# Groundwater Contaminant Identification and Modelling Using Principal Component Analysis and Geostatistical Techniques

Izinyon, O.C Meindinyo, L.I Ilaboya, I.R Department of Civil Engineering, Faculty of Engineering, PMB 1154, University of Benin, Benin City, Edo State. Nigeria

#### Abstract

Groundwater which constitutes high percent of the global fresh water is the most important source of drinking water, which when polluted, have acute effects on its users. Consequently, the quality and pollution of groundwater is a health concern in the world. The target of this research is to evaluate the quality of groundwater around the Niger Delta Basin Development Authority in order to identify and analyze the distribution of the critical contaminants that affect the overall quality of groundwater water around the study area. About hundred (100) boreholes spread to cover the study area were sampled. The water samples were analyzed using standard procedures for assessing drinking water qualities in order to understand the existing condition of groundwater within the study area. Statistical analysis of the groundwater quality data was done using average weighted index method to compute the water quality index, factor analysis using principal component method to identify the groundwater contaminants affecting the overall groundwater quality and geospatial analysis using kriging interpolation method to evaluate the spatial distribution of the selected groundwater contaminants. From the principal component analysis, result revealed that; nitrate, total dissolved solids, concentration of iron, total suspended solids and turbidity were the most important contaminants affecting the quality of the groundwater. Result of geospatial analysis using kriging interpolation revealed that; the water quality parameters showed relatively strong degree of spatial dependency which made it possible to generate the spatial distribution map for the selected water quality parameters.

**Keywords:** Geospatial analysis, Kriging interpolation, water quality index, Principal component analysis. **DOI**: 10.7176/JNSR/9-24-05

Publication date: December 31st 2019

#### 1. Introduction

The availability of water in adequate quality and quantity is essential for the existence of human life. Water is one of the basic elements that sustain life on earth. Man can survive for a considerable amount of time without food, but he will succumb within days without water. It is reported in (Nichole, 2004; Akporido, 2000) and Water for Life, (2007) that where there is water, there is life; and that life exists around numerous uses of water which make it important for survival and luxury. The average daily intake of water per person, directly or as contained in other foods, is about four litres (Water for life (2007). For those who live in hot dry climate, this daily requirement is at least double (Akporido, 2000; Water for Life, 2007). Whenever a number of people live together, a supply of water is always paramount. On this note, for all purposes for which water is required, the quality of water is of great importance (Ocheri et al., 2014). As world population is constantly growing, the demand for water increases every day.

Groundwater is one of the major sources of water. Groundwater is water stored under the earth surface in rocks and soil segments called aquifers and is less susceptible to contamination by the action of microorganism (Debels, et al., 2005). As groundwater use has increased, issue associated with the quality of groundwater resources have likewise grown in importance. Groundwater, depending on the depth and the prevailing soil condition is vulnerable to pollution by anthropogenic activities, such as agricultural sources, urbanization, un-engineered disposal of wastes including de-icing salt, oil and chemical spillage (Baalousha, 2010). These sundry activities are carried out by man without due regard to its consequences on the long run. Groundwater, once polluted can be very difficult to remediate back to its natural pristine state. For many years, attention has been directed at contamination from point sources. More recently, concerns have increased about nonpoint sources of contaminant and about the overall quality of groundwater resources (Faisal et al., 2003; Efe, 2003). Groundwater contamination can be detected by analyzing borehole water for a series of dissolved ion species. Hence water quality data sets typically contain many variables measured at several spatially scattered locations.

A review conducted in Ocheri, (2014), on groundwater quality in Nigeria urban areas revealed that Nigerian urban groundwater quality is influenced by the geological condition of the soil through which it flows, geochemical and physical process of the environment, rate of urbanization, industrialization and seasonal variability. Abams et al, (2007) and Agbalagba et al, (2011) conducted a review on the quality of groundwater in parts of Yenagoa and Port Harcourt and reported that the groundwater in these areas have been negatively impacted by anthropogenic activities in the areas.

# Materials and Methods Description of study area

The study area for this research is the Niger Delta Basin Development Authority. This study covers the original area of operation of the River Basin Authority, which is Rivers and Bayelsa State alone. The geographical coordinates of Rivers and Bayelsa states are 4.8581°N and 6.9209°E and 4.25°S and 5.37°W and 6.75°E respectively (Nwankwoala et al., 2011). The Niger Delta Basin is situated in the south-south geo-political zone of Nigeria. It is located in the rain forest region with relative humidity above 80% having an annual temperature range of 25°C to 31°C and annual rainfall of 4700mm on the coast to about 2400mm. The basin is characterized by two alternating climatic conditions of a long period of rainy season spanning from March to November, followed by a dry season spreading from November to March (Nwankwoala, et al., 2011). Figures 1 and 2 shows the Google earth and the study area maps respectively.



Figure 1: Google earth map of study area (Google .com)





The Niger Delta Basin is located on the continental margin of the Gulf of Guinea in equatorial West Africa. The Niger Delta lithofacies is made up of the three distinct vertical subdivisions viz. the Benin formation, the Agbada formation and the Akata formation. The Benin formation being the upper delta-top Lithofacies comprises of massive continental sands and gravels. The Agbada formation or facie consists of the pro-delta marine shales, with low stand turbidite fans which are deposited in a deep marine setting. In the Northern Delta Sector during the Oligocene times the Benin formation first occurs (Bolaji and Tse, 2009). Similarly, Paleocene age was established as the occurrence of the Akata formation in the proximal parts of the Delta. The Niger Delta complex geomorphologic features comprise of fresh water swamps, mangrove swamps, beaches, bars, and estuaries (Bolaji and Tse, 2009).

# 2.2 Sampling location and sample collection

The boundary of built up area (land use) within the study area was digitized and gridded at 2km interval to determine the sampling points and ensure uniform coverage. Water samples was collected systematically so as to have a general overview of the water quality condition within the study area. For accurate geo-referencing of the selected boreholes, Garmin hand held GPS receiver was employed to determine the geographical coordinates of each borehole. A section of the boreholes sampled including their location and geographical coordinates is presented in Table 1. One hundred (100) boreholes were systematically sampled with reference to location points at each season: Wet season (July to October 2018) and dry season (November to December 2018) in order to determine the physico-chemical and biological parameters of the groundwater samples. At every point of collection, the air tight, clean and dried plastic containers were rinsed two to three times with the borehole water to be sampled before collection. The samples were labelled properly and stored in air tight, clean and dried plastic containers before been transported to Water Resources and Environmental laboratory in the Department of Civil Engineering, University of Benin were the analysis were conducted in line with standard procedures and guideline recommended by World Health Organization (WHO). The water samples were analyzed in triplicates to obtain the mean value and standard deviation of each water quality test parameters. For the analysis of biochemical oxygen demand (BOD), the black bottles containing the water samples remained tightly closed prior to analysis in order to prevent photosynthetic and oxygen generation. In-situ parameters, namely; dissolved oxygen (DO), temperature, pH electrical conductivity (EC) and total dissolved solids (TDS) were determined in the field immediately after sample collection to avoid false measurement values (APHA, 2005).

Borehole Codes	Locations	Northings	Easting
1. Sample R1	Igbu Ahaoda	239820	561471
2. Sample R2	Mini Ama	269110	525361
3. Sample R3	Arukwo-Abua	235669	537656
4. Sample R4	Bakana	286341	528043
5. Sample R5	Edeoha-Ahoada	237214	556600
6. Sample R6	Edeoha-Ahoada	236203	556600
7. Sample R7	Okoboh-Abua	235766	540433
8. Sample R8	Buguma	262207	524264
9. Sample R9	Air force Base	280557	534103
10. Sample R10	Trans Amadi	279389	530030
11. Sample R11	Ipo-Ikwerre	274121	532098
12. Sample R12	Woji	286716	533642
13. Sample R13	Rumuokwurushi (1)	283293	536010
14. Sample R14	Amakiri Polo	286238	527163
15. Sample R15	Rukpokwu	289003	534162
16. Sample R16	Aggrey	280451	526634
17. Sample R17	NDBDA	278741	529397
18. Sample R18	Rumuokwurushi (2)	283012	536068
19. Sample R19	Amadi-Ama	279849	530118
20. Sample R20	Owodu	287302	531219
21. Sample R21	Okochiri	307314	519241
22. Sample R22	Trans Amadi (3)	278023	530112
23. Sample R23	Railway	279801	527029
24. Sample R24	Bundu	279684	525973
25. Sample R25	Oyorokoto	325714	496236
26. Sample R26	Kono Town	334047	508598
27. Sample R27	Oyigbo (1)	289245	538032
28. Sample R28	Ngo Town Andoni	323819	495804
29. Sample R29	Yegha Gokona	319044	517018
30. Sample R30	Oyigbo (2)	289599	538240
31. Sample R31	Nyokuru	339050	510170
32. Sample R32	Tegu-Gokana	316831	519746
33. Sample R33	Woji (2)	286421	533116

# Table 1: Coordinate Data of Sampled Boreholes (RIVERS)

# 2.3 Water Quality Analysis

A total of thirty-three (33) physico-chemical parameters and two (2) microbiological parameters were analyzed for each sampled domestic borehole to provide an insight into the overall quality of water within the study area. The physico-chemical parameters include: temperature, odour, colour/clarity, total hydrocarbon content (THC),

pH, Electrical conductivity (EC), Turbidity, Total suspended solid (TSS), Salinity, Alkalinity, Total Dissolve Solids (TDS), and Dissolved Oxygen (DO). Others are; Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Bicarbonate (HCO<sub>3</sub>), Sodium (Na), Potassium(K), Calcium(Ca), Magnesium (Mg), Chloride(Cl<sup>-</sup>), Phosphorus (P), Ammonium(NH<sub>4</sub>), Nitrite (NO<sub>2</sub>), Nitrate (NO<sub>3</sub>), Sulphate (SO<sub>4</sub>) and heavy metals, namely; Iron (Fe), Manganese (Mn), Zinc (Zn), Copper (Cu), Chromium (Cr), Cadmium (Cd), Nickel (Ni) and Lead (Pb). The microbiological parameters include: Total Coliform Counts (TCC) and E. Coli

# 2.3.1 Determination of in-situ parameters; (pH, EC, TDS, DO and Temperature)

For electrical conductivity (EC), total dissolved solids (TDS), pH, temperature and dissolved oxygen (DO), in-situ measurements were carried out since the measurement values of the parameter's changes with storage time (WHO, 2003). pH, electrical conductivity, temperature and total dissolved solids were measured using portable meter's (multi-parameters) while dissolved oxygen was examined using DO meter (Lutron DO-5509, Range 0 - 20mg/l) shown in Figure 3



# Figure 3: DO meter and multi portable meter

The multi portable meter probe was submerged in the water at 4cm and pH mode selected. Water sample was stirred gently and pH value displayed on the meter was allowed to adjust and stabilize before recording. Other measurements buttons were pressed successively and values recorded. The procedure was repeated three (3) times and the mean value calculated for each parameter. DO meter was also inserted into the water sample at about 10cm depth using the oxygen probe handle.

UNICAM 969 Atomic Absorption Spectrometer (AAS) shown in Figure 4 was used to determine the concentration of heavy metals such as; Iron (Fe), Manganese (Mn), Zinc (Zn), Copper (Cu), Chromium (Cr), Cadmium (Cd), Nickel (Ni), Lead (Pb), and Vanadium (V) while UV visible spectrophotometer (Thermo Scientific Spectronic 20D+) presented in Figure 5 was used to analyzed the level of phosphorous (P), Nitrate (NO<sub>3</sub>), Nitrite (NO<sub>2</sub>) and Sulphate (SO<sub>4</sub>). Other apparatus utilized included 250ml separating glass funnels, Cuvette, 10ml and 50ml pipette, 250ml conical flask, 50ml burette, 25ml and 50ml volumetric flask, glass beads, refrigerator, oven and whatman filter paper.





Figure 4: UNICAM 969 AA Spectrometer

Figure 5: UV Visible Spectrophotometer

Preparation of reagents and procedures employed in the laboratory for the analysis and determination of all water quality parameters followed the standard methods recommended by relevant authorities such as World Health Organization (WHO).

# 2.4 Groundwater contaminant identification

Statistical computations involving the application of principal component analysis for groundwater analysis was done using statistical package for the social sciences (SPSS 22 software). Principal component analysis (PCA) is a powerful tool that attempts to explain the variance of a large dataset of intercorrelated variables with a smaller set of independent variables. The Principal Components (PC) is the uncorrelated (orthogonal) variables obtained by multiplying the original correlated variables with the eigenvector, which is a list of coefficients (loadings or weightings). Thus, the PCs are weighted linear combinations of the original variables. PC provides information on the most meaningful parameters, which describe the whole data set while affording data reduction with a minimum loss of original information (Gajbhiye et al., 2014). Application of PCA to ground water analysis was conducted by using the following step by step methodology

- i. Computation of Anti-image correlation matrix
- ii. Computation of eigen values
- iii. Extraction of the component matrix
- iv. Extraction of the rotated component matrix
- v. Horizontal decentralization of the component matrix

# 2.4.1 Anti- Image Correlation Matrix

The anti-image correlation matrix was employed to test the suitability of principal component analysis (PCA) in explaining the underlying correlation between selected water quality parameters. The suitability of PCA is based on the values of the off diagonal matrix. If the off-diagonal matrix are very close to zero, then PCA is suitable otherwise we must think of another statistical method (Gajbhiye et al., 2014).

# 2.4.2 Computation of Eigen values

The eigen value is normally used as cutoff in factor analysis since it is the sum of the squared factor loadings of all variables. Factors with eigen values greater than 1 represent the number of component factors needed to describe the underlying variation of the groundwater quality. This are the component factors that contributes an adequate amount to the variation in the ground water quality. Factors with eigen value less than 1 means that such factor do not have any influence on the overall groundwater quality.

# 2.4.3 Computation of communualities

It is also important in the extraction phase to examine the communualities. The communalities is represented by the sum of the square loading for a variable across factor. Communality can range from 0 to 1. Communality of 1 means that all of the variation in the groundwater quality is explained by the component factors.

# 2.4.4 Extraction of component matrix

In other to identify the groundwater quality parameters that make up the members of each component factors, the component matrix was generated. The extracted component matrix was employed to understand the correlation between the groundwater quality parameters in each component group. In regression terms, the component matrix is the standardized regression coefficient between the observed values and the component factors. Higher factor loading indicates that a parameter is closely associated with the component factor. To determine the water quality parameters that make up each of the component factors, horizontal decentralization of the component matrix was done and the best favoured parameters was selected as member of that particular component factor.

www.iiste.org

#### 2.4.5 Extraction of the rotated component matrix

The rotated component matrix gives you the factor (s) you need to pay more attention to in other to improve the overall quality of the groundwater in any location. To understand the factors that is most critical towards improving the overall groundwater quality of the study area, vertical decentralization of the rotated matrix was done (Gajbhiye et al., 2014). For principal component analysis (PCA) to work, the mean was subtracted from each of the data dimensions. The mean subtracted represent the average across each dimension. The mean was computed from

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{1}$$

To compute the covariance matrix, equation (2) was employed as follows;

$$Var(X) = \frac{\sum_{i=1}^{n} \left( X_i - \bar{X} \right) \left( X_i - \bar{X} \right)}{n-1}$$
(2)

From equation (2), the covariance matrix was computed using

$$Cov(X,Y) = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right) \left(Y_i - \bar{Y}\right)}{n-1}$$
(3)

Covariance is normally measured in 2-dimensions such as (X, Y); (X, Z) or (Y, X). Provided the X and Y covariance matrix are of the same size, addition and multiplication of the covariance matrix was employed to calculate the eigenvector and eigenvalues of the covariance matrix. The notion of data compression and reduced dimensionality comes into PCA in computing the feature vector. To compute the feature vector, equation (4) was employed as follows

Feature Vector = (eig1, eig2, eig3 - eign) (4)

Once the component eigenvector is chosen, the feature vector was formed and then transposed to get the final solution as follows; Final Data = (Row Feature Vector) \* (Row Data Adjust) (5)

#### 2.5 Groundwater analysis using geostatistical techniques

The geostatistical analyst is an Arcmap extension that provides a wide variety of tools for spatial data exploration, identification of data anomalies and evaluation of error in prediction surface models, statistical estimation, and optimal surface creation. The geostatistical analyst can determine the approximate concentration of selected water quality parameters in an unsampled location and can determine the spatial distribution of the parameter i.e. where they may be moving by creating an optimal interpolated surface.

The step by step methodology involved in the creation of interpolation surface for selected groundwater quality parameters are:

- i. Collection of water sample from selected boreholes within the study area
- ii. Collection of spatial data (Northings, Easting and Elevation) of the selected boreholes within the study area using Garmin Hand-held GPS
- iii. Analysis of the physico-chemical and microbiological properties of the collected groundwater samples
- iv. Modelling the spatial distribution of selected groundwater parameters using kriging interpolation in ArcGIS

Kriging is a geostatistical interpolation technique that considers both the distance and the degree of variation between known data points when estimating values in unknown areas. It attempts to minimize the error variance and set the mean of the prediction errors to zero so that there are no over- or under-estimates. It is a robust interpolation tool which derives weights from surrounding measured values to predict values at unmeasured locations. The kriging weights are obtained from fitting of semi- variogram models, developed by viewing the spatial structure of the data. To create a continuous surface or map of any phenomenon, predictions are made from the models and the spatial arrangement of measured values. In this study, ordinary kriging method was employed for the spatial analysis of some selected groundwater quality parameters owing to its simplicity and prediction accuracy in comparison to other kriging methods. In addition, it is an error minimization technique where the values of a property (water quality) at unsampled location are predicted such that the variance of the estimated value is reduced to the barest minimum. The following steps are involved in the use of kriging interpolation method for the geospatial analysis of selected groundwater quality parameters

- i. Evaluation of normality test
- ii. Selection of attribute data and model interpolation method

- iii. Semivariogram fitting and testing
- iv. Cross validation
- v. Spatial Dependency determination
- vi. Creation of groundwater quality prediction maps

#### 2.5.1 Test of Normality

To test the normality of the water quality datasets, histogram plots of the different water quality parameters were generated using statistical package for the social sciences (SPSS version 22) software

#### 2.5.2 Selection of attribute data and model interpolation method

Geostatistics can reveal stochastic structure of groundwater quality variations in space and can showed a significant variation in the pH of groundwater within and around the study area. Table 2 presents the modelling step for the spatial variation of pH and other selected water quality parameters. It is observed that the attribute data selected is pH while the interpolation method is Kriging/Cokriging method.

# Table 2: Kriging modeling for the spatial variation of pH

lethods	Input Data		
Deterministic methods	Dataset		
Inverse Distance Weighting	Source Dataset	Sheet1\$ Events	
Global Polynomial Interpolation	Data Field	pH	
Radial Basis Functions	Dataset 2		
Local Polynomial Interpolation	Source Dataset	<none></none>	
Geostatistical methods	Dataset 3		
Kriging / CoKriging	Source Dataset	<none></none>	
Areal Interpolation Empirical Bayesian Kriging	Dataset 4		
	Source Dataset	<none></none>	and a second
iging / CoKriging			
ging is an interpolator that can be ex oss-correlation. Kriging uses statistical ging can require a lot of decision-maki	act or smoothed depending on I models that allow a variety of ing. Kriging assumes the data (	the measurement error model. It is very flexible and allows you to investigate graphs of spati f output surfaces including predictions, prediction standard errors, probability and quantile. The come from a stationary stochastic process, and some methods assume normally-distributed dated and the state of	al auto- and e flexibility of ta.
out Kriging / CoKriging			

The kriging interpolation method was selected based on its capability to determine the corresponding value of the attribute data in an unsampled location. The kriging method is also very flexible and allows you to investigate the graph of auto-and cross-correlation. In addition, the kriging method uses statistical models that allow a variety of output surface including predictions, standard error, probability and quantile. More also, unlike other methods of interpolation that assumes that the input data are normally distributed, kriging interpolation method recognizes the stochastic nature of the input data.

#### 2.5.3 Fitting and testing of semivariogram

The semivariogram/covariance model allows you to select the model that best described the groundwater quality data. In addition, the nugget, range and partial sill can also be determined. The interphase of Arcmap showing the semivariogram/covariance modelling of pH is presented in Figure 6



Figure 6: Semivariogram/covarince modelling of pH

# 2.5.4 Estimation of cross validation statistics

The Cross Validation statistics gives you an idea of how well the model predicts the values at the unknown locations. The essence of cross validation was to examine closely the fitted models and select the model that gives the best result (prediction). To select the model that gives the best result, goodness of fit statistics, namely; Root mean square error (RMSE), Mean square error (MSE), Root mean square standard error (RMSSE) and Average standard error (ASE) were employed

# 3. Results and Discussion

Principal component analysis (PCA) was performed to identify the most critical water quality parameter (s) that significantly influenced the overall quality of groundwater within the study area. To perform the analysis, water samples were collected from hundred (100) boreholes and twenty one (21) water quality parameters were monitored during dry and wet season. These water quality parameters include; pH, Electrical conductivity (EC), Salinity, Total Dissolve Solids (TDS), Dissolved Oxygen (DO), Bicarbonate (HCO<sub>3</sub>), Sodium (Na), Potassium(K), Calcium (Ca), Magnesium (Mg), Chloride (Cl<sup>-</sup>), Phosphate (PO<sub>4</sub>), Nitrate (NO<sub>3</sub>), Sulphate (SO<sub>4</sub>), Iron (Fe), Zinc (Zn), Copper (Cu), Turbidity, Total suspended solid (TSS), Temperature and Alkalinity.

# 3.1 Testing the suitability of PCA

To ascertain the potential of principal component method in handling the specific objectives of this study, Kaiser-Meyer-Olkin measure of sampling adequacy and anti-image correlation matrix were employed. The estimated statistics based on Kaiser-Meyer-Olkin measure of sampling adequacy is presented in Tables 3 **Table 3: Testing the suitability of principal component analysis** 

Kaiser-Meyer-Olkin Mea	asure of Sampling Adequacy.	.923
Bartlett's Test of	Approx. Chi-Square	2.935E3
Sphericity	df	210
	Sig.	.000

From the result of Table 3, it was observed that the significant value based on Bartletts test of sphericity was 0.000. Since the calculated (p-value) was less than 0.05, it was concluded that the model is significant and that principal component analysis (PCA) was suitable for the analysis. A further test of suitability of principal component analysis is the use of anti-image correlation matrix presented in Table 4

# Table 4: Anti-image correlation matrix

<u>Eile Edit View D</u> ata <u>T</u> ransform	Ele Edit View Data Transform Insert Format Analyze Graphs Lätties Add-gns Window Help											
🔶 🖶 🙆 🔕 🔒 🖬 🔶	💷 💒 📑 📀 🌑 🚛 🖷 🖼 📼	+++	+ -	00 💆 🛬 📮	) 📱							
}eneral Linear Model ▲									CI-4iI			
A Notes										COLLEISUOLU	Mau Ix-	
Active Dataset							Total					
Warnings		ъЦ	Nitroto	Electrical	Turbidity	Dissolved	Dissoved	Qodium	Quinhoto	Zine	Connor	Chlorida
Between-Subjects Factors	Correlation pH	1 000	- 544	- 168	135	552	00100	- 518	- 516	- 424	- 423	- 507
a Descriptive Statistics	Nitrate	544	1.000	205	100	005	.021	702	701	720	750	592
🗃 Box's Test of Equality of Cova	Electrical Conductivity	160	205	1 000	114	.003	.221	476	420	276	272	200.
🍘 Multivariate Tests	Turbidity	125	. 100	- 114	1.000	190	- 059	.473	.435	. 169	.077	- 150
Levene's Test of Equality of E	Dissolved Owner	.155	133	114	100	1.000	030	101	774	722	71.4	150
Tests of Between-Subjects E	Total Discoved Solide	.002	000	441	.130	1.000	270	007	//1	700	7.1.9	090
Parameter Estimates		.027	.221	.871	058	276	1.000	.234	.207	.201	.217	.107
General Estimable Function	Sulphoto	518	.702	.475	161	667	.234	1.000	.690	.646	.589	.641
E Lack of Fit	Zine	516	.781	.439	182	//1	.207	.690	1.000	.687	./1/	.648
	Zinc	424	.720	.375	159	733	.201	.646	.687	1.000	.735	.557
inulivariate rests	Copper	423	.759	.372	077	714	.217	.589	.717	.735	1.000	.496
Jog Jactor Analysia	Chloride	507	.582	.366	150	595	.167	.641	.648	.557	.496	1.000
商 Title	Iron	073	.272	.227	.001	267	.133	.223	.289	.219	.282	.161
R Notes	Carbonate	317	.603	.472	101	558	.254	.596	.622	.558	.614	.470
Active Dataset	Total Suspended Solids	.141	123	.072	018	.128	.071	069	150	148	059	124
Descriptive Statistics	Phosphate	492	.597	.401	194	605	.251	.643	.603	.505	.490	.498
a Correlation Matrix	Temperature	241	.356	.195	.062	336	.175	.279	.340	.295	.393	.232
🗿 Inverse of Correlation Matrix	Alkalinity	355	.624	.347	172	628	.136	.606	.640	.558	.586	.534
🗃 KMO and Bartlett's Test	Salinity	157	.228	.106	066	242	.051	.206	.227	.213	.188	.208
i Anti-image Matrices	Magnesium	497	.813	.480	145	790	.256	.699	.815	.677	.701	.577
👰 Communalities	Potassium	466	.710	.374	151	698	.143	.631	.788	.639	.644	.569
La Total Variance Explained	Calcium	263	.509	.334	.034	469	.162	.470	.542	.387	.452	.386

Result of Table 4 shows that principal component analysis is suitable for this analysis. The suitability of PCA is based on the fact that; the off diagonal matrix are less than one with a host of them very close to zero. Since the off diagonal matrix are less than one, it was concluded that PCA is good for the study.

# **3.2 Extraction of communualities**

It is also important in the extraction phase to examine the communualities. Communalities representes the sum of

the square loading for a variable across factor. Communality can range from 0 to 1. A communality of 1 means that all of the variation in the groundwater quality is explained by the component factor. The computed communualities is presented in Table 5

<u>F</u> ile <u>E</u> dit <u>∨</u> iew <u>D</u> ata <u>T</u> ransform	İnsert F <u>o</u> rmat <u>A</u> nalyze <u>G</u> r	aphs <u>U</u> tilitie:	s Add- <u>o</u> ns	<u>W</u> indo <sup>.</sup>
🗁 🖬 🖻 🕼 📑 💼 🔶 📂	🔲 🥌 📭 🔐 🧐 🚛	📑 🖬 🖬 🖉	s 💠 🔶	-
)eneral Linear Model 📃	Comm			
🖭 Title		Initial	Extraction	
notes 👘 👘	рН	1.000	.568	
Active Dataset	Nitrate	1.000	.789	
🔐 Warnings	Electrical Conductivity	1.000	.934	
Between-Subjects Factors	Turbidity	1.000	.714	
Descriptive Statistics	Dissolved Oxygen	1.000	.774	
Multiveriete Teete	Total Dissoved Solids	1.000	.955	
Levene's Test of Equality of E	Sodium	1.000	.687	
Tests of Between-Subjects E	Sulphate	1 000	795	
Parameter Estimates	Zinc	1.000	.655	
🛱 General Estimable Function	Copper	1 000	682	
🧧 Lack of Fit	Chloride	1 000	564	
- 🖆 Title	Iron	1 000	652	
🖳 🖵 🗃 Multivariate Tests	Carbonate	1.000	711	
.og	Total Suspended Solids	1 000	744	
actor Analysis	Phosphate	1.000	565	
Title	Temperature	1.000	.303	
C Notes	Alkalinity	1.000	.042	
C Deservictive Ctetistics	Salinity	1.000	.025	
Correlation Matrix	Magnacium	1.000	.329	
Inverse of Correlation Matrix	Botoccium	1.000	.858	
KMO and Bartlett's Test	Colsium	1.000	.735	
Anti-image Matrices	Calcium		.706	l.
Communalities	Extraction Method: Princ	cipal Compoi	nent Analysis	-

It was observed from the result of Table 5 that the initial communualities was 1.000 for all the ground water quality parameters employed for this analysis. Communualities of 1.000 is good since it indicates that the variation in groundwater quality around the study area can be explained with the aid of principal component analysis. Extraction communualities are estimates of the variance in each variables accounted for by the factors. High extraction indicates that the extracted components represents the variables well. If any extraction is very low (< 0.3), then one may need to extract another component factor. Since non of the extraction is less than 0.3, it was concluded that the extracted components are good enough to explain the underlying variation associated with the quality of groundwater.

#### **3.3 Extraction of Component Factors**

Using the method of principal component analysis, result of the component factor extracted is presented in Table 6

Table 6: Extraction of compoent factors using principal componet analysis method (PCA)

<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>D</u> ata <u>T</u> ransform	Įna	sert F <u>o</u> rmat	<u>A</u> nalyze <u>G</u> r	aphs <u>U</u> tilities .	Add-ons Window	Help						
🗁 🖶 🔮 🚨 🔒 📅 🌧 📌		* 📭 🔋	ିତ୍ର 🌑 🚽	🖣 🖓 🗗	++++	- 00						
eneral Linear Model	eneral Linear Model Total Variance Explained											
Title				Initial Eigenvalu	ies	Extractio	n Sums of Square	ed Loadings	Rotation	n Sums of Square	d Loadings	1
🔚 Notes		Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	
Marnings		1	9.632	45.865	45.865	9.632	45.865	45.865	8.414	40.064	40.064	1
Between-Subjects Factors		2	1.723	8.203	54.067	1.723	8.203	54.067	2.047	9.745	49.810	L
🛱 Descriptive Statistics		3	1.295	6.168	60.235	1.295	6.168	60.235	1.959	9.327	59.137	L
🗃 Box's Test of Equality of Cova		4	1.025	4.882	65.117	1.025	4.882	65.117	1.153	5.488	64.625	L
i Multivariate Tests		5	1.011	4.816	69.933	1.011	4.816	69.933	1.115	5.308	69.933	L
Levene's Test of Equality of E		6	.907	4.321	74.254							L
Tests of Between-Subjects E		7	.769	3.663	77.917							L
Parameter Estimates		8	.714	3.400	81.317							L
General Estimable Function		9	.667	3.176	84.493							L
Title		10	.506	2.409	86.903							L
Multivariate Tests		11	.472	2.247	89.150							L
.oq		12	.387	1.845	90.995							L
'actor Analysis		13	.337	1.603	92.598							L
🖆 Title		14	.294	1.399	93.997							L
notes 🛛 🔂		15	.273	1.300	95.297							L
Active Dataset		16	.258	1.229	96.526							L
Descriptive Statistics		17	.217	1.033	97.560							L
Correlation Matrix		18	.182	.865	98.425							L
Inverse of Correlation Matrix		19	.154	.735	99.160							
Anti-image Matrices		20	.108	.514	99.674							1
Communalities		21	.068	.326	100.000							
Total Variance Explained		Extraction 1	lethod: Prin	cipal Component	Analysis.							-
🚡 Scree Plot												

The extraction analysis determines how well the component factors explains the variation in the overall quality of groundwater from the hundred boreholes using the total varianced explained as presented in Table 6. Factors with eigen values greater than one represent the number of component factors needed to describe the underlying variation of the ground water quality. This are the component factors that contributes an adequate amount to the variation in the ground water quality. The eigen value is normally used as cutoff in factor analysis since it is the sum of the squared factor loadings of all variables. Factors with eigen value less than one means that such factor do not have any positive influence on the overall groundwater quality. From the results of Table 6, it was observed that five component factors had eigen value greater than one. These component factors contributed to change in the quality of the groundwater aroung the study area. This claim was further supported using the scree plot presented in Figure 7



Figure 7: Scree plot of component factors with most significant contributions

From the scree plot of Figure 7, it was again observed that five component factors possess very strong influence on the overall quality of the groundwater around the study location.

# **3.4 Extraction of Component Matrix**

In other to identify the groundwater quality parameters that make up the members of each component factors, the component matrix was generated as presented in Table 7

# Table 7: Component matrix showing the five component factors

<u>E</u> ile <u>E</u> dit <u>V</u> iew <u>D</u> ata <u>T</u> ransform	n Insert F <u>o</u> rmat <u>A</u> nalyze <u>G</u> r	aphs <u>U</u> tilitie	es Add- <u>o</u> ns	Window	Help			
🗁 🖬 📇 🖪 📑 👘 🔿	- 🛄 🐜 📭 💊 🌒 🖧	- 🖷 🖬	🗈 🖶 🔶	•	- 🛍 🗖	🗖 📮 🖡		
}eneral Linear Model ▲	Component Matrix <sup>a</sup>							
🖶 Notes			1	Component				
Active Dataset		1	2	3	4	5		
🐼 Warnings	pH	587	.368	.209	127	.170		
La Between-Subjects Factors	Nitrate	.878	111	040	.025	048		
Descriptive Statistics	Electrical Conductivity	561	733	- 238	111	113		
Box's Test of Equality of Cova	Turbidity	- 193	059	568	589	063		
I over a'c Test of Equality of E	Dissolved Oxygen	- 868	070	104	- 063	012		
Tests of Retween-Subjects F	Total Dissoved Solids	326	840	- 274	227	128		
Parameter Estimates	Sodium	817	- 024	- 129	- 014	- 044		
General Estimable Function	Sulphate	888	- 199	- 012	010	011		
🐱 Lack of Fit	Zinc	794	- 103	- 087	077	- 018		
Title	Copper	808	- 041	001	094	- 117		
🛛 🖵 🏣 Multivariate Tests	Chloride	715	- 131	- 186	046	005		
.og	Iron	.715	217	524	.040	206		
actor Analysis	Carbonate	769	172	204	- 245	- 072		
🔄 Title	Total Suspended Solids	.730	206	.204	416	626		
💼 Notes	Phoenhate	110	.390	.030	410	020		
Active Dataset	Temperature	.121	010	174	004	032		
Correlation Matrix	Alkalinity	.410	.000	.240	.434	447		
Inverse of Correlation Matrix	Salinity	.770	094	.007	137	.041		
KMO and Bartlett's Test	Magnaeium	.265	128	198	003	.401		
Anti-image Matrices	Botocolum	.913	011	.145	023	.050		
Communalities	Calcium	.842	119	.096	033	037		
🗃 Total Variance Explained		.656	.074	.471	117	.189		
🚡 Scree Plot	Extraction Method: Princ	ipal Compo	nent Analysi:	s.				
🕋 Component Matrix	a. 5 components extrac	ted.						

The extracted component matrix was employed to understand the correlation between groundwater quality parameters in each component group. In regression terms, the component matrix is the standardized regression coefficient between the observed values and the component factors. Higher factor loading indicates that a parameter is closely associated with the component factor. To determine the water quality parameters that is most highly correlated with the component factors, horizontal decentralization of the component matrix was done and the best favoured parameters was selected as member of that particular component factor. Result of the horizontal decentralization of the component matrix is presented in Table 8

	Variable	S		Co	mponet F	actors		
S/No	Variable Code	Variable Name	1	2	3	4	5	
1	$X_1$	pН		0.368				
2	$X_2$	Nitrate	0.878					
3	X3	EC		0.753				
4	$X_4$	Turbidity				0.589		
5	X5	DO			0.104			
6	$X_6$	TDS		0.840				
7	X7	Sodium	0.817					
8	$X_8$	Sulphate	0.886					
9	X9	Zinc	0.794					
10	$X_{10}$	Copper	0.806					
11	X11	Chloride	0.715					
12	X <sub>12</sub>	Iron			0.534			
13	X <sub>13</sub>	Carbonate	0.758					
14	X14	TSS		0.396				
15	X15	Phosphate	0.727					
16	X16	Temperature				0.454		
17	X17	Alkalinity	0.770					
18	X18	Salinity					0.451	
19	X19	Magnesium	0.913					
20	X20	Potassium	0.842					
21	X <sub>21</sub>	Calcium	0.656					

	1	-		
Table 8: Resul	t of horizontal	decentralization	of the comm	onent matrix

Results of Table 8 revealed that:

i. The first component factor is most highly correlated with nitrate, sodium, sulphate, zinc, copper, chloride, carbonate, phosphate alkalinity, magnesium, potassium and calcium

ii. The second component factor is most highly correlated with pH, electrical conductivity and total suspended soilds.

iii. The third component factor is most highly correlated with dissolved oxygen and concentration of iron

iv. The fourth component factor is most highly correlated with turbidity and tenmperature while

v. The fifth component factor is most highly correlated with salinity

# **3.5 Extraction of Rotated Component Matrix**

The rotated component matrix is presented in Table 9

# **Table 9: Rotated component matrix**

<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>D</u> ata <u>T</u> ransform /	İnsert F <u>o</u> rmat <u>A</u> nalyze <u>G</u> ı	aphs <u>U</u> tilitie	es Add- <u>o</u> ns	Window	Help		
	🔲 🏪 📭 👔 🚱 🗣 🍕	1 🖷 🖌	🗈 🖶 🔶	• •	- 🗰 🗖	1 🛬 🚍	<b></b>
eneral Linear Model		Rotated C	omponent M	latrix"			
A Notes				Component	1		1
Active Dataset		1	2	3	4	5	1
🐼 Warnings	На	715	.119	.181	.100	.008	1
🗃 Between-Subjects Factors	Nitrate	.859	.116	.177	079	002	1
im Descriptive Statistics	Electrical Conductivity	.335	.891	.159	.014	038	1
🗃 Box's Test of Equality of Cova	Turbidity	245	038	.123	096	.792	1
Multivariate Tests	Dissolved Oxygen	841	185	135	.121	.012	1
La Levene's Test of Equality of E	Total Dissoved Solids	.098	.971	.042	.011	.038	1
Parameter Fetimetee	Sodium	.787	.199	.135	055	082	1
General Estimates	Sulphate	.844	.125	.232	117	007	1
E Lack of Fit	Zinc	.781	.135	.109	122	.008	1
	Copper	.777	.123	.208	.012	.141	1
Multivariate Tests	Chloride	.721	.120	.031	154	078	1
log	Iron	.118	.077	.795	009	.015	1
'actor Analysis	Carbonate	629	.195	492	.176	- 065	1
🖆 Title	Total Suspended Solids	- 104	101	069	830	- 170	1
notes	Phosphate	704	193	098	- 029	- 146	1
Active Dataset	Temperature	472	108	- 064	278	572	1
Descriptive Statistics	Alkalinity	709	052	325	- 072	- 089	1
Correlation Matrix	Salinity	197	082	083	- 491	- 199	1
Inverse of Correlation Matrix	Magnesium	014	156	.000		066	1
Anti imaga Matricoa	Potassium	.014	046	.402	003	.000	1
Communalities	Celcium	.003	.040	.232	044	.027	1
Total Variance Explained	Systemation Mothod: Brin	inal Compo	007.	.073	002	.149	1
Scree Plot	Rotation Method: Varin	nax with Kais	ser Normaliza	s. ation.			
Component Matrix	a. Rotation converged i	n 6 iterations	2				
Reproduced Correlations	a. Hotallon contorgoa l						

The rotated component matrix gives you the factor (s) you need to pay more attention to in other to improve the overall quality of the groundwater in any location. To understand the factors that are most critical towards improving the overall groundwater quality of the study area, vertical decentralization of the rotated matrix was done and result obtained is presented in Table 10

# Table 10: Result of vertical decentralization of the rotated component matrix

	Variables			Componet Factors							
S/No	Variable Code	Variable Name	1	2	3	4	5				
1	X1	pH									
2	X2	Nitrate	0.859								
3	X3	EC									
4	$X_4$	Turbidity					0.792				
5	X5	DO									
6	$X_6$	TDS		0.971							
7	X <sub>7</sub>	Sodium									
8	$X_8$	Sulphate									
9	X9	Zinc									
10	$X_{10}$	Copper									
11	X <sub>11</sub>	Chloride									
12	X <sub>12</sub>	Iron			0.795						
13	X <sub>13</sub>	Carbonate									
14	$X_{14}$	TSS				0.830					
15	X15	Phosphate									
16	$X_{16}$	Temperature									
17	X17	Alkalinity									
18	X <sub>18</sub>	Salinity									
19	X19	Magnesium									
20	$X_{20}$	Potassium									
21	X <sub>21</sub>	Calcium									

From the result of Table 10, it was observed that; nitrate, total dissolved solids, concentration of iron, total suspended solids and turbidity were the most important variables affecting the quality of the groundwater within the study area. The component plot in rotated space is presented in Figure 8



#### Figure 8: Component plot in rotated space

From the component plot of Figure 8, it was observed that the combined isolated variables is X3 and X6 which represent electrical conductivity and total dissolved solids an indication that both variables are not only highly correlated, they also jointly contributes to variation in the quality of groundwater within the study area.

The final result of principal component analysis revealed that; nitrate, total dissolved solids, concentration of iron, total suspended solids and turbidity were the most important variables affecting the quality of the groundwater. To study the spatial distribution of these water quality parameters around the study location, geospatial analysis was done using kriging method of interpolation. The input variables for spatial analysis include; coordinates of the sampled boreholes (Northing and Easting), and the attribute variables (concentration of iron, nitrate, total dissolved solid, total suspended solids and turbidity).

# 3.6 Exploring Data Distribution

Exploring the data helped examine how the data is distributed and also to examine the global trends associated with the data. Geostatistical analysis provides many data exploration tools that helped examine the distribution of observed water quality parameters. For this study, frequency histogram plots were done to verify the distribution of the critical water quality parameters identified by principal component analysis and results are presented in Figures 8a, 8b, 8c, 8d and 8e respectively



Figure 8a: Histogram plot of Nitrate

Tab	Table									Histogram 🛛						<b>E</b>					
*=	- 1	à •   🏪 🎙	g 🖸 🤞	×							1	Freq	quency · 10 <sup>-1</sup>					Count	: 60	Skewness	:-3.1988
Exp	ort_O	utput								3	×	<b>1</b>						Max	: 1.24	1-st Quar	tile : 0.92
	FID	Shape *	CODE	NORTHINGS	EASTING	DESCRIPTIO	NITRATE	IRON	TSS	TDS /								Mean	: 0.94454	Median	: 1.01
	0	Point ZM	1	239820	561471	lgbu Ahaoda	0.06	0.0001	0.0001	46.78	1	2.4						Std. De	v.:0.19584		die:1.03
	1	Point ZM	2	269110	525361	Mini Ama	20.2	1.01	0.0001	184.8	II.										
	2	Point ZM	3	235669	537656	Arukwo-Abua	23.4	1.03	2.33	224.4	II.										
	3	Point ZM	4	286341	528043	Bakana	18.95	1.22	0.0001	236.28	1	1.8								_	
	- 4	Point ZM	5	237214	556600	Edeoha-Ahoada	27.2	1.05	5.61	250.8	II.										
	5	Point ZM	6	236203	556600	Edeoha-Ahoada	33.21	1.07	0.0001	6.67	II.										
	6	Point ZM	7	235766	540433	Okoboh-Abua	29.83	1.08	0.0001	189.72	1	1.2									
	7	Point ZM	8	262207	524264	Buguma	31.02	1.05	0.01	277.2	L										
	8	Point ZM	9	280557	534103	Air force Base	27.08	1.01	0.0001	104.5	I.										
Ц	9	Point ZM	10	279389	530030	Trans Amadi	12.33	0.98	0.0001	102.2	I.	0.6									
н	10	Point ZM	11	274121	532098	lpo-lkwerre	16.9	1.009	0.01	117.3	I.										
н	11	Point ZM	12	286716	533642	Woji	18.54	1.08	0.0001	133.9	L	l i						_			
н	12	Point ZM	13	283293	536010	Rumuokwurushi (1)	22.03	0.76	0.0001	122.1	L	1.	0.12	0.25	0.37	0.5	0.62	0.74	0.87	0.99	1 12 1 24
н	13	Point ZM	14	286238	527163	Amakiri Polo	33.3	1.03	0.01	187.2	L	ľ	0.12	0.20	0.01	0.0	Datacel	0.74	0.07	0.00	1.12 1.24
н	14	Point ZM	15	289003	534162	Rukpokwu	27.88	1.05	0.0001	193.9	L	-					Dataset				
н	15	Point ZM	16	280451	526634	Aggrey	33.01	0.98	0.0001	123.1	L	Tip:	Click or o	drag over	bars to sele	ect				Add	to Layout
H	16	Point ZM	17	278741	529397	NDBDA	26.7	1.06	0.01	144.03	1	Bar	e:	10		Staticti	ice.				
$\square$	17	Point ZM	18	283012	536068	Rumuokwurushi (2)	27.09	1.01	0.0001	109.1	1	1.00				- 3tausu					
H	18	Point ZM	19	279849	530118	Amadi-Ama	31.33	0.11	0.0001	111.1	1		Transformatio	n							
$\vdash$	19	Point ZM	20	287302	531219	Owodu	34.3	1.02	0.01	109.4	1	Tran	nsformation:	None	-						
L.	20	Point ZM	21	307314	519241	Ukochin	29.09	1.03	0.001	121.5											
1	-									,	1		Data Source					de star			
14	4	0	F FI	(30)	out of 60 Sel	ected)					1	Laye					Atu	Dute:			
Eve	ort O	utout									1	Exp	port_Output				<ul> <li>IR0</li> </ul>	ON			-
ext	0120	uthor									1	_					_		100 11		
	_										-	-					<u>۱</u>	_	(F)	194 18	0212 0 00

Figure 8b: Histogram plot of Iron



Figure 8c: Histogram plot of total suspended solids



Figure 8d: Histogram plot of total dissolved solids



# Figure 8e: Histogram plot of turbidity

From the result of Table 8a, it was observed that; though the concentration of nitrate was below the maximum allowable limit, the highest concentration of nitrate was recorded in water samples collected from Mini-Ama, Bakana and Woji. From the result of Table 8b, it was observed that; the concentration of iron in some of the borehole sampled was higher than the maximum allowable limit. The highest concentration of iron was recorded in water samples collected from Owodu, Okochiri, Rumuokwurushi, Rukpokwu, Amakin-Polo, Air force base, Buguma and Amakiri-polo. From the result of Table 8c, it was observed that; though the concentration of total suspended solids was below the maximum allowable limit, the highest concentration of TSS was recorded in water samples collected from Mini-Ama, Bakana, Woji, Owodu, Okochiri, Rumuokwurushi, Rukpokwu, Amakin-Polo, Air force base, Buguma and Amakiri-polo. Others are; Edeoha-Ahoada, Okoboh-Abua, Trans-Amadi, Aggrey, Amadi-Ama and Owodu. From the result of Table 8d, it was observed that; though the concentration of total dissolved solids was below the maximum allowable limit, the highest concentration of TDS was recorded in water samples collected from Rumuokwurushi, Amadi-Ama, Owodu, Trans-Amadi, Oyigbo and Yegha-Gokona. From the result of Table 8e, it was observed that; though the turbidity of the water samples was below the maximum allowable limit, the highest turbidity was recorded in water samples collected from Mini-Ama, Bakana, and Woji. Owodu, Okochiri, Rumuokwurushi, Rukpokwu, Amakin-Polo, Air force base, Buguma and Amakiri-polo others are; Edeoha-Ahoada, Okoboh-Abua, Trans-Amadi, Aggrey, Amadi-Ama and Owodu.

The histogram of Figures 8a, 8b, 8c, 8d and 8e revealed that the parameters are not normally distributed. This further support the claim that seasonal variation exist among the water quality parameters. The results are in line with that of previous study conducted by (Hooshmand et al., 2011) in which two geostatistical methods, kriging and co-kriging, were applied to estimate chloride content and sodium adsorption ratio (SAR) of ground water in the Boukan area of Iran. The histogram plot generated for the chloride concentration around the study area did not also obey normality an indication that ground water quality parameters varies with depth, time and season. In addition, in the study by (Ahmadi et al., 2008) in which geostatistical methods, namely; kriging and co-kriging were applied to study the maximum, minimum and mean ground water depths of 39 wells, the ground water depths were all observed to be spatially correlated an indication that they are not normally distributed. The basic steps invloved in the development of the spatial distribution map of the critical water quality parameters around the study parameters around the study location using kriging interpolation method is presented as follows.

# 3.6.1: Fitting and testing of semivariogram

Eight semivariogram models, namely; Circular, Spherical, Exponential, Gaussian, Hole effect, K-Bessel, J-Bessel and Stable) were fitted for each of the five critical water quality parameters used for geostatistical analysis during wet season in order to select the best fitted model for the parameter. Table 11 shows the result of the semivariogram models for nitrate and the corresponding values of nugget (the variability in the field data that cannot be explained by distance between the observations), major range (represents the distance at which two observations are unrelated/independent) and sill (the semi-variance at which the leveling takes place. The end point of range is the beginning point of sill. The different between the sill and the nugget is called partial sill).

S/N	Model Type	Nugget	Major Range	Partial Sill
1.	Circular	25.88929	122473.6	58.77035
2.	Spherical	24.03614	122473.6	44.43717
3.	Exponential	23.88726	122473.6	29.36212
4.	Gaussian	25.88929	122473.6	58.77035
5.	Hole effect	25.73296	122473.6	30.75632
6.	K-Bessel	25.81664	122473.6	54.20273
7.	J-Bessel	25.74801	122473.6	33.8048
8.	Stable	24.02447	122473.6	38.82232

#### Table 11: Semivariogram models for Nitrate

Fitted Semivariogram models presented in Table 11 provided information about the range, nugget and partial sill (Model Parameters) which were used to measure the degree of spatial dependency of sampled borehole points as a result of distance between them. It also provides the input parameters that were utilized for the kriging interpolation. To select the model that best described each water quality parameters and which was employed to generate the final prediction map, selected goodness of fit statistics generated from the cross-validation step were employed.

# **3.6.2: Estimation of cross validation statistics**

To select the most suitable model needed to generate the final prediction map, selected goodness of fit statistics, namely; Root mean square error (RMSE), Mean square error (MSE), Root mean square standardized error (RMSSE) and Average standard error (ASE) were employed and the calculated values corresponding to the different models is presented in Table 12

S/N	Model Type	RMSE	MSE	RMSSE	ASE
1.	Circular	6.828	-0.00610	1.245	5.369
2.	Spherical	6.944	-0.00292	1.255	5.408
3.	Stable	4.949	-0.00252	1.255	5.415
4.	Exponential	6.972	-0.00067	1.241	5.496
5.	Gaussian	6.828	-0.00610	1.245	5.369
6.	Hole effect	6.858	-0.00343	1.250	5.367
7.	K-Bessel	6.842	-0.00439	1.247	5.369
8.	J-Bessel	6.860	-0.00286	1.250	5.370

#### Table 12: Result of cross validation statistics

Following the application of different models on each water quality parameter, the errors were calculated using cross validation and the model that gives the best result was chosen. The criterion for selecting the best model is as follows:

- i. The mean standardized error should be close to 0,
- ii. The root mean square error and average standard error should be as small as possible and close to each other
- iii. The root mean square standardized error should be close to 1.

Based on the above criterion, and examining the results of Table 12, the stable model was selected as the best fit model for each selected water quality test parameters.

# **3.6.3: Examination of spatial dependency**

Fitted semi variogram chart can be used to examine the spatial dependence between the measured sample points. Based on the results of Tables 11 and 12, the stable model was selected as the best fit model and a summary Table was generated and presented in Table 13

Tal	hle	13:	Summary	table f	for	estimating	snatial	dependence
1	UIC	10.	Summary	tubic i		countaunt	Spatial	acpendence

S/N	Parameter	Best Model	Nugget	Major Range	Partial Sill
1.	Nitrate	Stable	24.02447	122473.6	38.82232
2.	Iron	Stable	26.0645	10564.9	19.0956
3.	TSS	Stable	3.44782	33089.6	55.0934
4.	TDS	Stable	13.4667	9.04586	6.78023
5.	Turbidity	Stable	1.0456	107783.3	32.9067

The sill (c) is the summation of nugget and partial sill while the ratio of Nugget to Sill  $(\frac{Cn}{C})$  was employed to

measure the degree of spatial structure (dependence) of a water quality parameter. If the ratio is less than 25%, the variable has strong partial dependence; between 25% and 75%, the variable has moderate spatial dependence, and greater than 75%, the variable shows only weak spatial dependence. The computed spatial dependence for the selected water quality parameters using the selected best fit model is presented in Table 14

Parameter	Best	Nugget	Partial	Sill (C)	[Cn/C]	Degree of Spatial
	Model	(C <sub>n</sub> )	Sill			Dependency
Nitrate	Stable	24.02447	38.82232	62.84679	0.38227	Strong
Iron	Stable	26.0645	19.0956	45.1601	0.577158	Strong
TSS	Stable	3.44782	55.0934	58.54122	0.058896	Strong
TDS	Stable	13.4667	6.78023	20.24693	0.665123	Strong
Turbidity	Stable	1.0456	32.9067	33.9523	0.030796	Strong

Results of Table 14 revealed that the water quality parameters showed relatively strong degree of spatial dependency which made it possible to generate the spatial distribution map for the selected water quality parameters.

# 3.6.4: Generation of water quality parameters distribution map

Finally the prediction map which can be employed to predict nitrate, concentration of iron, total suspended soilds (TSS), total dissolved solids (TDS) and turbidity of unsampled boreholes within the study area was generated and presented in Figures 9a to 9e respectively



Figure 9a: Final prediction map for Nitrate distribution



Figure 9b: Final prediction map for Iron distribution



Figure 9c: Final prediction map for TSS distribution



Figure 9d: Final prediction map for TDS distribution



Figure 9e: Final prediction map for Turbidity distribution

It was observed from the prediction map of Figures 9a to 9e that areas with red colour codes represent higher

values of Nitrate, Iron, TSS, TDS and Turbidity while areas with blue colour codes represent lower values of Nitrate, Iron, TSS, TDS and Turbidity. To validate the result of this study, a comparison was made with the results of previous studies as follows; in a study by Ramesh and Elango (2013). The study was carried out to analyze the spatial variability of groundwater quality for the region using samples collected from twenty-three (23) boreholes. The parameters analyzed included pH, EC, TDS, TH, Calcium, Magnesium, Bicarbonate, Chloride and Sulphate which were compared with permissible limits. Spatial interpolation technique using IDW approach was used to delineate the distribution of water pollutants. Results of the study revealed that the stable and K-Bessel model was the best fit model which is similar to the results of this study.

Research work done by Balakrishnan et al. (2011) on the use of GIS to map groundwater quality variation in Gulbarga City Karnataka India shows that GIS is a very powerful tool for investigating groundwater quality information. The study sampled 76(seventy-six) bore wells and open wells representing the entire corporation area. The analysis was based on the physico-chemical parameters which were restricted to measurement of hardness, salinity and determination of potential contamination by sewage. Spatial interpolation (IDW) techniques was used to delineate the locational distribution of groundwater pollutants and through GIS, overlay map of the critical parameters were also produced. Results obtained revealed a very strong spatial dependency of the measured parameters with a calculated (nugget to sill ratio) of less than 25% which is similar with the results of this study

# 4. Conclusion

The study was conducted to determine the critical contaminants affecting the overall quality of groundwater around the study area and develop a spatial map using kriging interpolation to visualize the distribution of these contaminants. Results of the study attest to the potential of factor analysis using principal component techniques in the assessment of water quality parameters with a view of selecting the parameters with critical influence on the water quality. More also, geospatial analysis using kriging interpolation revealed that; the water quality parameters showed relatively strong degree of spatial dependency which made it possible to generate the spatial distribution map for the selected water quality parameters.

# 5. References

- 1. Abams, T.K.S, Olu, A.W and Nwankwoala, H (2007): Groundwater Monitoring for Environmental Liability Assessment, J. Nig. Environ. Soc., vol. 4(1), 42-49
- 2. Abgalagba, O.E, Agbalagba, O.H, Ononugbo, C. P and Alao A.A (2011): Investigation into the Physicochemical Properties and Hydro-chemical Processes of Groundwater from Commercial Boreholes in Yenagoa, Bayelsa State, Nigeria. African Journal of Environmental Science and Technology Vol. 5(7), 473-481.
- 3. Ahmadi, S. H., and Sedghamiz, A (2008): "Application and evaluation of kriging and co-kriging methods on groundwater depth mapping". Environmental Monitoring and Assessment, vol. 138, 357-368.
- 4. Akporido, C.E., (2000). "A study of the Characteristics of Surface and Underground Water in Sapele and Ogharefe Oil Producing Areas of Delta State", The Nigeria J. Sci. Environ., 5(7): 65–75
- 5. APHA (2005), Standard Methods for the Examination of Water and Wastewater, 18th edition, American Public Health Association Washington D C
- 6. Baalousha, H. (2010): Assessment of a groundwater quality monitoring network using vulnerability mapping and Geostatistics; A case study from Heretaunga Plains, New Zealand. Agriculture and Water Management Journal, vol. 97, 240–246.
- Balakrishnan, P; Saleem, A; and Mallikarjun, N.D (2011): Groundwater quality mapping using geographic information systems (GIS); A case study of Gulbarga City, Karnataka, India, African Journal of Environmental Science and Technology, vol. 5(12), 1069 – 1084
- 8. Debels, P., Figueroa, R., Urrutia, R., Barra, R., and Niell, X., (2005), Evaluation of water quality in the Chillan River (Central Chile) using physicochemical parameters and a modified water quality index, Environmental Monitoring and Assessment, vol. 110, 301–322.
- 9. Efe, S.I., (2003). Water Quality and its Utilization in the Nigeria Rural Setting of Abraka, Delta State, Nigeria, the environmentalist Journal, vol. 26: 175 181
- 10. Faisal, K., Tahir, H. and Ashok, L., (2003), Water quality evaluation and trend analysis in selected watersheds of the Atlantic region of Canada, Environment Monitoring Assessment, vol. 88, 221-248.
- 11. Gajbhiye, S; Sharma S. K and Jha, M (2014): "Application of Principal Component Analysis in the assessment of water quality parameters", Science-fronts: A Journal of Multiple Science, vol. 4, 67-72.
- Hooshmand, A; Delghandi, M; Izadi, A and Aali, A. (2011): "Application of kriging and co-kriging in spatial estimation of groundwater quality parameters". African Journal of Agricultural Research, vol. 6(14), 3402-3408.
- Nichole, E.T., (2004), Water and Life, Journal of the Institute of Water & Environmental Management, vol. 3, 116-125
- 14. Nwankwoala, H.O.; Eludoyin, O.S. and Obafemi, A.A. (2012): "Groundwater Quality Assessment and

Monitoring Using Geographic Information Systems (GIS) In Port Harcourt, Nigeria." Ethiopian Journal of Environmental Studies and Management (EJEM) Vol. 5(4), 83-5967.

- Ocheri, M.I.; Odoma, L.A. and Umar, N.D. (2014): "Groundwater Quality in Nigeria Urban Areas: A Review". Global Journal of Science Frontier Research: Environment and Earth Science. Vol. 14, Issue 3, version 1.0, ISSN 2249-4626
- 16. Ramesh, K. and Elango, L. (2006): "Groundwater Quality Assessment in Tondiar Basin". International Journal of Environmental Pollution Vol. 26(6), 497-504
- 17. Water for Life (2007): http://www.Un.Org/works/water/index.html
- World Health Organization (2003): "Manual of Basic Technique for a Health laboratory" World Health Organization, 2<sup>nd</sup> Edition, Geneva.