ASSESSMENT OF THE SEASONAL VARIATION IN GROUND WATER QUALITY OF VILLAGES NEAR THE FLOOD PLAINS OF RIVERS NIGER AND BENUE USING MULTIVARIATE STATISTICAL ANALYSIS

SHAFIU, S.[1], PAUL, E.D.[2]*, AND OMONIYI, K.I.[2]
[1]Department of Science Laboratory Technology, Nuhu Bamalli Polytechnic, Zaria, Nigeria
[2]Department of Chemistry Ahmadu Bello University, Zaria, Nigeria

ABSTRACT

The use of different multivariate statistical approaches for the interpretation of the large data obtained for temporal and spatial variations of water quality datasets obtained during the monitoring of ground waters in some villages (Shintaku, Gbobe, Ganaja) near Rivers Niger and Benue and Lokoja metropolis were investigated. The dataset consists of analytical results from a seasonal survey conducted in wells near the Rivers Niger and Benue. Twenty two (22) parameters were monitored on 15 key sampling sites on seasonal basis. To establish the natural and anthropogenic factors or processes accountable for pollution in ground water, the dataset was treated using cluster analysis (CA), Pearson's correlation analysis, principal component analysis (PCA) and factor analysis (FA). Six latent factors were identified by principal component and factor analysis, which are responsible for the data structure explaining 88% and 86% of the total variance of the dataset. Cluster analysis, represented in dendrograms, showed four different groups of similarity between the sampling sites in the wet season and three in the dry season, reflecting the different physicochemical characteristics and pollution levels of the studied water systems. From factor analysis EC, TDS and TS in the dry season and Zn alone in the wet season were the most important parameters contributing to water quality variations in the ground water and positively contributed to water quality variations. It is established from the results that the pollution level and water quality characteristic varied seasonally and with specific ground water systems. A parameter that is most important in contributing to water quality variation for one season may not be important for another season. This study also demonstrates the necessity and effectiveness of multivariate statistical techniques for evaluation and monitoring pollution levels in groundwater.

Key Words: Variance, principal component analysis, factor analysis, cluster analysis physico-chemical parameter, anthropogenic effect.
***Correspondence**: elaoyi@gmail.com

INTRODUCTION

The quality of surface and ground waters is a very delicate issue and with an increased understanding of the importance of drinking water quality to public health and raw water quality to aquatic life, there is a great need to assess and manage surface and ground water quality. These are major issues which were profound impact on human lives, not only for its present use but also from the viewpoint of a potential source of water for future consumption [1-4]. Groundwater is of major importance and is intensively exploited for private, domestic and industrial uses [5]. Water intended for human consumption should be "safe and wholesome" i.e. free from pathogenic agents and harmful chemicals, pleasant to taste and useable for domestic purpose [7]. In this century, about one-third of the world's population depends on groundwater to meet their drinking and irrigation water demands [8].

Environmental pollution can result from natural phenomena, which including geogenic (precipitation rate, weathering processes, soil erosion) or anthropogenic (urban, industrial, agricultural activities and increasing exploitation of water resources) activities. However, river and groundwater interaction can alter water quality and quantity by as a result of anthropogenic activities. [9-12].

Due to their role in carrying off municipal and industrial wastewater and run-off from agricultural land in their vast drainage basins, rivers are among the most vulnerable water bodies to pollution [13]. These sources of contamination may be as a results of surface water runoff generated from urban, rural, and agricultural lands, discharge from ditches and creeks, groundwater seepage from malfunctioning septic tank systems, aquatic weed control, naturally occurring organic inputs, and atmospheric deposition which has been shown also to affect the quality of water [3]. The municipal and industrial wastewater discharge constitutes a constant polluting source, whereas, the surface run-off is a seasonal phenomenon [13, 14]. Since, rivers constitute the main inland water resources for domestic, industrial and irrigation purposes, it is imperative to prevent and control river pollution and to have reliable information on the quality of water for effective management [14].

Through the years, the groundwater monitoring networks in Nigeria have expanded tremendously, and many networks today consist of dozens, if not hundreds of sampling wells [3]. The application of different multivariate statistical techniques such as cluster analysis (CA) and principal component analysis (PCA), factor analysis (FA) offers a better understanding of water quality and ecological status of the studied systems. It allows the identification of the possible factors/sources that influence the water systems and offers a valuable tool for reliable management of water resources as well as rapid solutions of pollution problems [15, 16, 17]. The problems of indicator parameter or import monitoring station identification, data reduction and interpretation, and characteristic change in water quality parameters can be approached through the use of the PCA and FA [3]. Factor analysis (FA), which includes PCA is a very powerful technique applied to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible the variability present in data set. This reduction is achieved by transforming the data set into a new set of variables, the principal components, which are orthogonal (non-correlated) and are arranged in decreasing order of importance [14].

The aim of this study is to determine the nature and spatial distribution of pollutants and evaluate the seasonal correlations of water quality parameters in surface and groundwater in the villages around the Benue Lokoja River basin in Kogi State, Nigeria by applying the PCA and FA techniques. The data matrix obtained during the monitoring program, was subjected to different multivariate statistical techniques to extract information about the similarities or dissimilarities between sampling sites and variables responsible for spatial and temporal variations of water quality.

MATERIALS AND METHODS

Study Area

The study area includes three villages namely Gbobe, Shintaku and Ganaja near Rivers Niger and Benue flood plain located in Lokoja, Kogi State. The south bank of the river has two floodplains and one at the east floodplain. These areas are flooded in the rainy season (May-October) [18]. The other sampling sites are located within Lokoja metropolis, which are distances away from the rivers. Kogi State is found in the north central region of Nigeria; it occupies a land area of 29,833km^2 on 7^o 30'N and 6^o 42'E.

Table 1: Description of sampling sites for well water

S/n	Sample ID	Source	Proximity to the river (m)	Depth of well (m)
1	*K1*	Lokoja	>1000	9.2
2	*K2*	Lokoja	>1000	8.8
3	*K3*	Lokoja	>1000	6.8
4	*K4*	Lokoja	>1000	9.6
5	*K5*	Lokoja	>1000	4.9
6	*K6*	Lokoja	>1000	5.3
7	*K7*	Lokoja	>1000	6.0
8	*S1*	Shintaku	350	2.4
9	*S2*	Shintaku	850	5.2
10	*S3*	Shintaku	1000	5.4
11	*S4*	Shintaku	800	3.6
12	*G1*	Ganaja	650	1.85
13	*G2*	Ganaja	520	4.2
14	*G3*	Ganaja	300	8.4
15	*B1*	Gbobe	700	2.7

Monitoring sites and sampling techniques

In this study, eight well water samples were collected from three villages namely Shitaku, Ganaja, Gbobe near Rivers Niger and Benue and seven from the margin of Lokoja metropolis (300 m away from the Rivers Niger and Benue); these wells were chosen at different distances from the River. The region of sampling covered a wide range of catchments and ground water depths around the River basin. The main pollutant loads into the River include domestic wastewaters, agricultural runoff, animal husbandry and industrial effluents. Using standard analytical methods [18, 19], the samples were analysed for 19 parameters include pH, electrical conductivity, turbidity, total hardness, phosphate, nitrate, ammonia, dissolved oxygen, total suspended solid, total dissolved solid, chemical oxygen demand, biological oxygen demand, copper, nickel, cadmium, manganese, zinc, lead and total bacterial count. The sample bottles (1L) used for collection were washed thoroughly and sterilized with 2% HNO_3 to avoid contamination. The water samples were collected in triplicates and labeled accordingly. Collection of samples was done twice (during the wet and dry seasons) in 2014. A total of 90 samples were collected and analysed. The description of the parameters tested for the two seasons are presented in Table 2. Samples for heavy metals determination were preserved by adjusting the pH to 2 with analytical grade concentrated nitric acid. Samples for bacteriological analysis were preserved in well sterilized sample bottles and stored in ice box at 4^oC to 10^oC. The bacteriological analysis of the water samples were carried out within 24 hr of collecting the samples at the Nigerian Institute of Leather Research Science and Technology, Samaru, Zaria. Parameters with extremely low stability such as pH and temperature were measured in the field using field kits.

The multivariate analysis

Multivariate statistical methods for classification, modelling and interpretation of large datasets from environmental monitoring programs allow the reduction of the dimensionality of the data and the extraction of information that will be helpful for the water quality assessment and the management of surface waters [15]. Multivariate analysis of the well water quality data sets was performed through cluster and factor analysis techniques using Minitab (version 17) statistical software. Correlation structure between the variables was studied using the Pearson's correlation coefficient as a non-parametric measure of the correlation between the variables. In this study, the temporal variations of the ground water quality parameters were evaluated through season-parameter correlation matrix using the Pearson's (non-parametric) correlation coefficient. The water quality parameters were grouped in two different seasons (dry and wet) and each assigned a numerical value in the data file, which as a variable corresponding to the season was correlated (pairwise) with all the measured parameters. Factor and cluster analyses were combined to assess the degree of contamination, determination of chemical processes and to trace the diffusion paths. The statistical analyses were carried out for 22 parameters of 15 sites. Factor analysis was performed on correlation matrix of rearranged data (standardized data through z-scale transfer matrix) for the two seasons. Standardization tends to increase the influence of variables whose variance is small and reduce the influence of variables whose variance is large, eliminates the influence of different units of measurement and renders the data dimensionless [13].

The scree plots and the component loading plots provide a visual to the selection of the principal factors/components used. Principal component analysis/factor analysis was performed on correlation matrix of rearranged data (all observations for each group of sites), so that it explains the structure of the underlying data set. The correlation matrix was used since the units or the variables are greatly different. The correlation coefficient matrix measures how well the variance of each constituent can be explained by relationship with that of the others [14].

Cluster analysis

In hierarchical clustering, clusters are formed sequentially by starting with the most similar pair of objects and forming higher clusters, step by step. Hierarchical agglomerative CA was performed on the normalized data set (mean of observations over the whole period) by means of the Ward's method using squared Euclidean distances as a measure of similarity. The dendrogram provides a visual summary of the clustering processes, presenting a picture of the groups and their proximity [15]. Cluster analysis was applied to the river water quality data set with a view to group similar sampling sites (spatial variability) spread over the river stretch and in the resultant dendrogram, the linkage distance is reported as Similarities as a way to standardize the linkage distance represented on y-axis.

RESULTS AND DISCUSSION

Temporal variations of water quality parameters

Scree plots were used to identify the number of Principal Components (PC) to be retained in order to comprehend the underlying data structure [13]. In this study, the Scree plots are presented in Figures 1 and 2, which showed a pronounced change of slope after the fifth eigenvalue. The first six principal factors have eigenvalues greater than unity (a criterion used to determine the number of factors retained [3, 20]. These PCs have eigenvalues greater than or close to unity and explain 87.9% and 86.1% of the total eigenvalues of information contained in the original dataset for the dry and wet seasons, respectively. Therefore, the first six factors were used for further analysis. In PCA, eigenvalues are normally used to determine the number of PCs that can be retained for further study [3]. Projections of the original variables on the subspace of the PCs are called loadings and coincide with the correlation coefficients between PCs and variables. In other words, the component loadings are the linear combinations for each principal component, and express the correlation between the original variables and the newly formed components.

Table 2: Table of the physicochemical parameters of the wells water samples during the dry season

	L1	L3	L5	L6	L2	L4	L7	S1	S2	S3	S4	G1	G2	G3	B1
pH	6.9	6.38	7	7.5	7.4	6.7	7.4	5.9	6.2	6.7	6.9	6.3	6.8	6.4	7.43
EC uS/cm	676	185.1	167.7	185.1	725	231.95	985	985	957	343	347	941	320	53.5	216
Turbidity NTU	0.134	0.167	53.46	0.167	0.0625	0.088	24.45	6.46	9.51	6.58	5.025	6.05	5.05	7.635	130.45
Hardness mg/L	515.1	217.15	202.005	217.15	479.75	252.5	621.1	817.25	430.25	362.85	815.3	346.25	232.8	474.6	817.85
Phosphate mg/L	8.6	3.05	3.15	3.05	4.3	4.15	3.75	14.95	6.95	9.9	11.85	1.5	1.5	3.6	3.35
Nitrate mg/L	13.35	12.6	11.8	12.6	9.4	9.7	8.5	24.35	9.545	12.3	15	8.9	21.95	3.9	43.35
Ammonia mg/L	0.02	0.03	0.04	0.03	0.02	0.02	0.3	0.02	0.015	0.02	0.025	0.035	0.03	0.25	0.065
TSS mg/L	445	45	115	45	75	60	20	13	76	28	47.5	78	75	20	450
TDS mg/L	310	105	135	105	65	265	195	213.5	76	105	57.5	73	84	905	125
DO mg/L	0.45	0.45	0.4	0.45	0.45	0.25	0.75	0.35	0.55	0.55	0.75	0.55	0.085	0.4	0.12
BOD mg/L	0.1	0.25	0.15	0.25	0.2	0.15	0.15	0.35	0.25	0.33	0.405	0.39	0.035	0.25	0.03
COD mg/L	250	105	290	105	260	295	330	208.5	345.5	185.5	190	240	265	225	315
Cumg/L	0.003	0	0.002	0.003	0	0.001	0.002	0.001	0.003	0.006	0.005	0.002	0.003	0.006	0.003
Nimg/L	0.006	0.002	0.002	0.006	0.015	0.003	0.008	0.012	0.014	0.044	0.039	0.036	0.034	0.03	0.037
Znmg/L	0	0	0	0	0	0	0.01	0.01	0.02	0.024	0.03	0.04	0.01	0.006	0.002
Mnmg/L	0.003	0.004	0.002	0.001	0.009	0.013	0.006	0.012	0.007	0.004	0.006	0.009	0.009	0.008	0.01
Pbmg/L	0.001	0.001	0	0	0.001	0.002	0.002	0.003	0.005	0.006	0.001	0.008	0.009	0.003	0.002
Cdmg/L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TC MPN/100ml	2	3	4.5	2	12	6	7	33.5	12.5	13	21.5	64	26.5	3	31

Table 3: Table of the physicochemical parameters of the wells water samples during the wet season

	L1	L3	L5	L6	L2	L4	L7	S1	S2	S3	S4	G1	G2	G3	B1
pH	5.9	5.7	5.8	6.65	6.55	7.15	5.45	6.4	6.2	7	6.25	6.95	5.8	6.4	5.75
EC uS/cm	896	920	94.3	644	650	47.4	75.1	425	428	289.6	300	97.3	78.9	1398	1206
Turbidity NTU	0.611	0.711	0.712	0.621	0.624	53.5	0.812	100	111	4.11	4.18	140	1.38	1.042	2.83
Hardness mg/L	343.3	232.3	151.5	333.3	292.9	141.4	141.4	303	313.1	22.2	242.4	151.5	151.5	404	373.7
Phosphate mg/L	10.5	3.5	3.2	4.6	4.5	9.5	3.5	13.5	13.2	10.4	10.5	3.5	3.1	12	12.5
Nitrate mg/L	12.5	65	13.5	10.5	7.5	18	11.2	20.5	25	19.5	18.5	10.9	10.5	28.5	28.8
Ammonia mg/L	10.5	3.5	3.2	4.6	4.5	9.5	3.5	13.5	13.2	10.4	10.5	3.5	3.1	12	12.5
TSS mg/L	240	280	230	240	230	260	70	370	300	280	260	0.9	130	240	270
TDS mg/L	850	870	100	750	560	140	130	470	470	390	380	170	100	1560	1550
DO mg/L	0.9	1.2	1.2	1.2	1	1.1	1	1.2	1.1	1.1	1.1	0.9	1.1	0.9	1.1
BOD mg/L	0.4	0.4	0.4	0.4	0.2	0.4	0.3	0.3	0.3	0.6	0.5	0.3	0.3	0.3	0.2
COD mg/L	490	320	380	360	550	260	210	290	220	230	270	240	110	60	290
Cu mg/L	0.008	0.002	0.005	0.007	0.001	0.003	0.004	0.001	0.005	0.008	0.012	0.006	0.008	0.009	0.006
Ni mg/L	0.009	0.006	0.006	0.013	0.022	0.008	0.016	0.022	0.032	0.054	0.041	0.045	0.043	0.08	0.054
Zn mg/L	0	0	0.001	0.002	0.002	0.003	0.05	0.07	0.021	0.032	0.022	0.018	0.012	0.0068	0.0042
Mn mg/L	0.003	0.004	0.002	0.001	0.009	0.013	0.006	0.012	0.007	0.004	0.006	0.009	0.009	0.008	0.01
Pb mg/L	0.003	0.005	0.001	0.002	0.004	0.004	0.006	0.008	0.007	0.012	0.003	0.014	0.013	0.006	0.004
Cd mg/L	0	0	0.001	0	0	0.001	0	0.002	0.001	0.001	0.003	0.001	0.002	0.001	0.004
TC (MPN/100ml)	8	13	14	5	14.5	25	21.5	40.5	17	22	39.5	74	36	5.5	40

The loading plots of the first and second PCs are shown in Fig. 3 and 4. The component loadings can be used to determine the relative importance of a variable (water quality parameter) as compared to other variables in a PC and do not reflect the importance of the component itself [3].

Component loadings of the first two retained PCs for each season are presented in Figures 3 and 4. In the dry season, the PC1 and PC2 explained 23.5 % and 18.7% respectively of the total eigenvalues and was positively and largely contributed by the Chemical Oxygen Demand and was negatively affected by the Biological Oxygen Demand, Dissolved Oxygen, Total Suspended solid, pH, Manganese, Nickel and Ammonia NH_4^+. These components (PC1 and PC2) also reveal that the Total Bacterial Count, Lead and turbidity were less important in accounting for River water quality variations in the dry season since the loading (eigenvector) coefficients were low for these parameters. In the wet season, different component loading patterns were obtained for PC1 and PC2 (Fig. 4). PC1 and PC2 which explained 28.8% and 20.5% of the total eigenvalue was positively infected by the Dissolved Oxygen, Biological Oxygen Demand, electrical conductivity, phosphate and zinc and was negatively affected by Manganese, Nitrate and total hardness

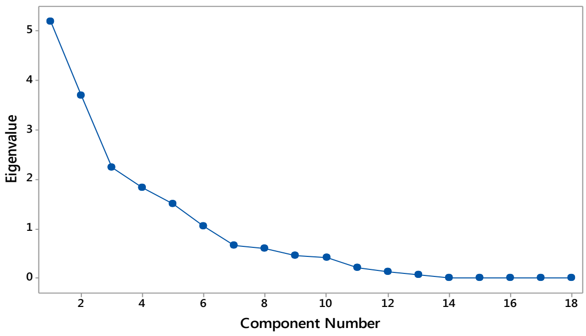

Fig. 1: Scree plot of all dry season components and their eigenvalues

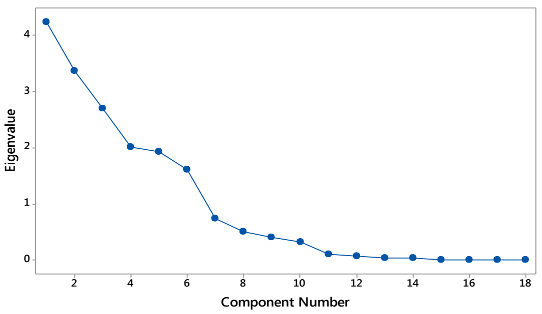

Fig. 2: Screeplot of all wet season components and their eigenvalues

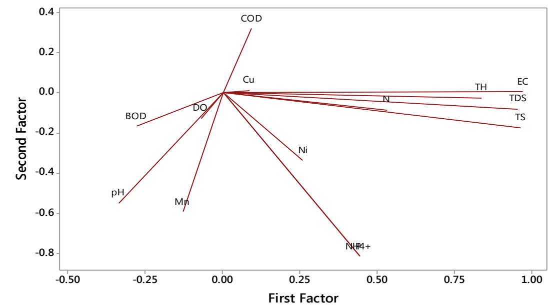

Fig. 3: Loadings plots for the first component (PC1) and the second component (PC2) in dry season

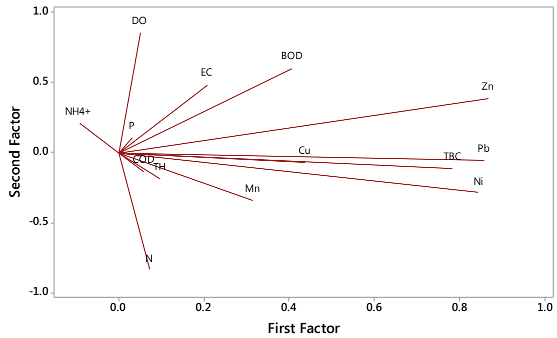

Fig. 4: Loadings plots for the first component (PC1) and the second component (PC2) in wet season

Seasonal Correlation of water quality parameters

Data in Table 6 provide the seasonal correlation matrix of the water quality parameters obtained from the PCA. In the dry season, the three most important water quality parameters were Total Solid, Total Dissolved Solid and Electrical conductivity as seen in the PCA results. The correlation coefficients between Total Solid and other parameters were less than or equal to 0.54 for the dry season except for Total Dissolved Solid (0.99), Electrical Conductivity (0.98) and Total Hardness (0.79). These correlations changed slightly in the wet season with a negative change in correlations in most of the parameters with Total Dissolved Solid (0.83), Electrical conductivity (0.26), and Total Hardness (0.21) as well as a positive increase in correlations with Turbidity (0.26), Chemical Oxygen Demand (0.18), Cupper (0.34) and Manganese (0.07). The negative increase in correlations with parameters such as Total Dissolved Solid, Electrical Conductivity and Total Hardness was as a result of increased volume of groundwater across wells, causing a dilution effect during the wet season. In the wet season, the correlation coefficients between Total Hardness and other parameters were relatively weaker and fell below 0.34. Interestingly, a good correlation was observed (0.83) between Total Solid and Total Dissolved Solid. This is because the total solid in the groundwater is dependent on the total dissolved solid, irrespective of the season.

From the results of the PCA (Fig. 4), the most important parameters in the wet season were Zinc, Biological Oxygen Demand, Electrical Conductivity and Dissolved Oxygen. The correlation coefficient between

Zinc and the other parameters were less than or equal to 0.67 with Biological Oxygen Demand and Dissolved Oxygen being the highest values at 0.67 and 0.66 respectively. The values of the correlation coefficient of Zinc in the dry season were relatively weaker with all values less than 0.39 and almost no correlation with neither Biological Oxygen Demand nor Nickel.

The correlation Tables showed that Total Hardness had relatively strong correlations with Electrical conductivity (0.63) and Manganese (0.60) in the dry season and also Nickel (0.68) and lead (0.62) in the wet season. Similar strong correlations were found between Total Hardness and Electrical Conductivity (0.8) and Total Dissolved Solid (0.78) in the dry season only.

Spatial similarities and site grouping

The relationships between the sites sampled are obtained through cluster analyses using Word's method (linkage between groups), Euclidian distance as similarity measure and synthesized in dendrograms. The dendrograms are performed separately for all sites, at different seasons for better assessment regarding the level of contamination and showed the result sequence in the association, displaying the information as degree of contamination. Cluster analysis rendered dendrograms (Fig. 5 and 6), where all the 15 sampling sites on the groundwater were grouped into three statistically significant clusters at similarity>90. The dendrograms indicate the status of pollution as well as the effect of contamination at the sites, In the dry season, the clustering procedure generated three groups of sites, as the sites in these groups have similar characteristic features and natural background source types as showed in PCs Tables.

Cluster 1 (Sites K1, K6, K3, G3 and B1)), cluster 2 (Sites K2, S1, S4 and S3) and cluster 3 (Sites KK5, K7, K4, G2 and G1) correspond to a relatively low pollution, moderate pollution and very high pollution and regions respectively. It implies that for rapid assessment of water quality, only one site in each cluster must be needed in spatial assessment of the water quality of the whole network. It is evident that the CA technique is useful in offering reliable classification of surface waters in the whole region and will make possible the design of a future spatial sampling strategy.

In the wet season, the degree of contamination is clearly reflected among the groups. The clustering procedure generated four major groups of sites, as the sites in groups have similar characteristics features (at similarity > 59.2) as shown in (Fig. 6) and natural background source types. In the wet season, the first group (K1, K2, K6, S3, G2, K5 and K4) may be considered as natural background type or less polluted group of stations and the second group may be related to slightly polluted stations, however the third group (K2, K7, S1, S2 and G1) as highly polluted group of stations, which are highly influenced by domestic wastes and other anthropogenic activities. This last group (G3) is the most polluted site in the wet season, which may be as a result of depth of the well in site G3.

Table 4: Seasonal correlation of water quality parameters in wet season

	pH	*EC*	*Turbi*	*TH*	*P*	*N*	*NH4+*	*TDS*	*TS*	*DO*	*BOD*	*COD*	*Cu*	*Ni*	*Zn*	*Mn*	*Pb*	*Cd*	*TBC*
pH	1																		
EC	-0.26	1																	
Turbi	0.38	-0.22	1																
TH	0.02	0.35	0.37	1															
P	-0.37	0.31	-0.20	0.61	1														
N	0.19	-0.12	0.75	0.49	0.12	1													
NH_3	0.18	0.04	0.12	0.20	-0.25	-0.25	1												
TDS	-0.22	-0.28	-0.09	0.06	-0.04	-0.29	0.60	1											
TS	0.00	-0.28	0.26	0.21	-0.07	0.08	0.41	0.83	1										
DO	0.00	0.41	-0.38	0.19	0.31	-0.58	0.27	-0.11	-0.29	1									
BOD	-0.50	0.25	-0.48	0.15	0.49	-0.39	-0.14	-0.06	-0.37	0.63	1								
COD	0.10	0.38	0.42	0.22	-0.15	0.13	0.25	0.02	0.18	-0.14	-0.49	1							
Cu	-0.01	-0.32	0.04	0.17	0.22	-0.05	0.24	0.40	0.34	0.19	0.20	-0.08	1						
Ni	-0.06	-0.09	0.21	0.33	0.11	0.30	0.00	0.00	0.02	-0.02	0.24	0.00	0.66	1					
Zn	-0.39	0.40	-0.19	0.20	0.24	-0.15	-0.08	-0.24	-0.37	0.49	0.67	0.01	0.39	0.66	1				
Mn	-0.33	0.25	0.10	0.35	0.08	0.29	0.01	0.12	0.07	-0.41	-0.04	0.42	-0.25	0.22	0.11	1			
Pb	-0.44	0.27	-0.16	-0.17	-0.13	0.03	-0.08	-0.10	-0.19	-0.19	0.09	0.19	0.24	0.60	0.62	0.35	1		
Cd	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1	
TBC	-0.31	0.42	0.17	0.28	0.03	0.35	-0.22	-0.31	-0.24	-0.09	0.32	0.12	-0.06	0.56	0.68	0.49	0.62	0.00	1.0

Table 5: Seasonal correlation of water quality parameters in dry season

	pH	***EC***	***Turbi***	***TH***	***P***	***N***	***NH3***	***TDS***	***TS***	***DO***	***BOD***	***COD***	***Cu***	***Ni***	***Zn***	***Mn***	***Pb***	***Cd***	***TBC***
pH	1																		
EC	-0.20	1																	
Turbi	0.42	-0.33	1																
TH	-0.21	0.80	-0.05	1															
P	0.20	0.40	0.23	0.44	1														
N	-0.26	0.48	-0.08	0.17	0.13	1													
NH3	0.20	0.40	0.23	0.44	1	0.13	1												
TDS	-0.16	0.98	-0.30	0.77	0.46	0.45	0.46	1											
TS	-0.14	0.97	-0.29	0.78	0.54	0.48	0.54	0.99	1										
DO	-0.11	-0.21	-0.09	-0.15	-0.06	0.31	-0.06	-0.19	-0.08	1									
BOD	0.29	-0.28	-0.21	-0.51	0.06	0.09	0.06	-0.28	-0.22	0.25	1								
COD	-0.01	0.08	-0.17	0.18	-0.17	-0.16	-0.17	-0.02	0.01	0.08	-0.04	1							
Cu	0.01	0.09	-0.28	0.03	0.19	-0.20	0.19	0.16	0.12	-0.28	0.44	-0.37	1						
Ni	0.19	0.35	0.03	0.15	0.40	0.00	0.40	0.44	0.39	-0.41	-0.09	-0.63	0.54	1					
Zn	0.00	-0.35	0.39	-0.22	0.27	-0.15	0.27	-0.32	-0.28	0.12	0.03	-0.29	-0.19	0.08	1				
Mn	0.28	-0.08	0.47	0.05	0.54	-0.06	0.34	-0.02	-0.02	-0.15	-0.51	-0.29	-0.34	0.27	0.29	1			
Pb	0.27	-0.33	0.47	-0.46	-0.07	-0.09	-0.07	-0.29	-0.35	-0.25	0.00	-0.54	0.07	0.48	0.39	0.33	1		
Cd	-0.09	0.07	0.07	0.13	0.47	0.03	0.47	0.21	0.23	0.19	-0.12	-0.33	0.34	0.50	0.15	0.44	0.10	1	
TBC	0.21	-0.40	0.60	-0.29	-0.03	-0.17	-0.03	-0.30	-0.36	-0.14	-0.16	-0.25	0.05	0.28	0.35	0.52	0.60	0.54	1

Table 6: Rotated factor loadings and communalities (Varimax rotation)

Dry Season Variables						**Wet Season Variables**					
Variable	Factor1	Factor2	Factor3	Factor4	Communality	Variable	Factor1	Factor2	Factor3	Factor4	Communality
pH	-0.337	-0.549	0.123	0.004	0.43	EC	0.208	0.484	-0.34	-0.691	0.87
EC	0.971	0.004	0.105	-0.055	**0.956**	TH	0.097	-0.19	-0.887	-0.209	0.876
TH	0.837	-0.028	0.017	-0.303	0.793	Phos	0.032	0.108	-0.866	0.113	0.776
Phos	0.442	-0.817	0.118	0.042	0.878	Nitr	0.072	-0.838	-0.342	-0.151	0.847
Nitr	0.531	-0.087	-0.355	0.276	0.492	NH4+	-0.092	0.206	0.06	-0.068	0.059
NH4+	0.442	-0.817	0.118	0.042	0.878	DO	0.05	0.856	-0.346	0.179	0.887
TDS	0.956	-0.081	0.176	-0.05	**0.954**	BOD	0.407	0.603	-0.414	0.285	0.781
TSS	0.964	-0.174	0.075	0.005	**0.965**	COD	0.059	-0.134	0.057	-0.692	0.504
DO	-0.07	-0.129	-0.69	0.387	0.648	Cu	0.437	-0.07	-0.156	0.65	0.643
BOD	-0.279	-0.168	0.011	0.905	0.924	Ni	0.844	-0.286	-0.183	0.26	0.895
COD	0.093	0.319	-0.551	-0.088	0.422	Zn	0.868	0.384	-0.225	0.032	**0.952**
Cu	0.087	0.011	0.762	0.529	0.868	Mn	0.315	-0.339	-0.151	-0.641	0.647
Ni	0.259	-0.338	0.799	-0.039	0.822	Pb	0.857	-0.055	0.291	-0.202	0.863
Mn	-0.129	-0.593	0.039	-0.702	0.863	TBC	0.783	-0.111	-0.157	-0.379	0.794
Variance	4.4541	2.2925	2.2181	1.9273	10.892	Variance	3.342	2.5054	2.3088	2.2389	10.395
% Var	0.318	0.164	0.158	0.138	0.778	% Var	0.239	0.179	0.165	0.16	0.742

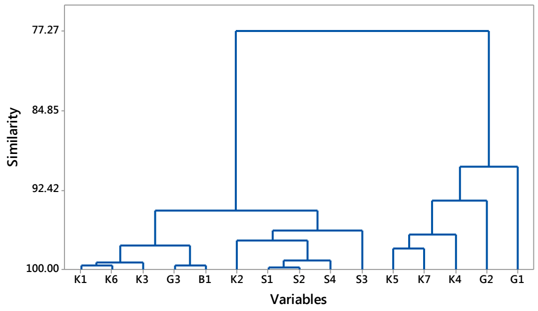

Fig. 5: Dendrogram of site similarities in dry season (Complete linkage, correlation coefficient distance)

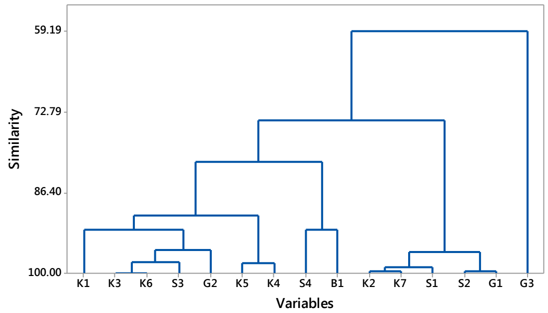

Fig. 6: Dendrogram of site similarities in wet season (Complete linkage, correlation coefficient distance)

Identification of important water quality parameters

Figures 3 and 4 show that PC1 and PC2 for both seasons were highly influenced (negatively or positively) by most of the variables, thus hindering the interpretation regarding which parameters are more important than the others in influencing water quality variations within a given season. Therefore, the PFA has been employed to circumvent the ambiguity in the data. The results of the PFA are presented in Tables 6 for the dry and wet seasons respectively showing the Varimax rotated factor loading correlation coefficients for the first four factors in both season. The reason to retain the first four factors for analysis is that these four factors account for 77.8%, and 74.28% of the total variances in both seasons. The rest of the 11 factors accounted for only small percentages of the total variances and had very low and insignificant correlation coefficients. The most important water quality parameters that can be used to evaluate seasonal variations of the well water quality are highlighted in Table 3. Hence any water quality parameter with an absolute factor correlation coefficient value greater than 0.95 was considered to be a most important parameter in contribution to variations in water quality. From the values of communality in the dry season, Electrical conductivity, Total Dissolved Solid and Total Solid are the most important parameters in contribution to water quality variations in the ground water. These parameters may be interpreted as representing influences from natural inputs. In the wet season, Zinc alone was identified as the most important parameters and positively contributed to water quality variations. The inorganic nutrients may be interpreted as representing influences from anthropogenic inputs in the wet season. Therefore, a quality parameter that is important in contribution to water quality variation for one season may not be important for another season [5].

CONCLUSION

The application of multivariate statistical techniques was used to establish the nature and spatial distribution of pollutants originated by natural and anthropogenic activities in the well waters of selected villages near the River Niger and Benue areas of Lokoja metropolis kogi state, Nigeria. The verifactors (Total Hardness, Total Dissolved Solid, Total Solid and Zinc) obtained from FA for the study sites, indicated that the genesis for contaminants responsible in the water quality variations for both wet and dry seasons are anthropogenic i.e. (industrial, domestic and agricultural activities), these are mainly related to trace metals leaching from the soil. .Although, factor analysis/principle component analysis did not resulted in a significant data reduction, but it helped to extract and identify the most important parameters responsible for the variations of the quality of groundwater at three different sampling sites. Therefore water quality design for monitoring network/strategy for the effective management of water resources in the villages is recommended for better surface water and groundwater.

REFERENCES

1. OMO-IRABOR, O.O., BAMIDELE, S., ODUYEMI, K., & AKUNNA, J. (2008). Surface and groundwater water quality assessment using multivariate analytical methods : A case study of the Western Niger Delta , Nigeria, **33**(8):666-673. http://doi.org/10.1016/j.pce.2008.06.019.
2. KORI, R., SAXENA, A. & UPADHAYAY, N. (2006). Groundwater Quality Assessment of Mandideep Industrial Area. National Serminar on Environmental and Development, Bhopal, India, p. 155.
3. OUYANG, Y. (2005). Evaluation of river water quality monitoring stations by principal component analysis. *Water Research*, **39**: 2621-2635. http://doi.org/10.1016/j.watres.2005.04.024.
4. BENGRAINE, K. & MARHABA, T.F. (2003). Using principal component analysis to monitor spatial and temporal changes in water quality. *Journal of Hazardous Materials*, **100**(1): 179-195. http://doi.org/10.1016/S0304-3894(03)00104-3
5. OUYANG, Y., NKEDI-KIZZA, P., WU, Q.T., SHINDE, D., & HUANG, C.H. (2006). Assessment of seasonal variations in surface water quality. *Water Research*, **40**: 3800-3810. doi:10.1016/j.watres.2006.08.030.
6. JINWAL, A. & DIXIT, S. (2008). Pre- and Post-Monsoon Variation in Physico-Chemical Characteristics in Groundwater Quality of Bhopal, India. *Asian Journal of Experimental Science* **22**(3): 311-316.

7. PARASHAR, C., DIXIT, S. & SRIVASTAVA, R. (2006). Seasanal variations in physico-chemical characteristics in upper lake of Bhopal, Asian. *Journal of Experimental Science*, **20**(2): 297-302.
8. ALLEY, W.M., HEALY, R.W., LABAUGH, J.W. & REILLY, T.E. (2002) Flow and storage in groundwate systems. *Science***, 296**: 1985-1990.
9. YUCE, G., & ALPTEKIN, C. (2013). *In situ* and laboratory treatment tests for lowering of excess manganese and iron in drinking water sourced from river-groundwater interaction. *Environmental Earth Sciences*, **70**(6): 2827-2837. http://doi.org/10.1007/s12665-013-2343-x.
10. EDET, A.E., WORDEN, R.H., MOHAMMED, E.A., PRESTON, M.R. (2012). Hydrogeochemical processes in a shallow coastal plain sand aquifer and tidal river systems (Calabar, Southeastern Nigeria): tracking wastewater and seawater pollution in ground and river waters. *Journal of Environmental Earth Science*, **65**(7): 1933-1953.
11. BORRELLI, N., OSTERRIETH, M., ROMANELLI, A., ALVAREZ M.F., CIONCHI, J.L., MASSONE, H. (2012). Biogenic silica in wetlands and their relationship with soil and groundwater biogeochemistry in the Southeastern of Buenos Aires Province. *Argentina. Environmental Earth Science*, **65:** 469-480.
12. GARCIA, M.G., LECOMTE, K..L., STUPAR Y, FORMICA, S.M., BARRIONUEVO, M., VESCO, M., GALLARA, R. & PONCE, R. (2012). Geochemistry and health aspects of F-rich mountainous streams and groundwaters from sierras Pampeanas de Cordoba. *Argentina. Environmental Earth Science*, **65:** 535-545.
13. SINGH, K.P., MALIK, A., & SINHA, S. (2005). Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques - a case study. *Analytica Chemica Acta*, **538**: 355-374. doi:10.1016/j.aca.2005.02.006
14. CHANDRA, U., KUMAR, S., RATH, P., BIHARI, B., & BHATTA, D. (2006). Application of factor and cluster analysis for characterization of river and estuarine water systems – A case study : Mahanadi River (India) *Argentina. Environmental Earth Science*, 434–445. doi:10.1016/j.jhydrol.2006.05.029
15. SHRESTHA, S., & KAZAMA, F. (2007). Assessment of surface water quality using multivariate statistical techniques : A case study of the Fuji river basin , *Japan. Environmental Modelling and Software*, **22**: 464-475. doi:10.1016/j.envsoft. 2006.02.001
16. SIMEONOV, V., STRATIS, J.A., SAMARA, C., ZACHARIADIS, G., VOUTSA, D., ANTHEMIDIS, A. & KOUIMTZIS, T. (2003). Assessment of the surface water quality in Northern Greece. *Water Research*, **37**: 4119-4124. 1/S0043-1354(03)00398-1
17. REGHUNATH, R, MURTHY., T.R.J. & RAGHAVAN, B.R.. (2002). The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. *Water Resources*, **36:** 2437-2442**.**
18. M. O. ISIKWUE, D. IORVER AND S. B. ONOJA (2011). Effect of Depth on Microbial Pollution of Shallow Wells in Makurdi Metropolis, Benue State, Nigeria. *British Journal of Environment & Climate Change,* 1(3): 66 73.
19. APHA, AWWA (2000). Standard Method for examination of Water and Wastewater investigation, Washington DC, USA, pp. 12-16.
20. YU, J.C., QUINN, J.T., DUFOURNAUD, C.M., HARRINGTON, J.J., ROGER, P.P. & LOHANI, B.N., (1998). Effective dimensionality of environmental indicators: a principal component analysis with bootstrap confidence intervals. *Journal of Environmental Management*, **53,** 101-111.