



# Modelling Groundwater Quality of Aba in Abia State Using Principal Component Analysis and Multiple Linear Regression

Ogbonnaya Paul Kanu <sup>a\*</sup>, Ejikeme Ugwoha <sup>b</sup>,  
Ngozi U. Udeh <sup>b</sup> and Victor Amah <sup>b</sup>

<sup>a</sup> Centre for Occupational Health, Safety and Environment, University of Port Harcourt, Nigeria.

<sup>b</sup> Department of Civil and Environmental Engineering, University of Port Harcourt, Nigeria.

## Authors' contributions

*This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.*

## Article Information

DOI: 10.9734/JERR/2023/v25i111019

## Open Peer Review History:

This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: <https://www.sdiarticle5.com/review-history/108553>

**Original Research Article**

**Received: 14/09/2023**

**Accepted: 20/11/2023**

**Published: 23/11/2023**

## ABSTRACT

The aim of this study was to model the groundwater quality of Aba in Abia state. To achieve the aim, thirty-two water samples were taken from sixteen boreholes during the rainy and dry seasons and analysed in the laboratory for pH, Electrical Conductivity, Total Hardness, BOD<sub>5</sub>, COD, Pb, Cd, Cr, NH<sub>3</sub>, TDS, SO<sub>4</sub>, NO<sub>3</sub> and PO<sub>4</sub>. Principal Component Analysis (PCA) and Multiple Linear Regression (MLR) were employed to extract the principal factors and develop a model for predicting water quality index for Aba, Abia State. In the dry season, water quality index could be estimated using the Water Quality Index (WQI) model with pH, PO<sub>4</sub>, COD, SO<sub>4</sub> and Pb with Adjusted R<sup>2</sup> = 0.999999999938 and standard error of 0.043868872. Meanwhile, in the rainy season, WQI could be estimated using the WQI model with Turbidity, PO<sub>4</sub>, NO<sub>3</sub>, COD, SO<sub>4</sub> and Pb with Adjusted R<sup>2</sup> = 0.999999997469 and standard error of 0.066697494. The one-way ANOVA for the parameters in

\*Corresponding author: Email: [ogbonnayakanu@gmail.com](mailto:ogbonnayakanu@gmail.com);

the dry season with  $p = 0.000 < 0.05$  indicated that leachate had a large effect on groundwater quality. During the rainy season, one-way ANOVA result with  $p = 0.000 < 0.05$  asserted that leachate had a large effect on groundwater quality.

*Keywords: Water quality index; Aba; groundwater quality; principal component analysis; multiple linear regression.*

## 1. INTRODUCTION

Water is one of the most requested of all human necessities and the most basic and critical requirement for biota [1]. Water also plays a vital role in the earth's ecosystem, as it is the principal component of the earth and an integral part of several natural resources. Water exists substantially either as surface or groundwater. Groundwater is water found in all voids of a geologic stratum [2]. It is considered as the most important source of freshwater especially for the arid and semi-arid regions attributable to the low precipitation rates in those regions [3]. Freshwater which is the form of water available for human consumption, exist in the more feasible form as rivers, streams and groundwater. Groundwater serves as the main reserve of freshwater and the main sources of water in the urban environment, which is used for drinking, industrial, agricultural and domestic purposes [4]. Groundwater is important to the entire world's population, but more important to semi-arid and arid regions where there is insufficiency of surface water [5].

Throughout the world, water is recognized as the most fundamental and indispensable of all natural resources. The demand for groundwater supply for drinking, domestic, agricultural, and industrial use has been on a steady increase owing to global population surge. Although not the largest source of water, groundwater is the largest available source of freshwater, making it an essential component of the water supply chain and a valuable natural resource [1].

Groundwater is a renewable resource, recharged by precipitation. It is often the first alternative choice of many consumers attributable to its perceived cleanness and safeness, since it is recharged, recycled and filtered through biological, natural processes that keeps it free from contamination and pollution. Groundwater becomes polluted from natural and anthropogenic sources, regardless. The safety of groundwater be it shallow or deep groundwater source, depends on the geology of the area, human activities/land use activities of the area,

environmental processes and meteorological condition of the area. It is for this reason that groundwater is assessed systematically and monitored continually, applying proven scientific methods, to ascertain its quality so as to ensure that it can be used for domestic purposes, and that no adverse effects are experienced attributable to its use.

Groundwater pollution is a big quandary especially in developing countries where there is inadequate infrastructure to treat and distribute water to the populace [6]. This is compounded by waste management challenges, which brings about groundwater pollution particularly by leachate from dumpsites.

Therefore, finding appropriate locations for suitable groundwater for drinking reasons is a significant difficulty for satisfying water demands. Groundwater quality declines for a variety of reasons, such as natural reversals in the flow of rivers, improper management of water bodies, climate variability, and human activities [7].

According to [8], people in rural areas of developing nations like Nigeria have access to cleaner water sources than those who reside in cities, where the issue of poor water quality is made worse by industrial effluents, municipal refuse dumps, and the disposal of toxic, metallic, and organic wastes.

The quality, quantity and accessibility of groundwater is requisite for holistic, sustained socio-economic and environmental development since it caters to the water needs of sectors across spheres of life [9]. An analytical tool and composite indicator known as a water quality index (WQI) gives end users information based on predetermined water content factors that are then transformed into a single unitless value. WQI offers the benefit of determining the condition of water quality without having to interpret each parameter separately. Nonetheless, more than 20 water quality indices were created and revived across the world until 1970. Because of its simplicity of use and scientific basis, several studies used various

WQIs to assess water quality. As a result of this research, a considerable amount of data on water quality is created, which must be gathered in order to determine the water quality status in a specific region [10]. Water quality indices (WQIs) are essential for making simpler the reportage of complex and technical water quality data, on a scientific bases of communication, using models that are capable of converting multi-variable water quality data to produce a single unitless digit score that describes overall quality of water, which is important for providing a structured platform to appraise and compare water quality of various water resources [11].

This study aims at modelling groundwater quality of the research area using Principal Component Analysis and Multiple Linear Regression. It provides critical data for protecting local groundwater resources from human contamination and building long-term harmony between people, society, and environment by raising awareness about the impact of people's insufficient waste management activities on groundwater quality. Access to clean drinking water and sanitation for all people everywhere is one of the United Nations' sustainable development goals (SDG 6). The study's findings could help policy makers establish and enforce appropriate environmental legislation to prevent groundwater contamination and boost the efficiency for groundwater quality assessment. It

demonstrates how environmental education may be expanded and where it might go to help people construct a better future. Future research in the area should include the measurement of more biological characteristics such Total coliform bacteria as microbial pollution of groundwater is unsafe for human health.

## 2. MATERIALS AND METHODS

### 2.1 Study Area

The study area, Aba city in Abia state is the main economic nerve centre of Abia State, Nigeria (Fig. 1). It serves the commercial needs of important cities in the south-eastern, and south-southern geopolitical regions of Nigeria. Over one million people are estimated to be living in the 236km<sup>2</sup> area. The vegetation and climate conditions are humid equatorial rainforest, with annual mean rainfall ranging from 2150 to 2460mm and mean daily temperatures ranging between 22°C and 33°C for the average minimum and maximum daily temperatures [12].

Aba lies approximately within longitude 7°19' E to 7°37' E and latitude of 5°3' N to 5°12' N in the Niger Delta Basin and is underlain by the Benin Formation. The aquifer type is mostly unconfined and the water table elevation range 26m –33m below ground level, with an average elevation depth of 28.6m [13,12].

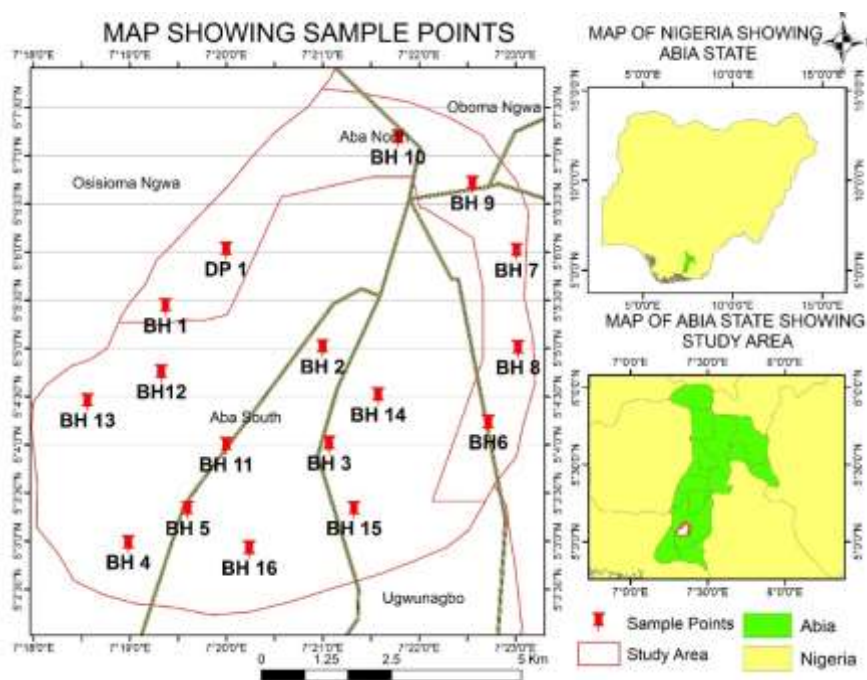


Fig. 1. Map of the study area showing sampling points

## 2.2 Sample Collection

Thirty-two Samples were taken from boreholes in Aba during the dry and rainy seasons, at evenly distributed sampling points by inserting a grid on the map of Aba using ArcGIS software and taking water samples from boreholes on the nodes of the grid for laboratory analysis. Borehole locations were selected based on spatial distribution on the grid of the map of Aba. Leachate sample was taken from the existing dumpsite in Aba.

The water samples were taken in properly sterilized bottles that had been thoroughly cleaned. Before obtaining a water sample, the bottles were rinsed three times with the groundwater sample. Following that, the obtained samples were put in an ice-filled cooler and transferred to the laboratory.

Other equipment and materials used during the study are as follows:

- i. Global Positioning system device (GPS App on mobile phone) was used to obtain the geographical position of the sampled boreholes and dumpsite.
- ii. Turbidity meter was used to read the in-situ turbidity of the groundwater samples.
- iii. pH Meter was used to determine the pH level of the groundwater samples.
- iv. Thermometer was used to read the temperature of the groundwater samples.

## 2.3 Sample Analysis

The concentrations of quality parameters (pH, Electrical Conductivity, TDS, Turbidity, PO<sub>4</sub>, NO<sub>3</sub>, COD, SO<sub>4</sub>, Total Hardness, BOD<sub>5</sub> and Pb) were analysed. All analyses were carried out in accordance with established methods and as described in the literature [14]. The study employed drinking water quality standards provided by the World Health Organization drinking water guideline to detect excessive quantities of these parameters.

## 2.4 Calculation of Water Quality Index

The physicochemical parameters (pH, EC, TDS, BOD<sub>5</sub>, COD, TH, SO<sub>4</sub>, NO<sub>3</sub>, Turbidity, Pb and PO<sub>4</sub>) were used to calculate the WQI in this study. The Water Quality Index (WQI) was calculated using the Weighted Average Water

Quality Index (WAWQI) method [15,5] which involved:

Step 1: Gathering information on the water quality metrics that will be used to calculate the WQI.

Step 2: Computation of k

Using Equation (1) to compute k.

$$k = \left(1 / \sum_{i=1}^n \frac{1}{s_i}\right) \quad (1)$$

where: k = proportionality constant

S<sub>i</sub> = Standard permissible limit

Step 3: Compute the nth parameter's quality rating, using Equation (2).

$$q_n = 100 \left(\frac{(v_n - v_{io})}{(s_i - v_{io})}\right) \quad (2)$$

where: v<sub>n</sub> = estimated concentration of the nth parameter of the given sampling location.

v<sub>io</sub> = ideal value of the nth parameter in pure water.

S<sub>i</sub> = standard permissible limit of the nth parameter.

Step 4: Using Equation (3), determine the unit weight of the nth parameter.

$$W_n = \left(\frac{k}{s_i}\right) \quad (3)$$

Step 5: Using Equation (4), calculate the Water Quality Index.

$$WQI = \left(\frac{\sum w_n * q_n}{\sum w_n}\right) \quad (4)$$

The water quality index rating, graded A – E, is shown on Table 1.

**Table 1. Water quality index (WQI) ratings**

WQI value	Water Quality Rating	Grade
0 - 25	Excellent	A
26 - 50	Good	B
51 – 75	Poor	C
76 - 100	Very Poor	D
Above 100	Unsuitable for drinking purpose	E

Source: [16,5]

## 2.5 Statistical Analysis and Model Development

The one-way analysis of variance (ANOVA) was utilized, to see if the parameter concentrations altered considerably [17]. Principal Component Analysis (PCA) and Multiple Linear Regression (MLR) were used to extract the principal factors and develop a model for predicting water quality index of the study area.

The Multiple Linear Regression model was given as shown below:

$$\text{Outcome} = \text{model} + \text{Error} \quad (5)$$

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_nX_n + \epsilon \quad (6)$$

$$Y = \beta_0 + \sum_{i=1}^n(\beta_iX_i) + \epsilon \quad (7)$$

where  $\beta_0$  is the sample intercept,  $\beta_1$  is the sample slope parameter for  $X_1$ ,  $\beta_2$  is the sample slope parameter for  $X_2$  and so forth, 'ε' represents the sample errors/residuals.

## 3. RESULTS AND DISCUSSION

The Water quality parameters measured included: Temperature, pH, Turbidity, Electrical Conductivity, Nitrate, Phosphate, Total Hardness, Carbonate, Sulphate, BOD<sub>5</sub>, COD, Ammonia, Lead, Cadmium, Chromium and TDS. Table 2 presents the location of the sampled boreholes in the study area.

**Table 2. Location of sampled boreholes and Ariaria dumpsite**

Borehole	Latitude	Longitude
BH1	5° 6' 30.528" N	7° 19' 41.952" E
BH2	5° 4' 59.412" N	7° 20' 59.928" E
BH3	5° 3' 59.256" N	7° 21' 4.032" E
BH4	5° 2' 57.588" N	7° 18' 59.004" E
BH5	5° 3' 18.756" N	7° 19' 35.976" E
BH6	5° 4' 12.216" N	7° 22' 43.032" E
BH7	5° 5' 59.424" N	7° 23' 0.708" E
BH8	5° 4' 58.908" N	7° 23' 1.608" E
BH9	5° 6' 41.148" N	7° 22' 33.348" E
BH10	5° 7' 9.804" N	7° 21' 46.98" E
BH11	5° 3' 58.5" N	7° 20' 0.06" E
BH12	5° 4' 43.752" N	7° 19' 19.524" E
BH13	5° 4' 25.932" N	7° 18' 33.588" E
BH14	5° 4' 29.532" N	7° 21' 34.56" E
BH15	5° 3' 18.756" N	7° 21' 19.44" E
BH16	5° 2' 54.096" N	7° 20' 14.352" E
DP1	5° 6' 34.02" N	7° 19' 41.628" E

## 3.1 Water Quality Indices for Dry and Rainy Season

Water quality index (WQI) was calculated using eleven parameters (pH, Electrical Conductivity, TDS, Turbidity, PO<sub>4</sub>, NO<sub>3</sub>, COD, SO<sub>4</sub>, Total Hardness, BOD<sub>5</sub> and Pb) [18]. The calculated water quality indices for dry and rainy seasons are presented in Table 3.

## 3.2 Developed Model for Groundwater Quality Index

The principal component analysis and multiple linear regression model development are presented in the Tables 4 – 12 and in the Figs. 2 - 5. Tables 4 & 6 show the eigenvalues of the extracted Factors, F1 – F10 of water quality parameters for dry and rainy seasons, respectively. Table 5 and 7 present the Varimax Rotated Component (VF1 – VF3) Matrix showing loadings of water quality parameters for dry and rainy season.

The model efficiency is validated by the coefficient of determination, R<sup>2</sup> and the Adjusted R<sup>2</sup>. The effectiveness is computed by regression coefficient and the standard error [10]. The selected dry season model had a coefficient of determination, R<sup>2</sup> = 0.99999999958541, Adjusted R<sup>2</sup> = 0.999999999938 and standard error of 0.043868872 as shown on Tables 12 and 13. This implies that 99.99999999% of the water quality index is explained by the model. The selected rainy season model had a coefficient of determination, R<sup>2</sup> = 0.99999999848, Adjusted R<sup>2</sup> = 0.999999997469 and standard error of 0.066697494 as shown on Tables 12 and 13. This implies that 99.9999998% of the water quality index is explained by the model..

From Tables 9 and 11, the models developed asserts that lead is quite significant and has a large impact on WQI. The presence of heavy metals such as lead in drinking water is a major public health concern, and water treatment techniques such as activated carbon filtration at the point of use can be used to limit its consumption [19].

In the dry season, the extracted factors VF2 and VF3 were rejected due to low coefficient of determination, R<sup>2</sup> of 0.42993 and 0.02620, and high standard errors of 4904.74 and 5896.73 respectively for the dry season.

**Table 3. Water quality index for dry and rainy season interpretation**

<b>Location_ID</b>	<b>WQI - Dry season</b>	<b>Interpretation</b>	<b>WQI - Rainy season</b>	<b>Interpretation</b>
BH1	21,531	Unsuitable for drinking purpose	5,545	Unsuitable for drinking purpose
BH2	3,847	Unsuitable for drinking purpose	352	Unsuitable for drinking purpose
BH3	2,492	Unsuitable for drinking purpose	851	Unsuitable for drinking purpose
BH4	1,518	Unsuitable for drinking purpose	1,020	Unsuitable for drinking purpose
BH5	3,562	Unsuitable for drinking purpose	198	Unsuitable for drinking purpose
BH6	12,628	Unsuitable for drinking purpose	86	Very poor
BH7	242	Unsuitable for drinking purpose	176	Unsuitable for drinking purpose
BH8	8,773	Unsuitable for drinking purpose	612	Unsuitable for drinking purpose
BH9	835	Unsuitable for drinking purpose	938	Unsuitable for drinking purpose
BH10	2,279	Unsuitable for drinking purpose	1,200	Unsuitable for drinking purpose
BH11	850	Unsuitable for drinking purpose	31	Good
BH12	1,236	Unsuitable for drinking purpose	194	Unsuitable for drinking purpose
BH13	5,120	Unsuitable for drinking purpose	185	Unsuitable for drinking purpose
BH14	1,219	Unsuitable for drinking purpose	242	Unsuitable for drinking purpose
BH15	3,075	Unsuitable for drinking purpose	359	Unsuitable for drinking purpose
BH16	3,105	Unsuitable for drinking purpose	253	Unsuitable for drinking purpose

**Table 4. Eigenvalues of Extracted factors of the water quality parameters for dry season**

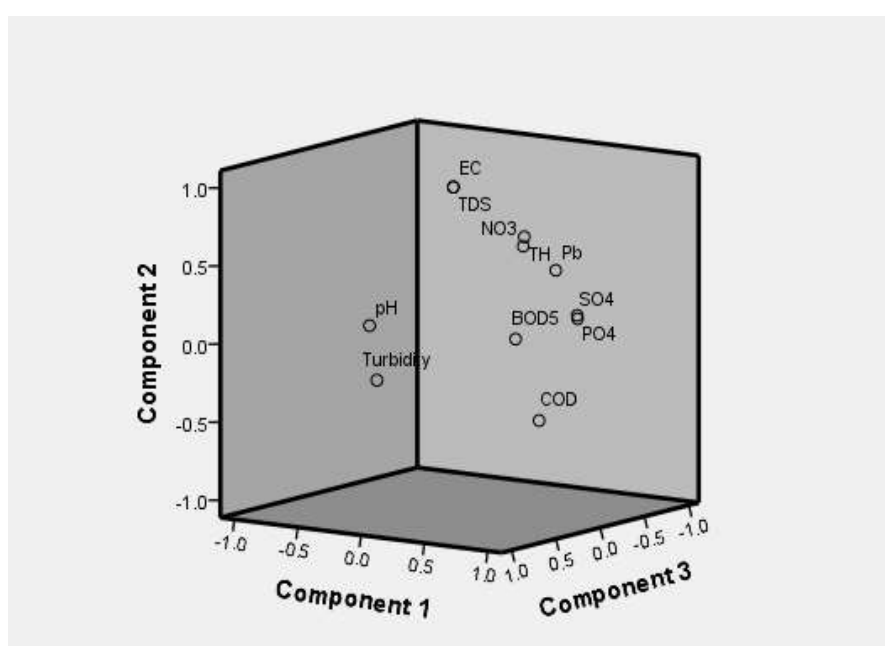
Component	Eigenvalues	% of Variance	Cumulative %
F1	4.489	40.813	40.813
F2	2.420	21.997	62.810
F3	1.268	11.530	74.340
F4	.966	8.779	83.119
F5	.701	6.374	89.493
F6	.466	4.233	93.726
F7	.433	3.940	97.665
F8	.183	1.668	99.333
F9	.073	.666	99.999
F10	.000	.001	100.000

**Table 5. Varimax rotated component matrix showing loadings of water quality parameters for dry season**

Parameter	Component		
	VF1	VF2	VF3
SO <sub>4</sub>	.961	.217	.041
PO <sub>4</sub>	.958	.238	.041
COD	.723	-.445	.134
Pb	.717	.488	-.061
pH	-.552	.047	.227
EC	-.016	.963	.048
TDS	-.016	.963	.048
NO <sub>3</sub>	.559	.647	.081
TH	.362	.645	-.212
Turbidity	-.064	-.164	.840
BOD <sub>5</sub>	-.020	-.113	-.661

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.



**Fig. 2. Component plot in rotated space using PCA for dry season**

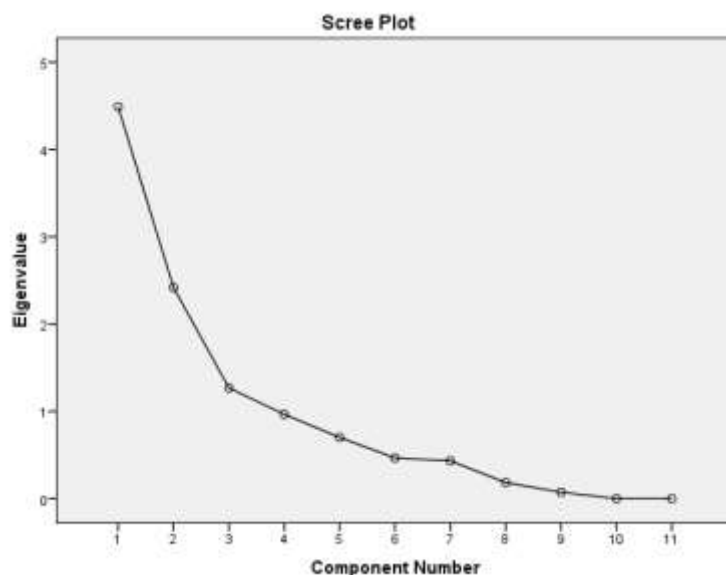


Fig. 3. Scree plot for dry season Eigenvalues using PCA

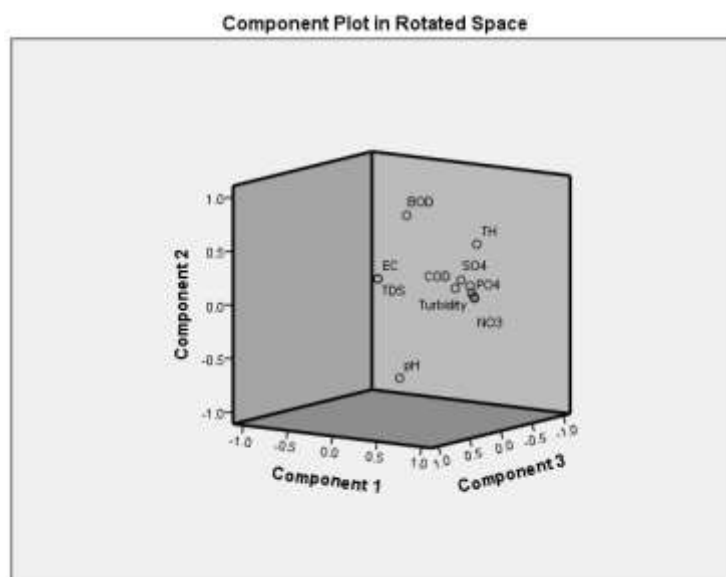


Fig. 4. Component plot in rotated space using PCA for rainy season

Table 6. Eigenvalues of extracted factors of the water quality parameters for rainy season

Component	Eigenvalues	% of Variance	Cumulative %
F1	6.326	57.505	57.505
F2	1.631	14.830	72.335
F3	1.058	9.614	81.949
F4	.795	7.225	89.174
F5	.610	5.543	94.716
F6	.371	3.376	98.092
F7	.159	1.443	99.535
F8	.049	.442	99.978
F9	.002	.020	99.997
F10	.000	.003	100.000



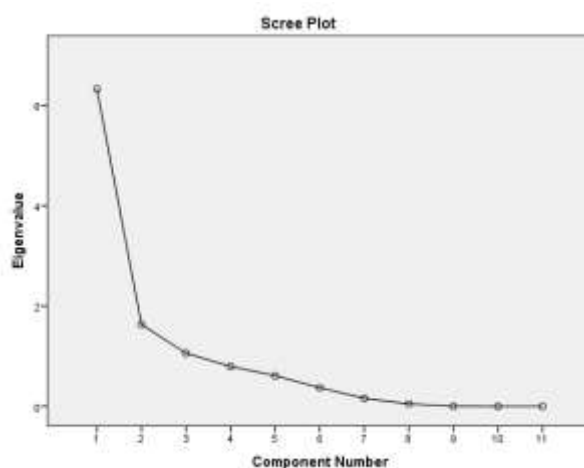
**Table 7. Varimax rotated component matrix showing loadings of water quality parameters for rainy season**

Parameter	Component		
	VF1	VF2	VF3
PO <sub>4</sub>	.957	.139	.199
NO <sub>3</sub>	.950	.156	.205
Turbidity	.940	.191	.227
Pb	.903	.249	.195
SO <sub>4</sub>	.773	.284	.154
COD	.569	.159	-.046
BOD <sub>5</sub>	.198	.837	.200
pH	-.216	-.791	-.275
TDS	.321	.350	.838
EC	.321	.350	.838
TH	.459	.488	-.546

Extraction Method: Principal Component Analysis.  
Rotation Method: Varimax with Kaiser Normalization.

**Table 8. Observed vs predicted WQI using MLR for dry season**

Location	Observed WQI	VF1 Predicted WQI	VF2 Predicted WQI	VF3 Predicted WQI
BH1	21,531	21,531	14,206	4,793
BH2	3,847	3,847	3,046	5,142
BH3	2,492	2,492	381	4,735
BH4	1,518	1,518	2,086	2,038
BH5	3,562	3,562	8,120	4,425
BH6	12,628	12,628	7,170	4,851
BH7	242	242	1,788	4,634
BH8	8,773	8,772	4,310	4,363
BH9	835	835	5,973	5,221
BH10	2,279	2,279	4,481	4,390
BH11	850	850	6,344	5,690
BH12	1,236	1,236	7,030	4,063
BH13	5,120	5,120	- 981	5,530
BH14	1,219	1,219	1,665	3,370
BH15	3,075	3,075	3,090	3,856
BH16	3,105	3,105	3,603	5,209



**Fig. 5. Scree plot for rainy season Eigenvalues using PCA**

**Table 9. Coefficients of multiple linear regression for dry season model**

	<b>Coefficients</b>	<b>Standard Error</b>	<b>t Stat</b>	<b>P-value</b>
Intercept	2.1876	0.2206	9.9181	0.0000
pH	-0.2992	0.0315	-9.5007	0.0000
PO <sub>4</sub>	0.8921	0.4882	1.8273	0.0976
COD	0.0007	0.0012	0.5822	0.5733
SO <sub>4</sub>	0.0015	0.0096	0.1605	0.8757
Pb	9843.0373	0.0444	221749.1848	0.0000

**Table 10. Observed vs Predicted WQI using MLR for rainy season**

<b>Location</b>	<b>Observed WQI</b>	<b>VF1 Predicted WQI</b>	<b>VF2 Predicted WQI</b>	<b>VF3 Predicted WQI</b>
BH1	5545	5545	1805	2943
BH2	352	352	1662	1631
BH3	851	851	1040	919
BH4	1020	1020	1379	550
BH5	198	198	580	1103
BH6	86	86	823	1944
BH7	176	176	399	-493
BH8	612	612	807	152
BH9	938	938	880	1175
BH10	1200	1200	1363	714
BH11	31	31	1029	-4
BH12	194	194	920	478
BH13	185	185	-229	-73
BH14	242	242	581	1051
BH15	359	359	-165	122
BH16	253	253	-633	31

**Table 11. Coefficients of multiple linear regression for rainy season model**

	<b>Coefficients</b>	<b>Standard Error</b>	<b>t Stat</b>	<b>P-value</b>
Intercept	0.035254272	0.088981796	0.396196	0.701189199
Turbidity	3.195065622	2.362583817	1.352361	0.209258699
PO <sub>4</sub>	0.220138224	0.977805336	0.225135	0.826903007
NO <sub>3</sub>	-0.021404109	0.124790785	-0.17152	0.867609487
COD	-0.000107496	0.002708228	-0.03969	0.969204995
SO <sub>4</sub>	0.000597849	0.001027713	0.581727	0.575029755
Pb	9843.911771	0.539270088	18254.14	0.0000

While for the rainy season, the extracted factors VF2 and VF3 were rejected due to low coefficient of determination, R<sup>2</sup> of 0.2585765 and 0.440177, and high standard errors of 1226.32953 and 1065.6124 respectively.

The visual validation of WQI model for dry and rainy seasons are presented in Figs. 6 and 7. The plots of the observed and predicted water quality index shows overlap and agreement for both seasons.

### 3.3 Validation of Model for Groundwater Quality Index

The selected water quality index models with the coefficient of determination, R<sup>2</sup> and its adjusted R<sup>2</sup> are presented in Table 13. The coefficients of the multiple linear regression for dry season and rainy season models are given in Tables 9 and 11, respectively.

### 3.4 Effect of Leachate on Water Quality Parameters

The influence of leachate on the different borehole characteristics may be understood from the Analysis of Variance of water quality parameters in the dry and rain seasons, as illustrated in Tables 14 and 15. The analysis of variance (ANOVA) showed that during the dry

season,  $p = 0.000 < .05$ . The consequence is to reject  $H_0$ , which asserts that leachate has no substantial effect on groundwater quality, and accept  $H_1$ , which asserts that leachate has a large effect on groundwater quality.

During the rainy season, ANOVA results were  $p = 0.000 < .05$ . The consequence is to reject  $H_0$ , which asserts that leachate has no substantial effect on groundwater quality, and accept  $H_1$ , which asserts that leachate has a large effect on groundwater quality.

**Table 12. Simulated water quality index models**

Extracted Factors	Model	Goodness of fit Statistics	Decision
<b>Dry Season</b>			
VF1	$f(x_1, x_2, x_3, x_4, x_5)$	$R^2 = 0.99999999958541$ Adjusted $R^2 = 0.99999999938$ Standard Error = 0.043868872	Model Accepted
VF2	$f(x_1, x_2, x_3, x_4)$	$R^2 = 0.429934250051281$ Adjusted $R^2 = 0.131728522797$ Standard Error = 4904.7360388	Model Rejected
VF3	$f(x_1, x_2)$	$R^2 = 0.0262076314606923$ Adjusted $R^2 = -0.123606579$ Standard Error = 5896.72690736	Model Rejected
<b>Rainy Season</b>			
VF1	$f(x_1, x_2, x_3, x_4, x_5, x_6)$	$R^2 = 0.999999998481656$ Adjusted $R^2 = 0.999999997469427$ Standard Error = 0.066697494	Model Accepted
VF2	$f(x_1, x_2)$	$R^2 = 0.2585765$ Adjusted $R^2 = 0.14451135$ Standard Error = 1226.32953	Model Rejected
VF3	$f(x_1, x_2, x_3)$	$R^2 = 0.440177$ Adjusted $R^2 = 0.2771276$ Standard Error = 1065.6124	Model Rejected

**Table 13. Selected water quality index models**

Model Equation	$R^2$	Adjusted $R^2$
<b>Dry Season</b>		
$y = 2.1876 - 0.2992 x_1 + 0.8921 x_2 + 0.00067 x_3 + 0.001544 x_4 + 9843.037 x_5$ where $y =$ Water quality index (WQI), $x_1 =$ pH, $x_2 =$ $PO_4$ , $x_3 =$ COD, $x_4 =$ $SO_4$ and $x_5 =$ Pb	0.99999999958541	0.99999999937812
<b>Rainy Season</b>		
$y = 0.03525 + 3.1951 x_1 + 0.22014 x_2 - 0.021404 x_3 - 0.0001075 x_4 + 0.000598 x_5 + 9843.912 x_6$ where $y =$ Water quality index (WQI), $x_1 =$ Turbidity, $x_2 =$ $PO_4$ , $x_3 =$ $NO_3$ , $x_4 =$ COD and $x_5 =$ $SO_4$ , $x_6 =$ Pb	0.999999998481656	0.999999997469427

**Table 14. One-Way ANOVA for dry season water quality parameters**

ANOVA: Single Factor				
Summary				
Groups	Count	Sum	Average	Variance
pH	16	103.46	6.46625	0.178665
Temp.	16	388.8	24.3	5.42
EC	16	3400.01	212.5006	19976.04
TDS	16	2176.006	136.0004	8182.187
Turbidity	16	6.299	0.393688	0.114718

ANOVA: Single Factor				
Summary				
Groups	Count	Sum	Average	Variance
NH <sub>3</sub>	15	3.556	0.237067	0.073562
PO <sub>4</sub>	16	4.606	0.287875	0.685022
NO <sub>3</sub>	16	80.986	5.061625	56.06293
COD	16	951.6	59.475	283.8588
SO <sub>4</sub>	16	170.479	10.65494	1808.973
Total Hardness	16	1532.71	95.79438	2439.972
BOD <sub>5</sub>	16	43.69	2.730625	0.91734
Pb	16	5.15847	0.322404	0.114408
Cd	16	0.166556	0.01041	0.000101
Cr	16	15.76931	0.985582	0.302964

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	904168	14	64583.43	29.44432	4.2E-43	1.736146
Within Groups	491323.5	224	2193.408			
Total	1395491	238				

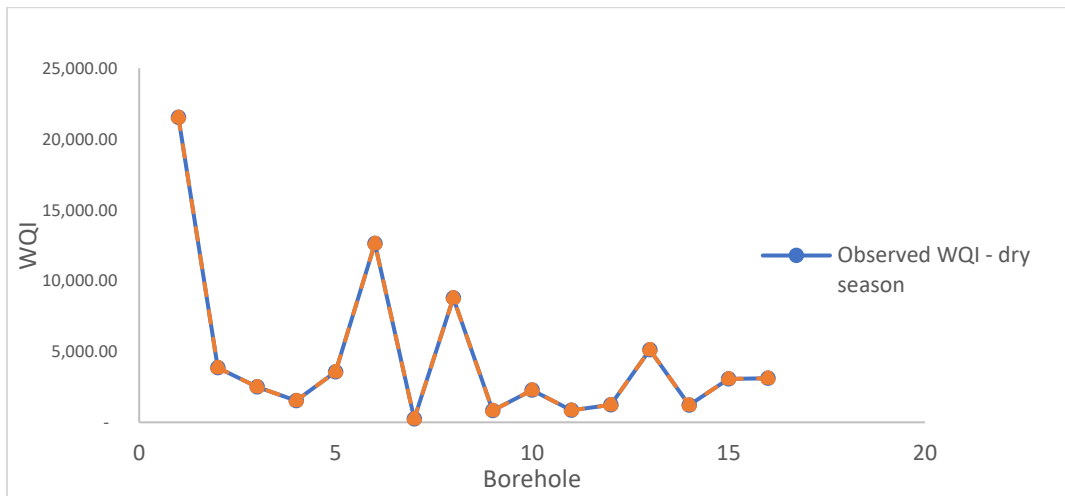


Fig. 6. Visual validation of WQI model for dry season

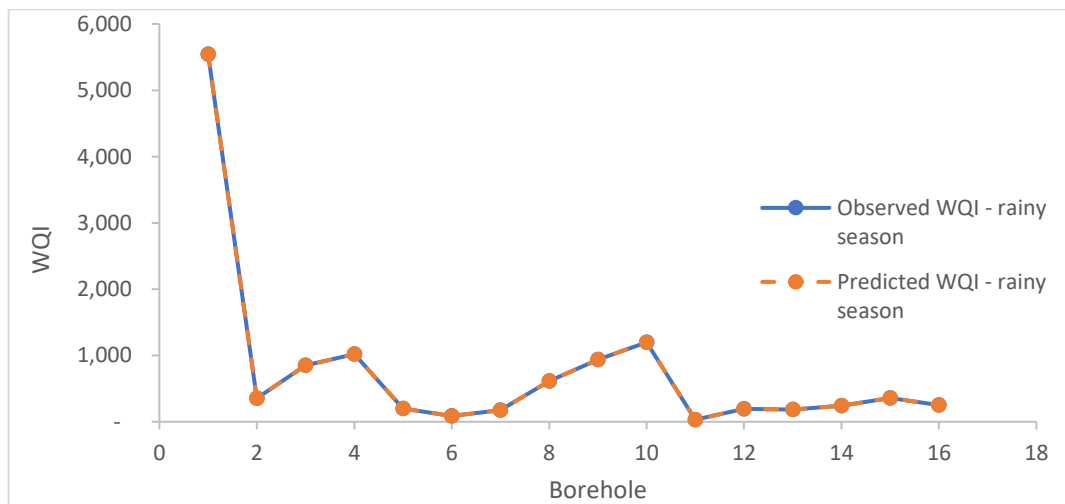


Fig. 7. Visual validation of WQI model for rainy season

Table 15. One-Way ANOVA for rainy season water quality parameters

<b>ANOVA: Single Factor</b>						
<b>Summary</b>						
<b>Groups</b>	<b>Count</b>	<b>Sum</b>	<b>Average</b>	<b>Variance</b>		
pH	16	110.73	6.920625	0.03574		
Temp.	16	358.6	22.4125	3.743833		
EC	16	4331.3	270.7063	26604.02		
TDS	16	2772.032	173.252	10897.01		
Turbidity	16	0.767	0.047938	0.013408		
NH <sub>3</sub>	16	0.951	0.059438	0.000692		
PO <sub>4</sub>	16	2.368	0.148	0.21115		
NO <sub>3</sub>	16	23.859	1.491188	23.28919		
COD	16	624.51	39.03188	111.9083		
SO <sub>4</sub>	16	735.25	45.95313	882.3156		
Total Hardness	16	1337.66	83.60375	1982.661		
BOD <sub>5</sub>	16	41.51	2.594375	0.55868		
Pb	16	1.24332	0.077708	0.018131		
Cd	16	1.724082	0.107755	0.1785		
Cr	16	4.45798	0.278624	0.06681		
<b>ANOVA</b>						
<b>Source of Variation</b>	<b>SS</b>	<b>df</b>	<b>MS</b>	<b>F</b>	<b>P-value</b>	<b>F crit</b>
Between Groups	1385633	14	98973.8	36.6515	3.53E-50	1.735948
Within Groups	607590.5	225	2700.402			
Total	1993224	239				

#### 4. CONCLUSION

This research was undertaken to perform the modelling of groundwater quality of Aba in Abia state. Groundwater samples were analysed in the laboratory for pH, EC, Total Hardness, BOD<sub>5</sub>, COD, Pb, Cd, Cr, NH<sub>3</sub>, TDS, SO<sub>4</sub>, NO<sub>3</sub> and PO<sub>4</sub>. Principal Component Analysis (PCA) and Multiple Linear Regression (MLR) were used to extract the principal factors and develop a model for predicting water quality index of Aba. In the dry season, water quality index can be estimated using the WQI model with pH, PO<sub>4</sub>, COD, SO<sub>4</sub> and Pb. Meanwhile, in the rainy season, water quality index can be estimated using the WQI model with Turbidity, PO<sub>4</sub>, NO<sub>3</sub>, COD, SO<sub>4</sub> and Pb. Though there were parameters with minimal concentrations, they could contribute to the deterioration of water quality in the future. This contaminated groundwater can be purified by using conventional methods such as Chemical Precipitation, Ion Exchange, Membrane Filtration (Ultrafiltration, Reverse Osmosis or Nanofiltration), Coagulation and Flocculation, Electrochemical Treatment, and Adsorption [20].

#### COMPETING INTERESTS

Authors have declared that no competing interests exist.

#### REFERENCES

- Adagba T, Kankara AI, Idris MA. Evaluation of groundwater suitability for irrigation purpose using gis and irrigation water quality indices. FUDMA Journal of Sciences. 2022;6(2):63–80.
- Todd DK, Mays LW. Groundwater Hydrology (3rd, Ed.). John Wiley and Sons; 2005.
- Kayemah N, Al-Ruzouq R, Shanableh A, Yilmaz AG. Evaluation of groundwater quality using Groundwater Quality Index (GWQI) in Sharjah, UAE. In E3S Web of Conferences. 2021;241: 01005. EDP Sciences.
- Okoye NM 1, Orakwe LC, Nwachukwu PC. Groundwater quality mapping using GIS: A case study of Awka, Anambra State, Nigeria. International Journal of Engineering and Management Research. 2016;6(2):579-584.
- Amah VE, Agu FA. Geostatistical Modelling of Groundwater Quality at Rumuola Community, Port Harcourt, Nigeria. Asian Journal of Environment & Ecology. 2020;37–47.
- Lateef A, Kolawole LL, Agosu SO. Quality assessment of some groundwater samples in Ogbomoso metropolis, Southwest

- Nigeria. *Journal of Environment and Earth Science*. 2012;2(6):39-49.
7. Bodrud-Doza MD, Islam ARMT, Ahmed F, Das S, Saha N, Rahman MS. Characterization of groundwater quality using water evaluation indices, multivariate statistics and geostatistics in central Bangladesh. *Water Science*. 2016;30(1):19–40.
  8. Ewuzie U, Nnorom IC, Ugbogu O, Onwuka CV. Hydrogeochemical, microbial and compositional analysis of data from surface and groundwater sources in Southeastern Nigeria. *Journal of Geochemical Exploration*. 2021;224: 106737.
  9. Singh R, Upreti P, Allemailem KS, Almatroudi A, Rahmani AH, Albalawi GM. Geospatial Assessment of Ground Water Quality and Associated Health Problems in the Western Region of India. *Water*. 2022;14:296.  
Available: <https://doi.org/10.3390/w14030296>
  10. Kadam AK, Wagh VM, Muley AA, Umrikar BN, Sankhua RN. Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India. *Modeling Earth Systems and Environment*. 2019;5(3):951–962.
  11. Banda TD, Kumarasamy MV. Development of Water Quality Indices (WQIs): A Review. *Polish Journal of Environmental Studies*. 2020;29(3).
  12. Ijioma UD. Delineating the impact of urbanization on the hydrochemistry and quality of groundwater wells in Aba, Nigeria. *Journal of Contaminant Hydrology*. 2021;240:103792.
  13. Agharanya UP, Dim CIP. Water quality assessment and resource potentials: the case of Aba–Urban and its Environs, Niger Delta Basin. *Water Resources*. 2018;45(2):250–267.
  14. Saleh HN, Valipoor S, Zarei A, Yousefi M, Asghari FB, Mohammadi AA, Amiri F, Ghalehaskar S, Khaneghah AM. Assessment of groundwater quality around municipal solid waste landfill by using Water Quality Index for groundwater resources and multivariate statistical technique: A case study of the landfill site, Qaem Shahr City, Iran. *Environmental Geochemistry and Health*. 2020;42(5): 1305–1319.
  15. Kasa T, Reddythota D. Investigation of the Wabe river water's suitability for drinking purposes and aquatic life and detection of pollution sources. *Applied Water Science*. 2023;13(7):154.
  16. Odia M, Nwaogazie IL, Nwachukwu EO, Awwiri GO. Modelling leachate pollution index and potential for selected municipal solid waste dump sites: A case study. *British Journal of Applied Science and Technology*. 2016;18(6): 1–16.
  17. Nwaogazie IL, Osamudiamen OP, Bovwe O. Modelling groundwater quality index based on sensitivity analysis for wet and dry seasons in Obio/Akpor local government area, Rivers state, Nigeria. *Nigerian Journal of Technology*. 2018; 37(3):799–805.
  18. Kanu OP, Ugwoha E, Udeh NU, Amah V. Groundwater quality assessment in Aba (Abia State) Using WQI and GIS Techniques. *Asian Journal of Environment & Ecology*. 2023;22(4): 1–15.  
Available: <https://doi.org/10.9734/ajee/2023/v22i4503>
  19. Jarvis P, Fawell J. Lead in drinking water— an ongoing public health concern? *Current Opinion in Environmental Science & Health*. 2021;20:100239.
  20. Pal DB, Tiwari AK, Giri DD. Various purification techniques of groundwater. *Groundwater Geochemistry: Pollution and Remediation Methods*. 2021;310-325.  
Available: <https://doi.org/10.1002/9781119709732.ch15>

**APPENDIX**

**Table A1. WHO Drinking Water Standard for Water quality parameters**

Parameters	WHO
pH	6.5 – 8.5
EC	1000
TDS	500
Turbidity	5
PO <sub>4</sub>	1
NO <sub>3</sub>	50
COD	250
SO <sub>4</sub>	250
TH	250
BOD <sub>5</sub>	5
Pb	0.01

**Table A2. Measured water quality parameters from sampled boreholes for dry season**

Borehole	Ph	Temp. °c	EC (S/cm)	TDS (mg/l)	Turbidity (NTU)	NH <sub>3</sub> (mg/l)	PO <sub>4</sub> (mg/l)	NO <sub>3</sub> (mg/l)	COD (mg/l)	SO <sub>4</sub> (mg/l)	Total Hardness (mg/l)	BOD <sub>5</sub> (mg/l)	Pb (mg/l)	Cd (mg/l)	Cr (mg/l)
BH1	5.74	22.3	315.7	202.048	0.315	1.863	3.39	23.14	96.41	170.15	183.01	2.53	2.18741	0.01631	6.03985
BH2	6.35	22	100.1	64.064	0.133	0.762	0.07	-0.075	67.84	0.009	84.1	3.01	0.39081	0.00427	0.32011
BH3	6.49	21	37.4	23.936	0.183	0.056	0.07	0.041	61.44	0.006	50.4	3.91	0.25314	0.00431	0.81973
BH4	6.65	22	190.8	122.112	1.392	0.527	0.056	0.83	55.68	0.014	54.1	1.98	0.15424	0.01203	0.91031
BH5	6.57	26.2	269.4	172.416	0.284	0.072	0.079	12.02	53.76	0.05	181.3	4.03	0.36182	0.00894	0.84391
BH6	6.31	26.8	558.9	357.696	0.365	0.143	0.14	10.53	17.68	0.023	150.4	1.89	1.28294	0.0362	0.52413
BH7	5.97	22.5	123.5	79.04	0.249	0.078	0.073	1.252	47.36	0.015	98.3	3.65	0.02451	0.01386	2.00317
BH8	6.82	27	255.9	163.776	0.241	0.814	0.056	0.751	60.16	0.019	79.6	4.62	0.89121	0.03102	1.93115
BH9	6.48	23.7	315.2	201.728	0.174	0.126	0.126	13.7	45.44	0.031	120.5	2.38	0.08475	0.00178	1.02131
BH10	6.74	26.4	254.3	162.752	0.428	0.089	0.056	0.625	62.72	0.009	77.6	2.85	0.23149	0.00394	1.25231
BH11	6.32	26.9	196.2	125.568	0.118	0.073	0.07	1.295	65.28	0.011	147.6	1.33	0.08631	0.00658	1.5231
BH12	6.93	23.6	311.2	199.168	0.591	0.112	0.112	16.35	75.52	0.031	52.6	2.47	0.12558	0.00821	1.42874
BH13	6.24	22.1	19.11	12.2304	0.187	0.056	0.056	0.112	52.87	0.012	44.2	1.24	0.52016	0.009346	0.98533
BH14	5.77	27	78.3	50.112	0.802	0.098	0.098	0.256	75.52	0.057	26.9	2.87	0.12374	0.00173	0.42103
BH15	7.35	22.3	58.2	37.248	0.713	0.511	0.056	-0.071	62.72	0.007	118.9	2.06	0.31238	0.01187	0.74987
BH16	6.73	27	315.8	202.112	0.124	0.039	0.098	0.23	51.2	0.035	63.2	2.87	0.31539	0.01247	1.03511

**Table A3. Measured water quality parameters from sampled boreholes for rainy season**

Borehole	pH	Temp. °c	EC (µS/cm)	TDS (mg/l)	Turbidity (NTU)	NH <sub>3</sub> (mg/l)	PO <sub>4</sub> (mg/l)	NO <sub>3</sub> (mg/l)	COD (mg/l)	S0 <sub>4</sub> (mg/l)	Total Hardness (mg/l)	BOD <sub>5</sub> (mg/l)	Pb (mg/l)	Cd (mg/l)	Cr (mg/l)
BH1	6.64	20.4	586.19	375.1616	0.480	0.035	1.870	19.57	57.13	130.93	150.96	3.54	0.56317	1.692082	0.00396
BH2	6.78	20.2	590.21	377.7344	0.043	0.074	0.029	0.284	52.72	64.25	45.30	3.87	0.03572	0.00115	0.25167
BH3	6.81	19.8	470.3	300.992	0.015	0.056	0.035	0.162	29.63	51.94	28.9	2.68	0.08641	BDL	0.11851
BH4	6.77	20.3	318.1	203.584	0.023	0.042	0.018	0.157	35.11	22.63	50.6	3.22	0.10365	0.00593	0.36973
BH5	6.83	24.7	286.4	183.296	0.01	0.084	0.062	0.139	48.39	19.84	105.2	1.79	0.02014	0.00167	0.84202
BH6	6.89	23.5	410.6	262.784	0.036	0.042	0.014	0.820	32.05	24.11	130.4	2.58	0.00873	0.00256	0.34516
BH7	6.88	20.0	153.8	98.432	0.018	0.028	0.033	0.402	22.69	52.69	22.8	1.63	0.01784	0.00487	0.10803
BH8	6.86	25.0	210.3	134.592	0.008	0.118	0.051	0.360	41.82	63.72	55.1	2.41	0.06213	BDL	0.08917
BH9	6.91	22.0	94.3	60.352	0.014	0.042	0.021	0.087	45.31	58.23	175.3	2.79	0.09524	0.00172	0.02635
BH10	6.87	25.1	177.2	113.408	0.021	0.056	0.068	0.565	27.03	42.91	110.9	3.64	0.1219	0.00163	0.52718
BH11	6.93	24.2	83.7	53.568	0.036	0.041	0.037	0.48	50.94	74.86	85.2	3.2	0.00315	0.00201	0.20963
BH12	6.86	23.0	182.6	116.864	0.027	0.098	0.026	0.466	24.13	13.22	90.3	2.65	0.01971	0.00658	0.01781
BH13	7.20	21.6	110.9	70.976	0.011	0.028	0.051	0.017	47.04	17.39	70.6	1.74	0.01877	0.00275	0.1931
BH14	6.94	23.6	250.1	160.064	0.003	0.084	0.013	0.144	36.17	29.18	113.2	2.29	0.02457	BDL	0.35178
BH15	7.13	21.2	186.2	119.168	0.017	0.052	0.018	0.097	40.51	43.63	60.8	1.56	0.03648	0.00113	0.19001
BH16	7.43	24.0	220.4	141.056	0.005	0.071	0.022	0.109	33.84	25.72	42.1	1.92	0.02571	BDL	0.81387

© 2023 Kanu et al.; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:  
 The peer review history for this paper can be accessed here:  
<https://www.sdiarticle5.com/review-history/108553>