 786-794.
- Pasquini, M.W., and Alexander, M.J. (2004). Chemical properties of urban waste ash produced by open burning on the Jos Plateau: Implication for agriculture. *Science of the Total Environment* 319(1-3), 225-240.
- Pepper, I.L, Rusin, P, Quintanar, D.R, Haney, C, Josephson, K.L and Gerba, C.P. (2004). Tracking the concentration of heterotrophic plate count bacteria from the source to the consumer tap. *Int. J. Food Microbiol.*, 92: 289–295.
- Rail, C.D. (2000). *Groundwater Contamination, Volume 1: Contamination, sources, and hydrology*. Technomic Publishing Company, Lancaster, Pennsylvania, pp. 187.
- Sha’Ato, R., Akaahan, T.J., and Oluma, H.O.A. (2010). Physico-chemical and bacteriological quality of water from shallow wells in two rural communities in Benue State, Nigeria. *Pakistan Journal of Analytical and Environmental Chemistry* 11(1), 73-78.
- Udousoro, I.I and Umoren, I (2014). Assessment of Surface and Ground Water Quality of Uruan in Akwa Ibom State of Nigeria. *Journal of Natural Sciences Research* www.iiste.org ISSN 2224-3186 ISSN 2225-0921 (Online) Vol.4, No.6, 2014 11-27
- UNEP/ERCE/UNESCO (2008). *Water quality for ecosystems and human health*, (2nd Ed.). [Online] Available: International Journal of Chemical, Environmental and Pharmaceutical Research <http://www.un.org/waterforlifedecade/quality.shtml> (June 16, 2012) 1(3), 35-44.
- USEPA, (1979). *National Secondary Drinking Water Regulations*. Office of drinking water, Washington D. C. 20460, 42p.
- USEPA. (2002). US Environmental Protection Agency. Integrated risk information system (IRIS) database. National Center for Environmental Assessment. Washington D.C. 2002).
- Ushie, F., Harry, T., and Affiah, U. (2014). Reserve Estimation from Geoelectrical Sounding of the Ewekoro Limestone at Papalanto, Ogun State, Nigeria. *Journal of Energy Technologies and Policy* 2224 ISSN 2225-0573 (Online) Vol.4, No.5, pp. 28.
- Verplanck, P.I., Nordstrom, D.K., Plumlee, G.S., Wanty, R.B., Bove, D.J., and Caine, J.S. (2006). Hydrochemical controls on surface and groundwater chemistry in natural acidic porphyry-related mineralized areas, southern Rocky Mountains. *Chinese Journal of Geochemistry* 25, 231-241.
- Wakernagel, H (1995). *Multivariate Geostatistics. An Introduction with Applications* Springer-Verlag Berlin, Germany.





- WHO(2004).Guidelinefordrinkingwater,Vol.1,Geneva,Switzerland.[Online]Available:<http://whqlibdoc.who.int/publications/2004/9241546387.pdf> (July 29, 2013).
- WHO (2011a). Guideline for Drinking Water Quality, (4th Ed.). [Online] Available: [http://www.who.int/water\\_sanitation\\_health/publications/2011/dwq\\_chapters/en/](http://www.who.int/water_sanitation_health/publications/2011/dwq_chapters/en/) (July 29, 2013).
- WHO (2011b). Manganese in Drinking-water: Background document for development of WHO guidelines for Drinking-water quality. Geneva, Switzerland, pp 29. Available: [http://www.who.int/water\\_sanitation\\_health/dwq/chemicals/manganese.pdf](http://www.who.int/water_sanitation_health/dwq/chemicals/manganese.pdf).Sept,17,2013
- Wu, T.N., Huang, Y.C., Lee, M.S., and Kao, C.M. (2005). Source identification of groundwater pollution with the aid of multivariate statistical analysis. *Water Science and Technology: Water supply* 5(6), 281-288.
- Yeung, I.M.H. (1999). Multivariate analysis of the Hong Kong Victoria Harbour water quality data. *Environmental Monitoring and Assessment* 60(3-4), 365-380.