ASSESSMENT OF GROUNDWATER QUALITY PARAMETERS OF JAEN DISTRICT, KANO STATE, NIGERIA

[1*]AHMAD ABUBAKAR SULEIMAN, [1]USMAN ALIYU ABDULLAHI, [1]AMINU SULEIMAN, [2]SULEIMAN ABUBAKAR SULEIMAN, [3]BINTA ABUBAKAR, [4]TASI'U MUHAMMAD, [4]NI'IMATULLAH KABIR ISAH, [5]ABDURRAHMAN HUSSAINI TAFIDA

[1]Department of Statistics, Kano University of Science and Technology, Wudil, Nigeria
[2]Hydrology Unit, Kano State Nigeria Erosion and Watershed Management Project (Kano - NEWMAP), Kano State, Nigeria
[3]Department of Science Laboratory Technology, Federal Polytechnic Kaltungo, Gombe State, Nigeria
[4]Department of Geography, Federal University of Lafia, Nassarawa State, Nigeria
[5]Admin Department, Kano Agricultural Supply Company, Kano State, Nigeria
[6]Diretorate of Academic Planning, Sule Lamino University Kafin Hausa, Jigawa State, Nigeria

[*]Corresponding Author: ahmadabubakar31@gmail.com

ABSTRACT

This study is based on physical and chemical parameters to assess groundwater quality of Jaen District, Kano state, Nigeria. Thirty groundwater samples were collected from different locations of the study area. The physical and chemical parameters such as electrical conductivity (EC), pH, total dissolved solid (TDS), carbonate (CO_3), magnesium (Mg), sodium (Na), potassium (K), calcium (Ca), bicarbonate (HCO_3), chloride (Cl), sulphate (SO_4), fluoride (F), nitrate (NO_3), and total hardness (TH) have been determined for drinking purposes. The correlation coefficients of different parameters were obtained. A linear regression equation was used to predict concentration of groundwater parameters that were found to have a higher level of significance in their correlation coefficients. Highly positive correlation coefficients are observed between EC and TDS; Mg and CO_3; Ca and CO_3; HCO_3 and CO_3; Ca and Mg; HCO_3 and Mg; Cl and Na; Ca and HCO_3; NO_3 and Cl. The results indicate that all the regression equations obtained are significant at 1%. The use of these linear regression equations in assessing groundwater quality could serve as an alternative method of monitoring processes.

Keywords: Jaen district, Groundwater parameters, Correlation coefficient, Regression analysis, Pollution

INTRODUCTION

Groundwater quality is the main water resource challenge as a consequence of rapid population growth, fast industrialization, urbanization, contamination movement from upland to lowland, and excessive use of fertilizers and pesticides in agriculture (Joarder *et al.,* 2008; Ibrahim *et al.,* 2021). Thus, increasing population, industrialization, and urbanization are the most important causes contributing to increased water contamination. The contamination in groundwater may spread waterborne diseases, kidney failure, gastroenteritis, maternal and infant mortality, cholera, typhoid fever, and giardiasis (Taiwo *et al*., 2015). Therefore, water quality monitoring is essential for the wellbeing of all people. Many treatment procedures are used to improve the quality of groundwater. Water should be free of organic and inorganic pollutants, pesticides, heavy metals, and other contaminants. All of its parameters should be within acceptable limits, including fluoride, iron, manganese, nitrate, pH, turbidity, alkalinity, chloride, total hardness, calcium hardness, and magnesium hardness (Dutta and Sarma, 2018**).**

The problems of groundwater quality are greater in industrial areas like Jaen in Sharada Phase III, one of the major industrial sites in Nigeria. Some of the industrial activities in the Jaen district include textile finishing, metal production, detergent and soap manufacturing, paper production, and food processing. Most of the residents in this area use groundwater as the major source of drinking water and for domestic activities. A few studies were carried out to monitor the quality of groundwater in the area. For example, previous studies (Mustapha, 2015; Suleiman *et al.,* 2020a) reported that groundwater contamination in the area was due to anthropogenic and industrial activities. Groundwater contamination in the area is also due to the transport of hazardous and harmful materials produced by residential and industrial activities by leaching into the underground within the area, making it vulnerable to the dangers of polluted water from industrial sources.

Investigations into the quality of drinking water samples have been continuously performed by researchers around the world (Heydari *et al.,* 2013). In recent years, only a few approaches for determining water quality have been used, such as the projection pursuit technique and neural networks (Salman and Ruka'h 1999). Water Quality Index has also been considered to obtain information about water quality (Sadat-Noori *et al.,* 2014). Water quality data has been modeled using various statistical approaches (Juahir *et al.,* 2011; Shrestha and Kazama 2007; Suleiman *et al.,* 2020b; Ibrahim *et al.,* 2021). Multivariate statistical techniques have been used for analyzing complex water quality datasets with minimal loss of significant information (Lee *et al.,* 2001; Ravikumar and Somashekar 2017; Gulgundi and Shetty 2018). These approaches take time to estimate pollutants in groundwater, and when human health is at stake, instant information about water parameter(s) is required.

Regression equations can be used to estimate concentrations of pollutants in the shortest period of time. Thus, this technique will help determine the pollutants' levels in the water quality monitoring process and also serve as a powerful tool for understanding and identification of possible pollutants affecting water quality datasets with optimum cost. In the present research, correlation and regression techniques have been applied to estimate the water quality parameters in the Jaen district.

MATERIALS AND METHODS

Study Area

Jaen is a District in Gwale local government area within Kano city, Nigeria. It is situated at a latitude $11^{\circ}58'N$ and longitude $8^{\circ}30'E$. Because of the research area's climatic conditions, which include seven months of dry weather and five months of rainy weather, and the fact that the entire study area is built on a basement rock complex, tube wells, open wells, and hand pumps are used to provide the huge volumes of water required for home and industrial usage in the study region. The sampling locations of the study area are shown in Figure 1.

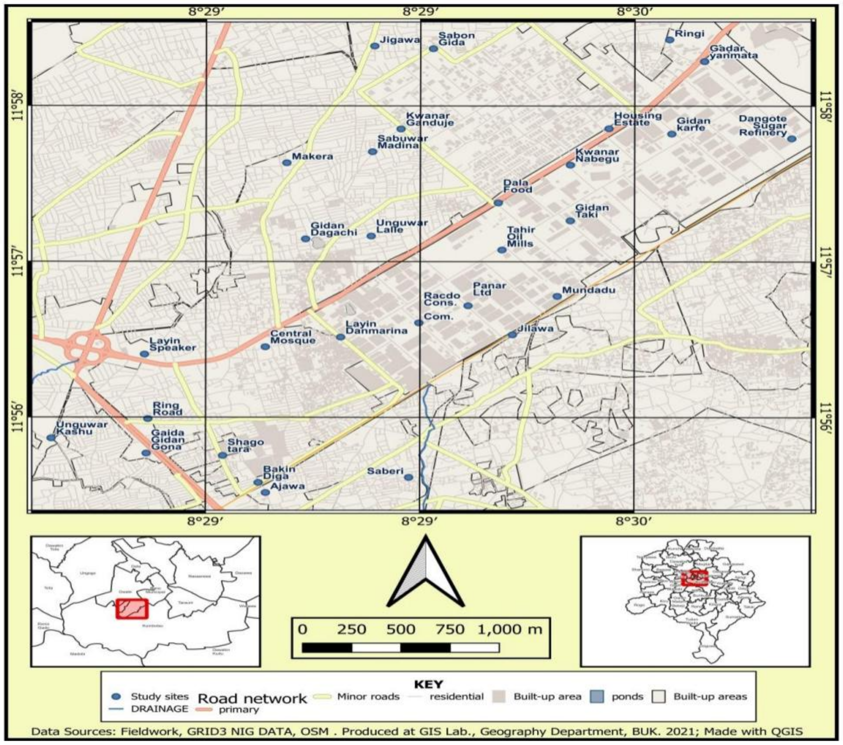

Figure 1: Study area map with sampling stations

Sample Collection and Analysis

Thirty groundwater samples were collected from various locations within the study area in the month of August (2020). To establish a good sampling representation for estimating the quality of groundwater in the study area, the sampling sites were carefully selected with the purpose of covering residential and industrial areas. The latitude and longitude of sampling points were determined on a map using the Geographical Position System (GPS). Water samples were taken from hand pumps and open wells, then poured into cleaned plastic bottles and stored in an ice box in accordance with the standard method (American Public Health Association, 2005). The collected samples were sent to the laboratory where they were analyzed for electrical conductivity, pH, total dissolved solid, carbonate, magnesium, sodium, potassium, calcium, bicarbonate, chloride, sulphate, fluoride, nitrate, and total hardness. Electric conductivity, pH and total dissolved solid were measured and recorded in their respective stations immediately after sampling. Most parameters are expressed in milligram per liter (mg/L), except EC (µS/cm), pH and TDS (NTU).

Correlation Analysis

Correlation is a statistical tool that studies the relationship between two variables. Correlation analysis involves various methods and techniques used for studying and measuring the extent of the relationship between two or more variables. The degree to which the two variables are interrelated is measured by the coefficient of correlation. It is denoted by r_{XY} and is defined as provided in Equation (1).

$$r_{XY} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{n}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{n}\right]}} \quad (1)$$

where X and Y are the two variables, n is the number of pairs of observations. The value of r_{XY} ranges from -1 to +1, or near to one indicates a strongest positive linear correlation between two parameters compared and $r_{XY} = -1$ or near to -1 reveals strongest negative linear correlation. The details on correlation analysis can be found in this paper (Suleiman *et al.,* 2020a). In this paper, we applied the correlation coefficient to measure the interrelated variation among the groundwater parameters.

Regression Analysis

In groundwater, a relationship exists among many parameters. For example, the relationship between the sodium concentration of a given sample and that of chloride may result in sodium chloride (NaCl). Regression analysis enables us to find a mathematical equation for the relationship between two or more variables as well as use this equation to make predictions about one variable (Suleiman, 2015). In regression analysis, there are two types of variables: independent variable and dependent variable. The independent variable is also known as a regressor, predictor, or explanatory, while the dependent variable is called a regressed or explained variable. The regression adjusted R^2 tells us whether regressors are good at predicting or explaining the values of the dependent variable in the sample of the data at hand (Joarder *et al.,* 2008). The values of adjusted R^2 (0.91–0.99) usually quantify a tight linear trend and could be an effective parameter in estimating a good fit for the model (Suleiman *et al.,* 2021a). The mathematical representation of the regression model is given in Equation (2).

$$Y = a + bX \quad (2)$$

where Y and X are the respective dependent and independent variables of the regression model, a is the intercept of the model and b is the slope of the model. The values of a and b are obtained as given in Equation (3) and Equation (4), respectively.

$$a = \frac{\sum Y}{n} - \frac{b\sum X}{n} \quad (3)$$

Similarly,

$$b = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2} \quad (4)$$

The details of these equations were provided in (Dalah *et al.,* 2020). All computations in this study were performed using Minitab 16.

RESULTS AND DISCUSSION

The data of groundwater samples collected from different locations of the study area are summarized in Table 1, Table 2, Table 3 below, which included parameters of electrical conductivity (EC), pH, total dissolved solid (TDS), carbonate (CO_3), magnesium (Mg), sodium (Na), potassium (K), calcium (Ca), bicarbonate (HCO_3), chloride (Cl), sulphate (SO_4), fluoride (F), nitrate (NO_3), and total hardness (TH).

Table 1 : Descriptive statistics of the groundwater parameters

	Concentration			
Parameter	Min.	Max.	Mean	Std. Dev.
EC	209.0	1490.0	672.7	302.6
pH	5.6	7.2	66713	0.4114
TDS	104.0	731.0	339.2	151.3
CO3	0.180	4.380	1.491	1.233
Mg	0.0300	0.7400	0.2514	0.2076
Na	0.00	32.00	7.50	8.02
K	0.1000	1.3000	0.4200	0.3566
Ca	0.5893	0.4964	0.5893	0.4964
HCO3	0.220	5.340	1.255	1.212
Cl	0.000	4.600	1.262	1.176
SO4	6.00	55.00	22.30	10.75
F	0.00000	0.03000	0.00800	0.00761
NO3	0.0000	0.9300	0.1957	0.2214
TH	0.01000	0.04000	0.01733	0.00828

Table 2: Correlation coefficients among different groundwater quality parameters

Parameter	EC	pH	TDS	CO3	Mg	Na	K	Ca	HCO3	Cl	SO4	F	NO3	TH
EC	1													
pH	-.30	1												
TDS	**.99**	-.31	1											
CO3	-.16	.03	-.16	1										
Mg	-.16	.02	-.16	**1.00**	1									
Na	-.07	.03	-.07	-.20	-.21	1								
K	-.03	-.03	-.03	-.17	-.17	.27	1							
Ca	-.16	.01	-.16	**.99**	**.99**	-.19	-.17	1						
HCO3	-.19	-.21	-.19	**.81**	**.81**	-.22	-.19	**.81**	1					
Cl	.02	.16	.03	-.07	-.07	**.79**	.15	-.06	-.01	1				
SO4	-.18	.31	-.18	-.13	-.13	-.03	.04	-.13	-.17	-.13	1			
F	-.17	.09	.08	-.09	-.09	-.40	-.15	-.09	.01	-.17	.12	1		
NO3	.53	.02	.17	-.34	-.35	.48	.26	-.34	-.26	**.53**	.24	-.07	1	
TH	.21	.17	-.19	.45	.46	.19	-.15	.42	.23	.24	-.13	-.25	.02	1

Table 3: Correlation coefficients, regression equations and adjusted R^2 for different pairs of parameters which have significant value of correlation

Pairs of parameters	Correlation Coef. Value	Regression Equation	Adjusted R^2 value	p – value
EC – TDS	0.994	$EC = -5.45 + 2.00\,TDS$	0.999	0.000
Mg – CO_3	1.000	$Mg = 0.000273 + 0.168\,CO_3$	1.000	0.000
Ca – CO_3	0.994	$Ca = -0.0093 + 0.401\,CO_3$	0.993	0.000
HCO_3 – CO_3	0.810	$HCO_3 = 0.067 + 0.797\,CO_3$	0.644	0.000
Ca – Mg	0.994	$Ca = -0.0098 + 2.38\,Mg$	0.993	0.000
HCO_3 – Mg	0.813	$HCO_3 = 0.062 + 4.75\,Mg$	0.649	0.000
Cl – Na	0.793	$Cl = 0.390 + 0.116\,Na$	0.616	0.000
HCO_3 – Ca	0.812	$HCO_3 = 0.086 + 1.98\,Ca$	0.648	0.000
NO_3 – Cl	0.528	$NO_3 = 0.0700 + 0.0995\,Cl$	0.254	0.003

The descriptive statistics of the different physicochemical parameters of groundwater constituents such as electrical conductivity (EC), pH, total dissolved solid (TDS), carbonate (CO_3), magnesium (Mg), sodium (Na), potassium (K), calcium (Ca), bicarbonate (HCO_3), chloride (Cl), sulphate (SO_4), fluoride (F), nitrate (NO_3), and total hardness (TH) in Jaen district for drinking purposes are given in Table 1. The EC value in the study area lies in the range of 209 to 1490, indicating high mineralization in the study area for drinking purposes. The high EC value exhibited by these samples may be due to dissolution of ionic heavy metals from manufacturing activities of heavy machines that later found their way into groundwater via leaching of sub-soil layers (Eruola and Adedokun, 2012). The pH varies from 5.6 to 7.2, indicating the slightly alkaline nature of groundwater. The TDS value ranged from 104 in Ajawa to 731 in Jigawa. Some of the groundwater samples show TDS values below and above the permissible limit for drinking, i.e., 500 mg/L. A higher TDS value is generally harmful to those suffering from kidney and heart disease (Shrestha and Basnet, 2018).The maximum TDS values revealed by these locations are an indicator of salty water that can be due to the discharge from industrial treatment plants causing leaching of soil contamination and pollution of groundwater by point-source. This is in agreement with our early study in the area (Suleiman *et al.,* 2020a).

The novel calculation of the correlation coefficient and regression analysis provide a systematic way for the rapid estimation of water quality. The correlation matrix for different groundwater quality parameters for the Jaen district is depicted in Table 2. The level of significance is taken at 5%. A perfect positive correlation coefficient is obtained between Mg and CO_3 (1.00), while a highly positive correlation coefficient is obtained between EC and TDS (0.99), Ca and CO_3 (0.99), Ca and Mg (0.99), HCO_3 and Mg (0.81), Ca and HCO_3 (0.81), HCO_3 and CO_3 (0.81), Cl and Na (0.79), NO_3 and Cl (0.53). Magnesium, carbonate and bicarbonate are strongly correlated with calcium, and this can be attributed to the dissolution of salts and inorganic pollution load in the water, which indicates industrial activities were responsible for contamination of the assessed groundwater parameters in the study area (Suleiman *et al.,* 2020a). The high values of correlation coefficient of carbonates and bicarbonates are the main cause of alkalinity in the groundwater. A highly positive correlation between sodium and chloride indicates that the groundwater in our study area tends to produce common salt, which may induce hypertension in the human brain if consumed in excess (EPA, 2001). Nitrate shows a significant positive correlation with chloride, and is can be explained as the soil present in this area mainly constitute high concentration of domestic waste and thus soil consist of high nitrate concentration. Hence, it is evident that the correlation study of the groundwater parameters is essential in the monitoring of water resources. After developing the correlation matrix, scatter plots for significantly correlated parameters were constructed as shown in Figures 2-10, respectively. In these figures, dependent variables were plotted along the X-axis and independent variables along the Y-axis. All the plotted graphs revealed a direct linear positive relationship between the groundwater parameters.

Based on the information given in Table 2, a simple linear regression analysis has been performed for the groundwater parameters having correlation coefficient values greater than 0.5. The different dependent groundwater parameters were obtained using the regression equations by substituting the values of the independent parameters in the equations as shown in Table 3. We provide in the first column the pairs of significant parameters. In the second column we present the values of significant correlation coefficient as depicted in Table 2. In the third column the results for simple linear regressions are presented. All the regression equations computed are significant at 1%. The fourth column gives the results of adjusted R^2 .The adjusted R^2 tells us whether the independent variable is good at explaining or predicting the values of dependent variable in the samples (Bhatnagar and Devi, 2012). The regression equations were evaluated using adjusted R^2 to determine the proportion of variability in the estimated groundwater parameters explained by the regression equations. The adjusted R^2 values ranged from 0.254 for nitrate to 1.00 for magnesium concentration. The adjusted R^2 values of different groundwater quality parameters are listed in Table 3. It is evident that carbonate is the most appropriate parameter predicting or explaining 100 and 99 and 81% values of the magnesium, calcium and bicarbonate, respectively for the collected groundwater samples in study area. Similarly, magnesium is best parameter that predicts more than 99 and 81% of calcium and bicarbonate, respectively. Total dissolve solid explains more that 99% of the electrical conductivity, bicarbonate estimates more than 81% of calcium, sodium explains 79.3% of chloride, and chloride determines 52.8% of Nitrate concentration.

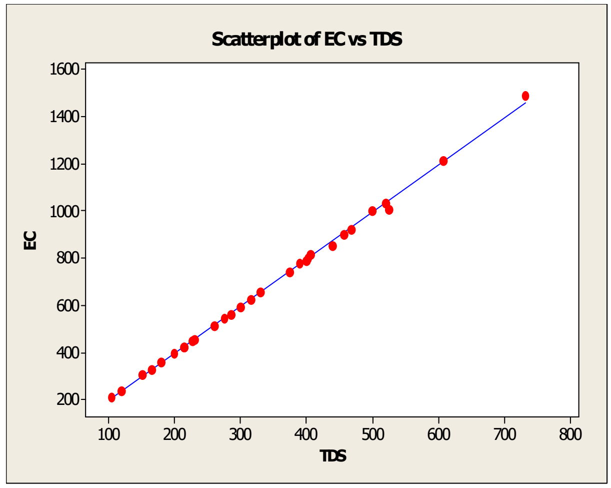

Figure 2: A graph of electrical conductivity and total dissolve solid

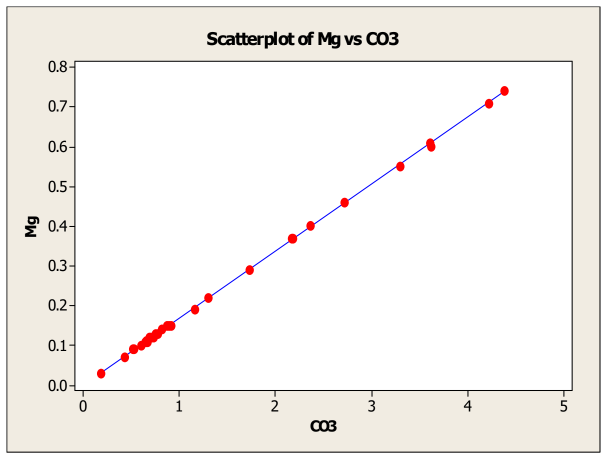

Figure 3: A graph of magnesium and carbonate

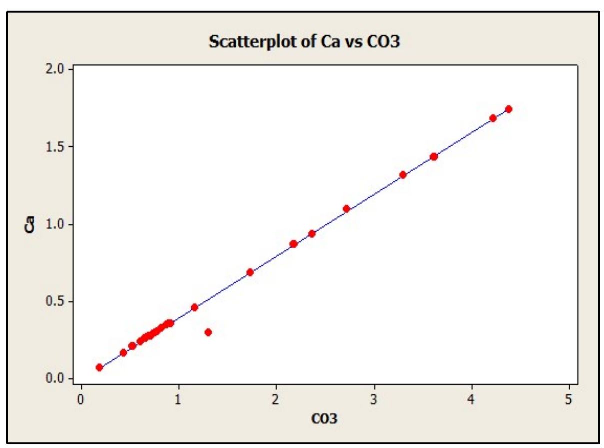

Figure 4: A graph of calcium and carbonate

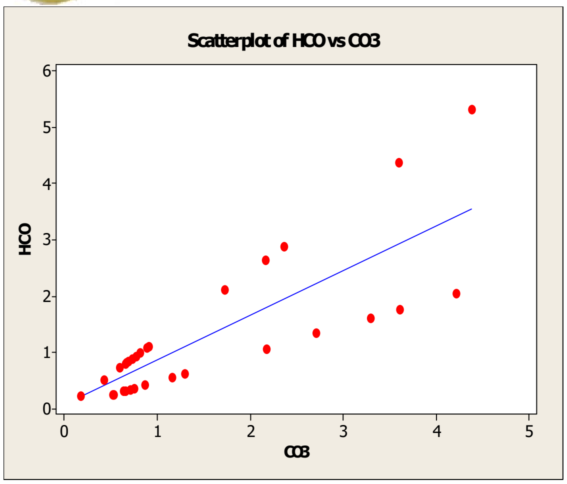

Figure 5: A graph of bicarbonate and carbonate

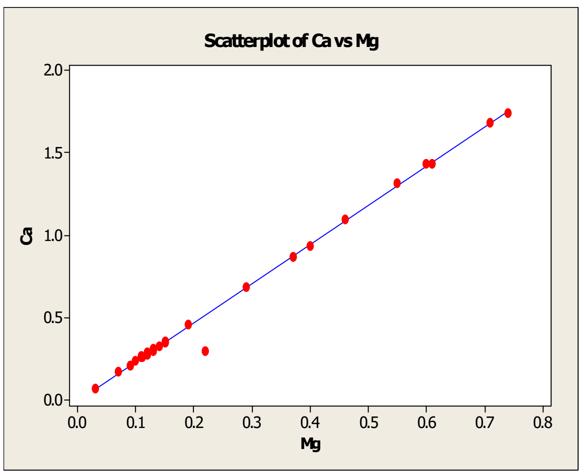

Figure 6: A graph of calcium and magnesium

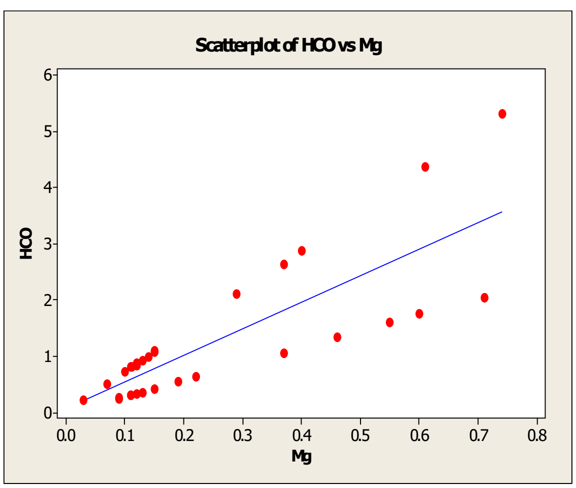

Figure 7: A graph of bicarbonate and magnesium

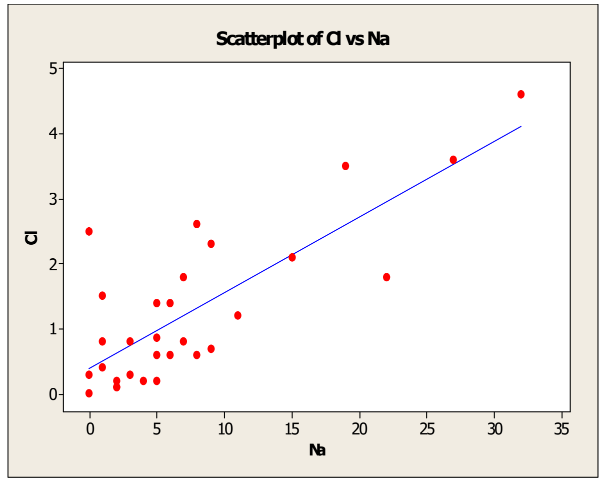

Figure 8: A graph of chloride and sodium

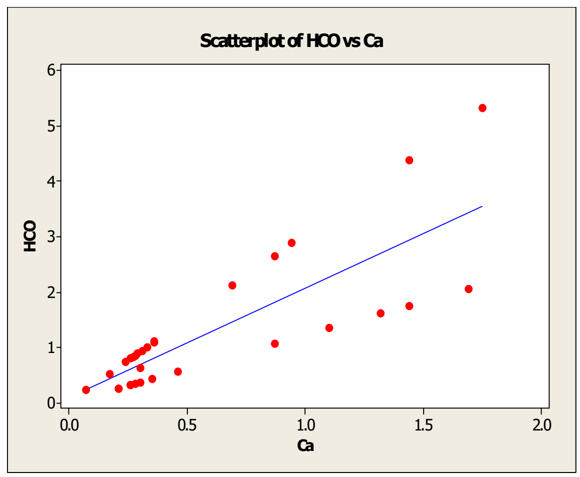

Figure 9: A graph bicarbonate and calcium

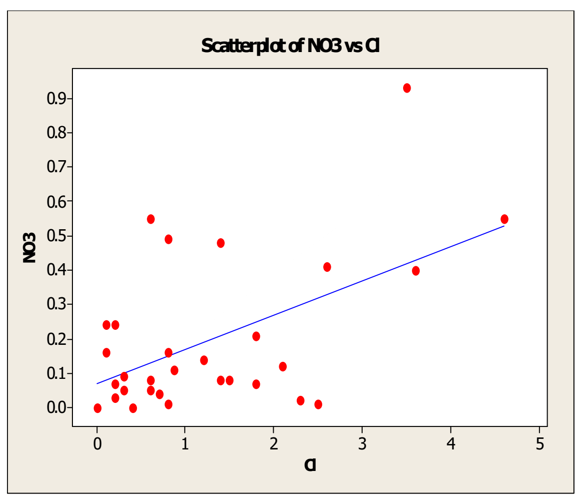

Figure 10: A graph of nitrate and chloride

CONCLUSION

The Jaen district is one of the industrial areas in Nigeria likely suffering from water pollution. A systematic study of linear regression analysis was established to predict the concentration of different parameters, which were found to have a higher level of significance in their correlation coefficient. Highly positive correlation coefficients are observed between EC and TDS; Mg and CO_3; Ca and CO_3; HCO_3 and CO_3; Ca and Mg; HCO_3 and Mg; Cl and Na; Ca and HCO_3; NO_3 and Cl. In most of the locations in our study area, problems with total dissolved solids and bicarbonate were found. All the regression equations computed are significant at 1%. It is evident that carbonate is the best parameter, predicting or explaining 100 and 99 and 81% of the magnesium, calcium, and bicarbonate values, respectively. Magnesium is the best parameter that predicts more than 99 and 81% of calcium and bicarbonate, respectively. Total dissolved solid accounts for more than 99% of electrical conductivity, while bicarbonate accounts for more than 81% of calcium, sodium accounts for 79.3% of chloride, and chloride accounts for 52.8% of nitrate concentration. This kind of study can be applied to other industrial regions of the country to monitor the level of significance of groundwater parameters.

Acknowledgments

This work is a part of the project IBR allocated to Kano University of Science and Technology, Wudil, granted in 2020 by the Tertiary Education Trust Fund (Tetfund). The authors are thankful to the editors and reviewers for their contributions.

REFERENCES

Bhatnagar, A., and Devi, P. (2012). Application of correlation and regression analysis in assessing lentic water quality: a case study at Brahmsarovar Kurukshetr, India. *International Journal of Environmental Sciences*, 3(2): 813-820.

Dalah, C.M., Singh, V.V., Abdullahi, I., and Suleiman, A.A. (2020). The study of HIV/AIDS trend in Yobe state for the prescribed period (1999 – 2019). *Int. J. Stat. Appl,* 10(2020): 10–16.

Dutta, B., Sarma, B. (2018). Correlation study and regression analysis of ground water quality assessment of Nagaon town of Assam, India. *International Journal of Engineering Research & Technology*, 7(6): 320-331.

EPA [Environmental Protection Agency] (2001). Parameters of water quality-interpretation and standards. *Environmental Protection Agency*, Ireland.

Eruola, A.O. and Adedokun, N.A. (2012). Analytical assessment of cadmium, lead and iron in hand dug wells of Ilaro, South Western Nigeria.*Glob J Sci Front Res Chem,* 12(6): 23-29.

Gulgundi, M. S. and Shetty, A. (2018). Groundwater quality assessment of urban Bengaluru using multivariate statistical techniques. *Applied Water Science,* 8(1): 43-55.

Heydari, M.M., Abbasi, A., Rohani, S.M., and Hosseini, M.A. (2013).Correlation study and regression analysis of drinking water quality in Kashan city, Iran. *Walailak J Sci & Tech*, 10(3): 315-324.

Ibrahim, A., Suleiman, A.A., Abdullahi, U.A., and Suleiman, S.A. (2021). Monitaring groundwater quality using probability distribution in Gwale, Kano State, Nigeria. *Journal of Statistical Modeling and Analytics*, 1(2), 99-113.

Joarder, A. M., Raihan, F., Alam, J. B. and Hasanuzzaman, S. (2008). Regression analysis of ground water quality data of Sunamganj District, Bangladesh. *Int. J. Environ. Res.,* 2(3): 291-296.

Juahir, H., Zain, S.M., Yusoff, M.K., Hanidza, T.T., Armi, A.M., Toriman, M.E., and Mokhtar, M. (2011). Spatial water quality assessment of Langat river basin (Malaysia) using environmetric techniques. *Environ Monit Assess,* 173(11): 625–641.

Lee, J.Y., Cheon, J.Y., Lee, K.K., Lee, S.Y., and Lee, M.H. (2001). Statistical evaluation of geochemical parameter distribution in a groundwater system contaminated with petroleum hydrocarbons. *J Environ Qual,* 30(5): 1548–1563.

Mustapha, A. (2015). Evaluation of river water physicochemical properties and environmental consequences of Sharada industrial estate, northwest Nigeria. *Journal of Ecology,* 2(2015): 157-162.

Shrestha, A.K., and Basnet, N. (2018). The correlation and regression analysis of physicochemical parameters of river water for the evaluation of percentage contribution to electrical conductivity. *Journal of Chemistry.*

Shrestha, S. and Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environ Model Softw*, 22(2): 464–475.

Standard Methods for the Examination of Water and wastewater (2005). American Public Health Association/American Water Works Association/water Environment Federation. 21st edn, Washington DC, USA.

Ravikumar, P. and Somashekar, R.K. (2017). Principal component analysis and hydrochemical facies characterization to evaluate groundwater quality in Varahi river basin, Karnataka State, India. *Appl Water Sci,* 7(2): 745–755.

Sadat-Noori, S.M., Ebrahimi, K., and Liaghat, A.M. (2014). Groundwater quality assessment using the water quality index and GIS in Saveh- Nobaran aquifer, Iran. *Environ Earth Sci,* 71(9): 3827–3843.

Salman, S.R. and Ruka'h, Y.A. (1999). Multivariate and principal component statistical analysis of contamination in urban and agricultural soils from north Jordan. *Environ Geol,* 38(3): 265–270.

Suleiman, A.A., Suleiman, A., Abdullahi, U.A., and Suleiman, S.A. (2021). Estimation of the case fatality rate of COVID-19 epidemiological data in Nigeria using statistical regression analysis. *Biosafety and Health,* 3(2021): 4-7.

Suleiman, A.A., Ibrahim, A., and Abdullahi, U.A. (2020a). Statistical explanatory of assessment of groundwater quality in Gwale LGA, Kano state, northwest Nigeria. *Hydrospatial Analysis,* 4(1), 1-13.

Suleiman, A.A., Singh, V.V., Ibrahim, A., Abdullahi, U.A., and Suleiman, S.A. (2020b). Assessment of probability distribution of groundwater quality data in Gwale area, north – western Nigeria. *Annals of Optimization Theory and Practice*. 3(1): 37-46.

Suleiman, A.A. (2015). Analysis of multicollinearity in multiple regressions. *International Journal of Advanced Technology in Engineering and Science,* 3(1), 571 - 578.

Taiwo, A.M., Towolawi, A.T., Olanigan, A.A., Olujimi, O.O. and Arowolo,T.A. (2015). Comparative assessment of groundwater quality in rural and urban areas of Nigeria. *Research and practices in water quality*.